SWAN: System for Wearable Audio Navigation

Jeff Wilson[1], Bruce N. Walker[2,3], Jeffrey Lindsay[2], Craig Cambias[3], and Frank Dellaert[3]
[1]*Biomedical Interactive Technology Center,* [2]*School of Psychology, and* [3]*School of Interactive Computing - Georgia Institute of Technology*
jeff.wilson@bitc.gatech.edu , bruce.walker@psych.gatech.edu

Abstract

Wearable computers can certainly support audio-only presentation of information; a visual interface need not be present for effective user interaction. A System for Wearable Audio Navigation (SWAN) is being developed to serve as a navigation and orientation aid for persons temporarily or permanently visually impaired. SWAN is a wearable computer consisting of audio-only output and tactile input via a handheld interface. SWAN aids a user in safe pedestrian navigation and includes the ability for the user to author new GIS data relevant to their needs of wayfinding, obstacle avoidance, and situational awareness support. Emphasis is placed on representing pertinent data with non-speech sounds through a process of sonification. SWAN relies on a Geographic Information System (GIS) infrastructure for supporting geocoding and spatialization of data. Furthermore, SWAN utilizes novel tracking technology.

1. Introduction

There is a critical need for navigation and orientation aids for the visually impaired. This need applies to individuals who have suffered physical vision loss (e.g. full or partial blindness) as well as those affected by temporary loss of vision such as firefighters in a smoke-filled environment. For a person with vision loss, the two fundamental tasks of navigating through a space and knowing what is around her can be a great challenge. It is therefore highly important to develop a system that communicates a range of information about the environment in a non-visual manner, to allow a person greater knowledge, connection to, and more effective navigation through the space. Of the candidate alternative display modalities, audition is the obvious choice because of the excellent human ability to localize complex environmental sounds.

There are approximately 11.4 million people with vision loss in the United States, 10% of whom have no usable vision; and by 2010 these numbers will nearly double [1-3]. Many can remain very productive even with diminished eyesight, so long as they are able to get to and from work, and move about the office safely and effectively. Spatial orientation is the major mobility problem encountered by all individuals with profound vision loss [4, 5], but is especially difficult for people whose vision loss occurs later in life [5, 6].

Wayfinding (the ability to find one's way to a destination) is dependent on the ability to remain oriented in the environment in terms of the current location, heading, and direction of a destination. Even experienced blind pedestrians exhibit movement error large enough to occasionally veer into a wall or into a parallel street when crossing an intersection [7]. Evidence suggests this behavior may be due to slight deviations in step direction that accrue over time and which blind pedestrians may lack a reliable orienting stimulus to correct [8]. These problems persist when the person is indoors. While there has been a great deal of research in the area of electronic travel aids for obstacle avoidance, there has not been comparable research in the development of orientation devices that keep one apprised of location and heading [5].

In addition to visually impaired workers, there are many situations where sighted individuals cannot use vision for navigation. In some cases, such as a firefighter in a smoke-filled building, or soldier during a night mission, vision may not be useful due to environmental constraints (e.g., smoke or darkness). Individuals in such situations can be faced with serious consequences should they become disoriented. Sonification of navigation data can help these individuals find their way, while being aware of there environment at all times.

1.1 Previous work

To date, the primary presentation modality of existing path planning auditory displays has been synthesized speech that is used to speak instructions to the user. The Personal Guidance System (PGS) [9, 10] is typical of the modern devices: the computer creates words that seem to come from the same place as the object or feature to which they refer via virtual speech beacons. "Doorway here" would sound as if it came from the real doorway. See, also, the Mobility of Blind and Elderly People Interacting with Computers (MoBIC) system [11]. The Drishti system [12, 13] utilizes the ESRI ArcSDE spatial database engine, voice recognition, and synthetic speech.

There are two commercially available navigation systems of note for the visually impaired. One is the Humanware Trekker and the other is the Sendero BrailleNote GPS. Both support basic navigation tasks via speech presentation and are intended as supplements to a user's existing obstacle avoidance techniques. Both have path planning and path recording modes and virtual map browsing.

