3-D SOUND NAVIGATIONAL AID FOR THE BLIND BASED ON REAL-TIME SONAR RANGE MEASUREMENTS

Armando B. Barreto, Maroof H. Choudhury and Carlos A. Ordóñez
Digital Signal Processing Laboratory
Electrical & Computer Engineering Department
Florida International University
Miami, FL, 33174, USA

KEYWORDS

3-D sound, Multimedia, Disabilities, Ranging, Sonar, Blind Navigation, Electronic Travel Aides (ETAs).

ABSTRACT

This paper reports the design, implementation and evaluation of a real-time system that uses 3-D sound to inform a blind user of the location of obstacles around him/her and to provide a "spatial auditory beacon", indicating the North direction. The 3-D sounds are created through the use of Head-Related Transfer Functions (HRTFs) implemented in a portable computer. The distance from the current location to the nearest obstacles in six directions around the user are measured continuously by six sonar ranging modules, arranged radially on a headgear worn by the user. The system also incorporates a digital compass to continuously update the relative location of the North direction, around the user. The system is portable and completely battery-operated. The system has been verified by blind navigation tests on four subjects, yielding an average navigation efficiency of 93.1%.

I. INTRODUCTION

In attempting to navigate through his/her environment, a blind person faces two main types of challenges: 1) Identify the location of obstacles, and 2) Figure out a safe route through the obstacles towards the intended destination, in reference to his/her position. Electronic systems commonly known as Electronic Travel Aides (ETAs) assist visually impaired people in both of these tasks. Several devices have been developed in the past as an electronic replacement of the guide dog. The Personal Area Locator (PAL) [Dodson et al., 1999] designed primarily for outdoor navigation applications, utilizes GPS (Global Positioning System) and RDS (Radio Data Services) in addition to a compass for directional orientation detection. The 'Navbelt' [Shoval et al., 1998] utilizes sonar technology to evade obstacles in indoor environments. Similarly, the system described here utilizes sonar technology to estimate the distances between the subject and the closest obstacles in his/her surroundings, creating an estimated panoramic map of those surroundings.

Our system has been designed to continuously monitor and report sets of six range measurements [Choudhury and Barreto, 2002], which are communicated to the user as spatialized sound by a 3-D sound rendering system, also developed in our laboratory [Ordóñez et al., 2002]. 3-D sound is also used to provide a "spatial auditory beacon" to the user, upon request. This beacon consists of a sound that appears to originate from the North direction.

II. SYSTEM LAYOUT & FUNCTIONAL DESCRIPTION

The overall layout of the system is shown in Figure 1. The two key computational components are the 3-D SOUND RENDERING ENGINE (3DSRE), which was prototyped in a portable computer, using the Matlab® simulation environment, and the SONAR AND COMPASS CONTROL UNIT (SCCU), implemented as an embedded system using an OOPIC (Object Oriented Programmable Integrated Circuit)[Savage Innovations, 2001]. This is a Microchip PIC16F77-based microcontroller [Microchip, Inc., 2000] running a real time Operating System (OS). The six sonar modules and the digital compass unit are connected to the SCCU, which controls them and receives information from them.

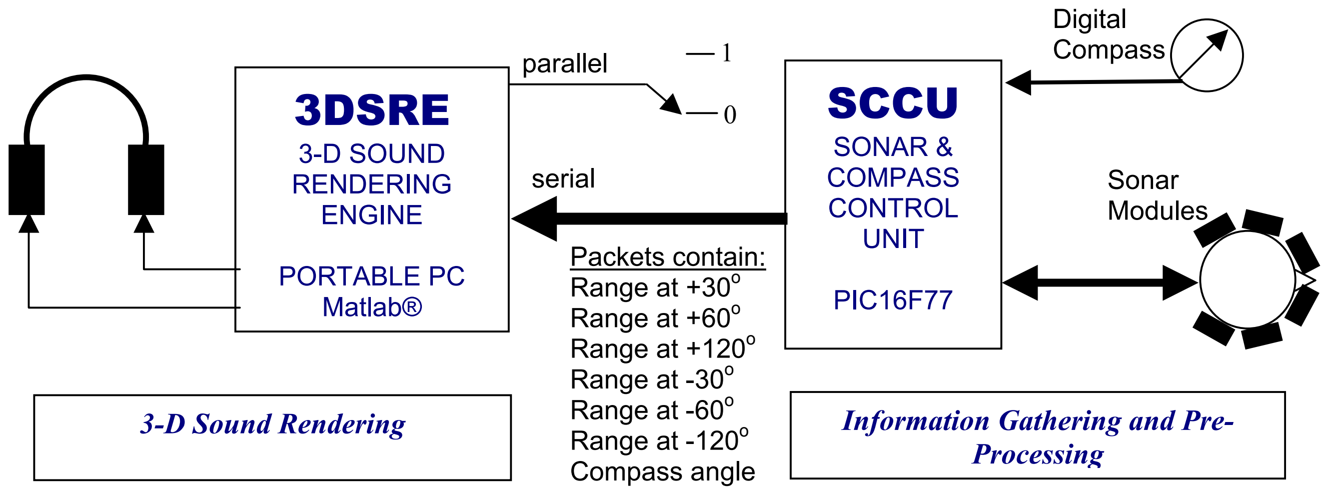

Figure 1. Block Diagram for the 3-D Sound Blind Navigation System.

Figure 1 shows that the functionality of the overall system is cleanly divided into two main functions, and it indicates that each main function is performed by each of the two main computational blocks:

Information Gathering and Pre-processing: The SCCU controls the triggering sequence of the six sonar modules and reads the range information from them. Similarly, it continuously reads the most recent information provided by the Digital Compass. The SCCU produces a packet of information containing the range measurements in six predetermined directions (with respect to 0^o *azimuth*, which is the direction ahead of the subject, at any time), as well as the most recent *heading* angle, which is the facing direction of the traveler's head with respect to the North direction [Loomis et al., 1998]. This packet is delivered every 0.5 seconds (irrespective of the current operational mode of the system) to the 3DSRE through a serial link.

3-D Sound Rendering: The 3DSRE has been implemented in a portable computer (PC), through a Matlab® program that a) Determines the current operational mode of the system, reading the state of the push-button connected to the PC's parallel port and operated by the user; b) reads the packets of information transmitted from the SCCU, into the PC's serial port; and c) creates the binaural left and right output signals to render six superimposed spatialized sounds to reflect the distances from the user location to the nearest objects in the corresponding directions ("Obstacle Map Mode"), or a single spatialized sound, simulated as if originated in the North direction, with respect to the user ("North Beacon Mode").

Further details on the implementation of both of these main functions are provided in the following sections.

III. INFORMATION GATHERING AND PRE-PROCESSING

The information gathering and Pre-processing function requires the SCCU to continuously interact with the sensors required for two tasks:

- 6-channel sonar range measurement
- Heading angle measurement (with respect to North)

Furthermore, the information resulting from these two tasks is assembled by the SCCU in packets that are sent to the 3DSRE every 0.5 seconds through a standard serial link.

III.1 6-channel range measurement

The omni-directional ranging system utilizes six Devantech SRF04 sonar modules to retrieve ranging information in six directions around the subject. These sensor modules are set in a headgear worn by the user (Figure 2). They are intended to sense obstacles in the following azimuthal directions $+30^o$, $+60^o$, $+120^o$, -30^o, -60^o, and -120^o, (where 0^o azimuth indicates the direction straight-ahead with respect to the user and 90^o azimuth is directly to the right of the subject), as shown in Figure 3.