There are several drawbacks to using speech sounds in this way. Speech beacons are harder to localize in a virtual environment than non-speech beacons [14]. Users also give speech beacons low ratings for quality and acceptance [14]. The speech-based interface cannot display a large amount of information, as two or more speech beacons presented simultaneously are difficult to attend to [e.g., 15]. It is also difficult to use a speech-based interface for navigation and carry on a conversation at the same time [see, e.g., 16, 17]. Further, spoken messages in such a system are each generally more than a second long, so the system is often talking. For occasional spoken directions (e.g., "Turn left"), this is not a major issue. However, if the system is simultaneously presenting other sounds representing the upcoming curb cut, a low hanging branch, etc., the inherent inefficiency of speech can result in a cluttered listening environment. Recent work has also shown that using non-speech cues leads to better performance than speech cues in situations where there is a significant cognitive load such as simultaneous tasks [18]. One possible alternative may be spearcons, [19] compressed speech stimuli that have shown superior performance to artificial speech, auditory icons [20] and earcons [21].

While it is true that presenting a number of non-speech sounds around the user could also lead to a busy listening experience, the acoustic flexibility and brevity of non-speech sounds provides the designer with considerably more control. An immediately recognizable sound similar to trickling water could be an aesthetic and effective means of indicating the location of a fountain, without speaking "Drinking Fountain" aloud. One final concern with spoken navigation commands that are not spatialized is that it simply takes many words to describe non-rectilinear movement: a 20 degree turn must be described as, "Veer to the left," or "Turn a little bit to the left." In our experience, simply walking toward a beacon sound is easier than translating "57 degrees" into a movement action. Thus, while speech-based navigation sounds have been useful in some cases, there is a need to understand how to utilize non-speech sounds as well, and to implement them into a practical system.

2. SWAN implementation

The System for Wearable Audio Navigation (SWAN) addresses the limitations of previous speech-based navigation aids by using non-speech audio presentation of navigation information whenever possible. SWAN provides an auditory display that enhances the user's ability to (1) keep track of her current location and heading as she moves about, (2) find her way around and through a variety of environments, (3) successfully find and follow a near optimal and safe walking path to her destination, and (4) be aware of salient features of her environment.

SWAN supports these goals through sophisticated position tracking technologies, sonification of navigation routes and environmental features, and implementation of a database of information relevant to the user's navigation needs. SWAN allows users to record their movements or paths through the environment. These paths are used to create a personally relevant set of maps for the user. Additionally, the user can annotate objects found within the environment including locations, features, and obstacles. This could include, for example, a particular bus stop, a favorite coffee shop, or a section of sidewalk prone to flooding after rain showers. The map can also be queried for directions to a particular location. To indicate paths and features in the environment the user is presented with a set of non-speech audio cues that guide them along the paths, all the while sonifying nearby features and obstacles.

SWAN has been developed in two phases. The initial phase focused on creating a wearable system that supports only one user, and allows that user to utilize a personal database of paths and features. The current (second) phase serves as a mobile wireless platform for sharing information among many users via a remote GIS database. Furthermore, SWAN now serves as a platform for experimenting with novel

tracking technologies, thus the system is constantly undergoing changes and upgrades. The following discussion represents the full range of capabilities available in the SWAN project; any specific system build may include all or a subset of features.

2.1. Overview of hardware

SWAN consists of a portable computer, an audio processor, audio output hardware, tactile input devices, and position/orientation tracking. All the components are stored in a backpack, allowing for mobile use (see Fig. 1). A wireless Internet connection allows for constant connectivity.

2.1.1. Base system. The core SWAN system currently consists of a Sony Vaio UX280P micro-PC. The Vaio has 1 GB of RAM, a 40 GB hard drive, and utilizes a 1.2 GHz Intel Core Solo Processor (U1400). Also included is 802.11a/b/g and Verizon EDGE support. The Vaio runs WindowXP and the SWAN application is built upon the Windows .NET platform.

2.1.2. Position tracking hardware. The existing SWAN system combines multiple GPS receivers with either a digital compass (PNI, Inc., model TCM3) or an Intersense InertiaCube2 orientation sensor to provide a reasonable estimate of location and accurate head orientation. The iCube2 is sampled at approximately 30 Hz, and makes use of an internal Kalman filter that allows it to merge orientation data from multiple MEMS sensors. The compass supports a similar update rate. GPSs utilizing the SirfStar III chipset are employed for global positioning. SirfStar III is capable of 1Hz position updates and also supports the Wide Area Augmentation System (WAAS). Additional sensors, such as light meters and thermometers, can also be added to the sensor package. A multi-camera computer vision system is also being developed, and that system will enable indoor and outdoor tracking, without having to instrument the environment with RFID tags, visual fiducials, etc. The sensor fusion software component, called *Merge*, can support a range of sensor inputs.

2.1.3. Audio hardware. Audio generation hardware consists of the generic onboard sound chipset built into each mobile computer's main board. Additionally, an external USB sound card may be used.

The auditory display makes use of 3D spatialized audio cues. Presenting sounds that seem to come from arbitrary spatial locations around the user (and seemingly external to the head) requires at least two speakers and special software (e.g., Open Audio Language/OpenAL) to mimic the timing and spectral changes that cue sound location. Audio presentation can be accomplished by way of headphones, which also serve as the physical mount point for the iCube2 tracker or compass. Unfortunately, while headphones are usually quite effective for spatialized audio, they block out important ambient sounds and thus are inappropriate for this specific audio navigation application. That is, end users refuse to wear headphones that block out the ambient sounds. In the past, other researchers have suggested a number of potential solutions to this critical problem, including shoulder-mounted speakers [22], speakers suspended away from the ear [23], and even using "audio spotlights" to target the ears from some distance [24].

None of these suggestions have met all the criteria for a successful display, especially in terms of user acceptability. By contrast, the SWAN team is beginning to address this issue successfully through the use of bone conduction headphones (or bonephones). Stereo bone conduction headphones are used to present stereo sound to the listener without obstructing their regular perception of environmental sounds. The bonephones transmit vibrations through the skull into the inner ear, resulting in the perception of regular sound. The primary advantage of this technique is that external sound sources are not occluded, as is the case with headphones. Extensive psychophysical assessment of the bonephones' spectral and temporal characteristics has been conducted [25]. Also there is ongoing work in the determination of any "bone-related transfer function" (BRTF) or cross-talk cancellation that may need to be implemented [26]. This solution provides excellent flexibility in sound design, yet is extremely discreet.
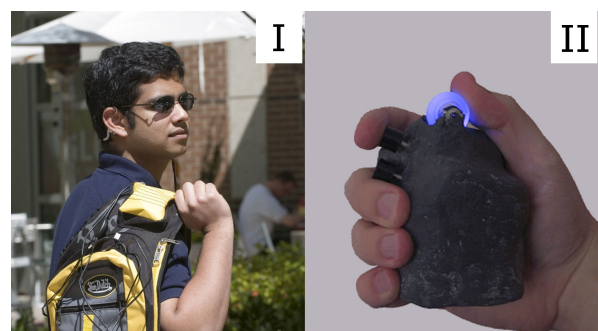

Figure 1. (I.) A SWAN developer is shown testing the SWAN system. (II.) Manual controller for the audio menu interface is shown in use.

2.1.4. Input hardware. A small microphone is used for recording voice annotations, however the bulk of input to the system is via physical handheld devices. In order to navigate the SWAN system's audio menus, a custom hardware input device has been developed for tactile interactions with SWAN (see Fig. 1). Using the thumb-wheel and two buttons, users can scroll up and down through the audio menu choices and then go forward or backward through different sets of choices. The device is derived from the electronics found in a computer mouse. An upper layer mouse filter device driver was written to direct input from the device to SWAN and not to the OS user interface.

2.2. Tracking via sensor fusion

The tracking framework in SWAN, called Merge, provides pose information consisting of location and orientation (i.e., the x, y, z location and roll, pitch, and yaw of the user's head). Merge is named after its task of merging data from multiple sensors into a single pose estimate, otherwise known as sensor fusion.