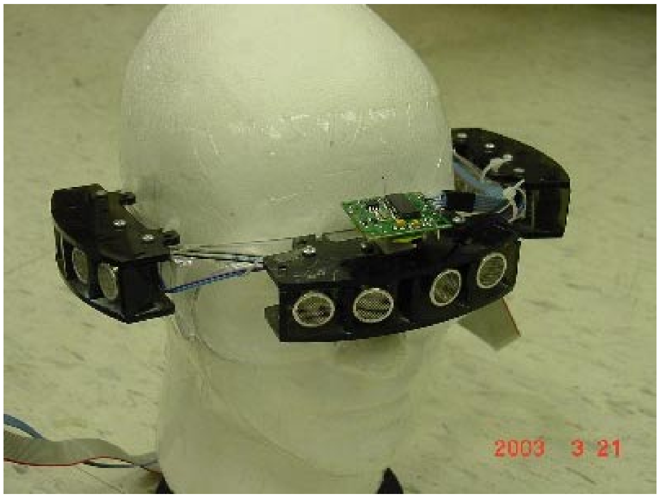

Figure 2. Headgear housing the six sonar modules and the digital compass unit

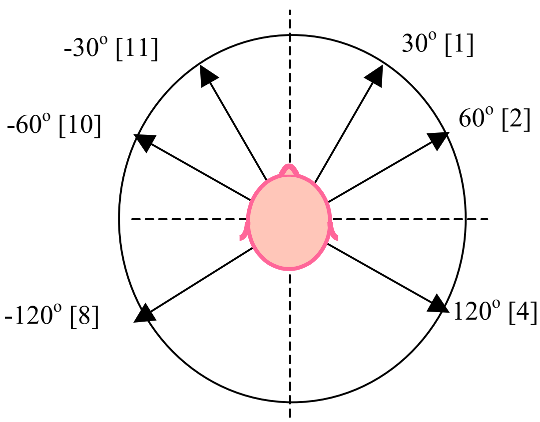

Figure 3. Main azimuthal directions explored by the system

The Devantech SRF04 sonar module transmits an ultrasonic frequency of 40 KHz, and has a guaranteed range of operation from 3 cm to 3 meters. Its sensitivity is such that it detects a 3 cm diameter stick at 2 m range. Triggering the sonar requires a TTL-level pulse of a minimum duration of 10 microseconds. After it has been triggered the SRF04 will provide a TTL-high level in its “Echo output”, until the ultrasonic echo has been detected. At that time, the “Echo output” will switch to TTL-low. The complete timing sequence for this sonar module is shown in Figure 4. Figure 5 shows the nominal beam pattern for this sonar module [Acroname Inc., 2001].

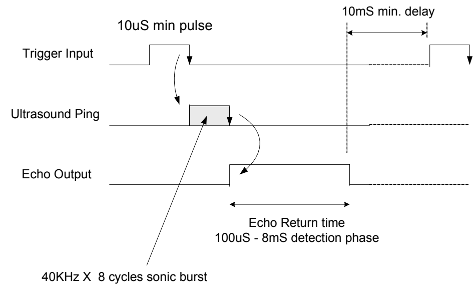

Figure 4. Timing diagram for the Devantech SRF04

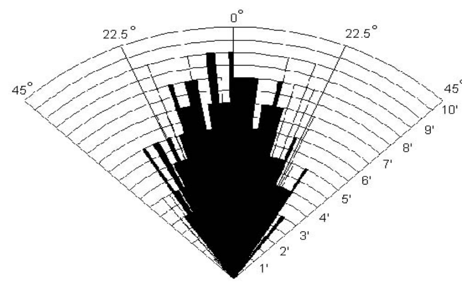

Figure 5. Beam pattern for the Devantech SRF04 sonar

The triggering of the sonar modules and the acquisition of the ranging information obtained from each one of them are under the control of the SCCU. In addition to the OOPIC, the SCCU includes an 8-bit counter for each of the sonar modules. Each one of these counters is enabled/disabled by the “Echo output” line form the corresponding sonar module and driven by a 6.75 KHz clock, in such a way that the maximum count reached in the counter represents the range, in inches, detected by the associated sonar module. The clock period is approximately 148 μS, which represents the echo return time for a distance of one inch. The 8-bit counters therefore can effectively measure up to a distance of 255 inches. An over-range prevention scheme employed in the logic circuit suspends the count operation of an individual counter when it reaches a count of 255.