To estimate the most probable pose given information from a variety of sensors we use a Bayes filter approach. Bayesian probability theory is the theoretical framework in which to properly account for different sensor accuracies. Because some of the sensors we use cannot be well represented using Gaussian uncertainty models, we use a particle filter method known as Monte Carlo Localization [27, 28].

The particle filter central to Merge represents the probability density over poses using a set of samples or particles, which are transformed at each time-step in a two-phase process. First, Merge considers the movement of the user by randomly displacing each pose according to a motion model. The model represents typical walking speed. This step increases uncertainty, reflected by an expanding *cloud* of particles, each representing a hypothesized pose. Uncertainty is then decreases in the second step by weighting each pose by its likelihood given the sensor measurements.

Merge is designed to allow for a wide array of sensors. Sensor data affects the estimated pose through a likelihood function. As long as this model can be created, any sensor can be added to Merge, and each additional sensor, whether or not it is unique, decreases theoretical error. Currently, Merge makes use of any number of GPS and orientation devices. Other sensors are under development, such as temperature, pressure, etc., which can be added directly to the model, refining pose estimates even further.

A unique feature of Merge is its ability to take into account an a priori map of the environment should it exist [27, 28]. By taking into account the fact that users cannot travel through walls, or are unlikely to walk in the middle of the street, the uncertainty on pose can be decreased beyond what GPS can provide. A map of the Georgia Tech area was created by other researchers at the institute by creating a vector map overlaid upon georeferenced aerial photographs. This map was then rasterized, with various regions colored to an appropriate probability value.

This final pose estimate contains some amount of error that can be reduced by increasing the number of particles. However, this increases running time, which in turn also affects accuracy, so it is important to use an optimum number of poses. The user can choose the number to use, tailoring Merge to the performance characteristics of a specific hardware system.

2.3. Audio presentation

The success of the SWAN system is due to the design of its auditory display. To satisfy our requirements it needs to be exclusively auditory, and display navigation cues as well as other non-navigation sounds that convey time- and task-relevant information about the surroundings. It must allow for high bandwidth flow of information without over-taxing cognition or attention. The user must continue to hear speech and non-speech environmental sounds, and to carry on with other tasks simultaneously.

2.3.1. Path presentation (beacons). When presented to SWAN users, routes are composed of path segments joined together by nodes, or waypoints. Users need to know where the path starts, which direction the next waypoint is located, and how far it is. In order to accomplish this task, Merge determines the user's location, which is compared to the destination, along with any intervening waypoints. To guide the user along the path, SWAN presents a spatialized non-speech beacon sound that continually indicates the desired direction [29]. The beacon's virtual location is updated as the person's location or head orientation changes so as to always be a constant virtual distance away from the user so as to not introduce any volume attenuation (similar to the "spatial lead" mode in [30]). The user simply walks toward the sound, moving from waypoint to waypoint, until the destination is reached. When a user approaches a waypoint, the beacon tempo increases until the waypoint is reached, at which point there is a subtle success chime, and the beacon shifts to direct the user toward the next waypoint. For practical purposes, the waypoint itself is considered to be a disk, about 1.5 meters in diameter. If the user misses and overshoots the waypoint, the beacon sound changes

timbre to emphasize that the waypoint is now behind the listener (a form of front-back disambiguation), for faster recovery. Such a system has proven highly effective for navigation, is easily learned, and allows for successful transit of arbitrarily complex (and non-rectilinear) paths. We have also found that such a dynamic and failsafe system design is nearly impossible to implement with speech commands. Indeed, any missed waypoints (inevitable in a real-world application) result in considerable difficulties for users of a speech-based system.

2.3.2. Environment (features). As the user transits the path, features in the environment are also represented using spatialized non-speech sounds. SWAN can represent most anything in the environment that is relevant to the user; park benches, restrooms, bus stops, etc. Points of interest, obstacles, and surface transitions (e.g., change from level sidewalk to a descending stairway) are also represented. A combination of auditory icons, earcons, and spearcons are used to indicate the various features and presented simultaneously with beacons.

2.3.3. Ad hoc audio annotations. A method for general annotation of the environment is also supported. Using a menu system, a user may create an item in the GIS database to represent some feature in the environment at their current location. The user can enter information about the feature. By classifying the feature into a predefined category, the feature can then be associated with an easily identifiable non-speech auditory cue that can be played when encountered.

2.3.4. Audio menus. There are several ways to control the system (e.g., tell SWAN where you would like to go). One method is via a speech recognition/voice command interface. We have included this functionality via the Microsoft speech tools, as well as some more sophisticated word spotting methods. However, in practice, such an input method is neither reliable nor private enough to be seriously considered for continuous user input to a navigation system such as SWAN. Thus, we have also implemented a simple and effective audio menu system, to be used instead.

User interaction with SWAN is predominantly directed by the use of audio menus, which are navigated with the handheld physical device, described previously. The audio menu consists of a spoken text description for each menu item. The spoken text is either a pre-recorded annotation or text-to-speech (TTS) audio.

Menu items are generally preceded by earcons or spearcons, to enhance the menu. Earcons are brief,

music-like sequence of chords that facilitate identification of menu items, once the user becomes familiar with them [see, e.g., 20, 21]. However, recent research in our lab has shown that earcons can be difficult for users to learn, and lead to slow performance. We have determined that spearcons are more effective for menu enhancement [19], thus we are now augmenting the SWAN audio menus with spearcons.

The audio menu supports preemption by the user, meaning that items can be skipped over or selected without listening to the entire annotation. Preemption and techniques described above allow the user to quickly navigate the audio menu interface. The menu system supports the actions: move forward, move backwards, select current item, and back out of current menu level. Each level in the menu is represented by a circular doubly-linked list. Therefore the user can traverse through the menu and start over at the beginning without backing up. A unique audio notification is played when the user goes from the last item to the first, or vice versa, to help avoid confusion.

2.4. Path planning and map building

Path planning is a critical part of navigation for the visually impaired, and is thus a key feature of the SWAN project. Specifically, path planning encompasses the notion of the user determining step-by-step directions for traversing from point A to point B. First, the user must specify an origin and destination that defines a path. SWAN then calculates a path so the user may travel to her desired destination.

We have implemented path planning methods in which simple movement graphs are constructed from GPS data as specified by the user. The user has the option of manually or automatically dropping graph nodes (i.e., waypoints) based on the current GPS position. Each consecutively stored waypoint is assumed to have a line of sight devoid of obstacles with the previous waypoint and therefore adjacency information can be stored, thus forming a path. The automatic mode of waypoint dropping calculates metrics of user movement to determine appropriate points at which to drop a node. Waypoints can be specifically named by text entry or voice annotation and therefore can serve as origins or destinations in path planning. Once the path is stored however, the user may recall the path stored in SWAN. Additionally, a method of merging maps as well as finding self-intersections is implemented via simple 2D edge intersections and merging of overlapping nodes.

Currently, Dijkstra's Algorithm is used to determine the shortest path between the origin and desired

destination and is similar to the method used in PGS [10]. SWAN users have a method of managing the size of their map due to the ability to create multiple differing map files and simply load whichever map is applicable to their current geographic position.

While SWAN is now heavily integrated with ESRI GIS components, the ESRI route planning tools are not used. Generally, these routines are specific to automobile navigation and are not necessarily easily applicable to pedestrian path planning. Another point of consideration is that GIS data does not accurately denote the most suitable pedestrian paths. Sidewalks may or may not be available in readily available GIS databases. Even so, these databases may not effectively represent scenarios such as discrete positioning within a wide sidewalk or patio, walkways available to everyone on private property, cutting through a building, favoring a sidewalk on a particular side of a street for improved safety, etc. Often sidewalk data in widely available GIS databases only includes public sidewalks implicitly assumed to be adjacent to roads and represented by line segments (no information to represent the exact dimensions of the sidewalk).

The previously mentioned Drishti system implements a novel route server built with ESRI tools that can present pedestrian directions to the user that take into account the user's preference for avoiding obstacles rather than simply returning the shortest path [13]. However, Drishti relies on representing path information as line segments buffered to walkway width and is therefore only a rough approximation of where a pedestrian can traverse.

For these reasons, we seek to represent available pedestrian paths with the true boundaries of pedestrian-friendly areas rather than approximating with line segments. Providing path planning for this sort of map representation is common in robotics [31] and video games [32]. The A* algorithm is often used in when the map is known and static [33, 34].

Our tracking framework, Merge, uses a 2D probability map that can also serve the purpose of providing an accurate map that can aid in the determination of a path. We can simply use adjacency information of the pixels that form the raster map images to facilitate use of the A* algorithm. However, we may be able to improve the efficiency of the searches by utilizing a non-uniform region segmentation instead of uniformly-sized square pixels.

3. Evaluation

In addition to the development and engineering efforts, we have been engaged in an active program of evaluation of the SWAN system [35]. A mixture of virtual environment and real-world testing is being conducted; employing both sighted and visually impaired participants since both are potential user populations for the device. We have focused mostly on the auditory interface, in order to determine how effective the non-speech sounds are in guiding a user along a path. Experimental studies have helped us determine the optimal kinds of non-speech audio [29], as well as the best ways to design the details of interaction, such as capture radius (the actual size of the waypoint) [36], to maximize speed or accuracy, as appropriate. One example of this concerns the capture radius, where there has been found to be a tradeoff between speed of navigation and accuracy of navigation based on varying radii used. Larger radii tend to promote faster navigation at the cost of accuracy, while smaller radii lead to more accurate but often slower navigation. Thus far, the system—and the auditory interface in particular—has proven itself to be simple and effective in guiding users along both simple and complex paths. The current stages in evaluation include studies of how effective the system can be in the presence of external noises sources, and an assessment of the cognitive workload the system imposes [37].

It is also important to note that other complementary, robust lines of research into non-visual navigation systems are longstanding and ongoing, such as the research surrounding the PGS system [38]. This research has also demonstrated the potential efficacy of such a navigation system [39], as well as examining other key aspects such as user hardware preferences [40]. In addition this group has also investigated more basic perceptual factors a non-visual navigation system must take into account such as auditory distance perception. A good summary of this work is found in Loomis, et al [40].

4. Future work

One key future direction for work is the evaluation of the use of new sensors for the system. No one sensor performs well in all of the given situations the SWAN is likely to be used in. SWAN is inherently robust in its use of multiple sensors. Nevertheless, ongoing studies are aimed at discovering what additional sensors are useful for position tracking and in what situations they are most effective. Our belief is that with the correct sensors, locations will be able to be identified using their unique multi-sensor signature. The more precision and accuracy the system has in terms of position tracking, the more reliable the output.

Cameras, as part of a computer vision system, are one type of sensor with great potential benefit to SWAN. We are developing methods to estimate a user's position in real time based on images from multiple small, mounted cameras [41]. In order to make use of computer vision, some prior knowledge must be obtained. We are building 3D databases containing reference points, paired with image feature descriptions. These databases will enable the SWAN computer vision system to estimate the user's location based on images matched in the database. To make this practical for pedestrian navigation, we are developing a camera rig capable of processing images fast enough for real time localization. When integrated into SWAN, this vision system is expected to yield excellent location and orientation information, particularly indoors. When entering a building GPS may lose satellite signals, lowering accuracy if a fix can even be obtained. In such a situation, the vision system would still provide sensor data for position estimates, allowing SWAN to carry on uninterrupted. If SWAN can tell when it is indoors using other sensors, for instance by the amount of ambient light, then vision can be given more importance, scaling back reliance on GPS as appropriate.

5. Conclusion

We have developed a System for Wearable Audio Navigation (SWAN) with powerful localization capabilities, and a novel and effective auditory interface. SWAN aids a user in safe pedestrian navigation and includes the ability for the user to create and store personal pedestrian navigation paths. Emphasis is placed on the use of non-speech sounds through a process of sonification. SWAN serves as a foundation for research into a variety of aspects of psychoacoustics, human computer interaction, and novel tracking technology.

6. Acknowledgements

7. References

[1] National Center for Veteran Analysis and Statistics, National Survey of Veterans: Assistant Secretary for Policy and Planning, Department of Veterans Affairs, Gov't Printing Office, 1994.

[2] W. De l'Aune, Legal Blindness and Visual Impairment in the Veteran Population 1990-2025. Decatur, GA: VA Rehabilitation R&D Center, 2002.

[3] G. L. Goodrich, Growth in a Shrinking Population: 1995-2010. Palo Alto, CA: Palo Alto Health Care System, 1997.

[4] S. LaGrow and M. Weessies, Orientation and mobility: Techniques for independence: Royal New Zealand Foundation for the Blind, Dunmore Press, 1994.

[5] B. Blasch, W. Wiener, and R. Welsh, Foundations of Orientation and Mobility, 2nd ed. New York: American Foundation for the Blind, 1997.

[6] S. B. Levy and A. R. Gordon, "Age-related vision loss: Functional implications and assistive technologies," International Journal of Technology and Aging, vol. 1, pp. 16-125, 1988.

[7] D. Guth and R. LaDuke, "Veering by blind pedestrians: Individual differences and their implications for instruction," Journal of Visual Impairment & Blindness, vol. 89, pp. 28-37, 1995.

[8] C. S. Kallie, P. R. Schrater, and G. E. Legge, "Variability in Stepping Direction Explains the Veering Behavior of Blind Walkers," Journal of Visual Impairment & Blindness, vol. 33, pp. 183-200, 2007.

[9] J. M. Loomis, R. Klatzky, and R. Golledge, "Navigating without vision: basic and applied research," Optometry and Vision Science, vol. 78, pp. 282-289, 2001.

[10] J. M. Loomis, R. G. Golledge, and R. L. Klatzky, "GPS-based navigation systems for the blind," in Fundamentals of Wearable Computers and Augmented Reality, 1 ed, W. Barfield and T. Cauldell, Eds. Mahway, NJ: Lawrence Erlbaum Associates, Inc., 2001, pp. 429-446.

[11] T. Strothotte, S. Fritz, R. Michel, A. Raab, H. Petrie, V. Johnson, L. Reichert, and A. Schalt, "Development of dialogue systems for a mobility aid for blind people: initial design and usability testing," in Proceedings of ASSETS '96: The Second Annual Conference on Assistive Technology, New York, 1996.

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

[13] S. E. Moore, "Drishti: An Integrated Navigation System for the Visually Impaired and Disabled," Master's Thesis Gainesville, FL: University of Florida, 2002.

[14] T. V. Tran, T. Letowski, and K. S. Abouchacra, "Evaluation of acoustic beacon characteristics for navigation tasks," Ergonomics, vol. 43, pp. 807-827, 2000.

[15] G. H. Mowbray, "Simultaneous vision and audition: The comprehension of prose passages with varying levels of difficulty," Journal of Experimental Psychology, vol. 46, pp. 365-372, 1953.

[16] C. D. Wickens, Engineering psychology and human performance, 2nd ed. New York, NY: HarperCollins Publishers, 1992.

[17] J. Lindsay and B. N. Walker, "The Effect of a Speech Discrimination Task on Navigation in a Virtual Environment," Submitted for publication.

[18] R. L. Klatzky, J. R. Marston, N. A. Giudice, R. G. Golledge, and J. M. Loomis, "Cognitive load of navigating

without vision when guided by virtual sound versus spatial language," Journal of Experimental Psychology: Applied, vol. 12, pp. 223-232, 2006.

[19] B. N. Walker, A. Nance, and J. Lindsay, "Spearcons: Speech-based earcons improve navigation performance in auditory menus," in International Conference on Auditory Display, London, England, 2006, pp. 63-68.

[20] J. Marila, "Experimental comparison of complex and simple sounds in menu and hierarchy sonification.," in International Conference on Auditory Display, Kyoto, Japan, 2002, pp. 104-108.

[21] S. A. Brewster, P. C. Wright, and A. D. N. Edwards, "An evaluation of earcons for use in auditory human-computer interfaces," in SIGCHI Conference on Human Factors in Computing Systems, Amsterdam, 1993, pp. 222-227.

[22] N. Sawhney and C. Schmandt, "Nomadic Radio: Scalable and Contextual Notification for Wearable Audio Messaging," in ACM Conference on Human Factors in Computing Systems (CHI'99), New York, 1999, pp. 96-103.

[23] D. A. Ross, "Implementing assistive technology on wearable computers," in IEEE Personal and Ubiquitous Computing, 2001, pp. 2-8.

[24] Holosonics, "Audio Spot Light." http://holosonics.com

[25] B. N. Walker and R. M. Stanley, "Thresholds of audibility for bone-conduction headsets," in International Conference on Auditory Display, Limerick, Ireland, 2005, pp. 218-222.

[26] R. M. Stanley, "Toward adapting spatial audio displays for use with bone conduction: The cancellation of bone-conducted and air-conducted sound waves," in Psychology Atlanta, Georgia: Georgia Institute of Technology, 2006.

[27] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte Carlo Localization for Mobile Robots," in IEEE International Conference on Robotics and Automation (ICRA), 1999.

[28] S. M. Oh, S. Tariq, B. Walker, and F. Dellaert, "Map-Based Priors for Localization," in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2004.

[29] B. N. Walker and J. Lindsay, "Effect of beacon sounds on navigation performance in a virtual reality environment," in International Conference on Auditory Display, Boston, MA, 2003, pp. 204-207.

[30] J. M. Loomis, R. G. Golledge, and R. L. Klatzky, "A geographical information system for a GPS based personal guidance system," International Journal of Geographical Information Science, vol. 12, pp. 727-749, 1998.

[31] J.-A. Meyer and D. Filliat, "Map-based navigation in mobile robots - ii. a review of map-learning and path-planning strategies," Journal of Cognitive Systems Research, vol. 4, pp. 283-317, 2003.

[32] I. L. Davis, "Warp Speed: Path Planning for Star Trek: Armada," in Association for the Advancement of Artifical Intelligence (AAAI) 2000 Spring Symposium on Interactive Entertainment and AI Proceedings, 2000.

[33] R. Dechter and J. Pearl, "Generalized best-first search strategies and the optimality af A*," Journal of the ACM, vol. 32, pp. 505-536, July 1985.

[34] K. I. Trovato and L. Dorst, "Differential A*," IEEE Transactions on Knowledge and Data Engineering, vol. 14, pp. 1218-1229, 2002.

[35] B. N. Walker and J. Lindsay, "Navigation performance with a virtual auditory display: Effects of beacon sound, capture radius, and practice," Human Factors, vol. 48, pp. 265-278, 2006.

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

[37] B. N. Walker and J. Lindsay, "The effect of a speech discrimination task on navigation in a virtual environment," in Human Factors and Ergonomics Society, San Francisco, 2006.

[38] J. Loomis, R. Golledge, R. Klatzky, J. M. Speigle, and J. Tietz, "Personal guidance system for the visually imapired," in First Annual International ACM/SIGCAPH Conference on Assistive Technologies (ASSETS94), Marina del Rey, CA, 1994.

[39] J. Marston, J. Loomis, R. Klatzky, R. Golledge, and E. Smith, "Evaluation of Spatial Displays for Navigation without Sight.," ACM Transactions on Applied Perception, vol. 3, pp. 110-124, 2006.

[40] R. G. Golledge, J. R. Marston, J. M. Loomis, and R. L. Klatzky, "Stated preferences for components of a Personal Guidance System for non-visual navigation," Journal of Visual Impairment & Blindness, vol. 98, pp. 135-147, 2004.

[41] F. Dellaert and S. Tariq, "A Multi-Camera Pose Tracker for Assisting the Visually Impaired," in 1st IEEE Workshop on Computer Vision Applications for the Visually Impaired, 2005.