III.2 Measurement of user heading with respect to North

The headgear worn by the user also incorporates a Devantech CMPS01 Magnetic Compass module. This module is based on two Philips KMZ51 magnetic field sensors, which are sensitive enough to detect the Earth's magnetic field. These two sensors are mounted at right angles, with respect to each other, so that the direction of the horizontal component of the Earth's magnetic field, with respect of the sensor reference direction can be computed. This direction is continuously updated and made available to the SCCU as a 1-byte, 255-step compass reading, which can represent up to 360^o of deviation from the compass reference direction to the North direction (1 full revolution). This heading with respect to the North direction is included as the last data item in each packet sent from the SCCU to the 3DSRE.

IV. 3-D SOUND RENDERING

The packet of information built by the SCCU and received by the 3DSRE allows the system to implement two modes of operation:

"Obstacle Map Mode", in which six 3-D sound sources are created at the azimuth angles explored by the six sonar modules and at distances defined by the six range values included in the latest SCCU packet received. The composite left and right binaural channels that result from superimposing the left and right channels simulated for each one of the six sources are delivered to the headphones worn by the user. This is the default mode of operation of the system.

"North Beacon Mode", in which only a single 3-D sound source is created at one of 12 predefined azimuth values (at 30^o intervals) around the subject. The pre-defined azimuth location to use is the one that most closely reflects the North direction with respect to the current orientation of the user's head (as reported by the heading value received in the latest SCCU packet received). This mode of operation is adopted only while the user is pressing a push button connected to the parallel port of the PC in which the 3DSRE is implemented.

The functionality of both modes is based on the application of the Head-Related Transfer Functions (HRTFs) for different locations in the horizontal plane. The following subsections outline the concept of Head-Related Transfer Functions and then explain how it is used in each mode of operation.

IV.1 Sound Spatialization through the use of Head –Related Transfer Functions

Head-Related Transfer Functions (HRTFs) are widely used for the synthesis of binaural sound, which can produce in the listener the illusion of a sound that originates at a virtual location around him/her.

Sounds traveling from a physical source to the listener's eardrums undergo modifications due to reflections and interactions with the listener's torso, shoulders, head and the pinnae or outer ears. The Head-Related Transfer Functions (HRTFs) capture the nature of these modifications. The left and the right HRTFs contain all the information required by the listener to localize the source of the sound. Conversely, a single channel digital sound can be processed by these HRTFs to generate left and right audio signals that will make the listener believe that the sound emanates from the corresponding virtual source location.

For example, if the sound source is in the left hemisphere, with respect to the listener, the sounds arriving at the right ear (contralateral) are delayed and also attenuated, with respect to the sounds reaching the listener's left ear. This small but significant change in phase and intensity of sound is interpreted as a change in direction by the brain. In the horizontal plane, the most important cue that helps a person localize a sound source is the Inter-aural Time Difference (ITD) [Kuhn, 1977]. The Inter-aural Intensity Difference (IID) also plays an important role. Both of these sound localization cues are contained in the HRTFs.

The HRTFs of given subject, for a given location around him/her, can be obtained experimentally by playing a broad-band signal through a speaker situated at the location under study, and placing small microphones at the entrance of the ear canals of the subject. We have used the Ausim3D HeadZap system to obtain the HRTFs of a representative subject at 12 different azimuth angles, in the horizontal plane (uniformly separated by intervals of 30^o). This system measures a 128-point impulse response for each of the left and the right ear using a sampling frequency of 44.1 KHz. Golay codes are used to generate a broad-spectrum stimulus signal delivered through a Bose Acoustimass speaker. The response is measured using miniature blocked-meatus microphones placed at the entrance to each ear canal. This response is processed to extract the direct path response, remove speaker and microphone effects, and equalize the resulting Head-Related Impulse Responses (HRIRs), which are the time-domain expressions of the HRTFs [AuSim 3D, 2000].

Once the left and right HRIR sequences for a given azimuth θ, $h_{L\theta}(n)$ and $h_{r\theta}(n)$, have been obtained experimentally, a single monaural sound sequence s(n) can be convolved with each one of them, to yield the left and right components of a binaural pair of sounds, $y_{L\theta}$ and $y_{R\theta}$. When delivered through headphones to the listener, these two binaural components will make the listener believe that the sound s(n) emanates from the corresponding virtual source location, at azimuth θ :

$$y_{L\theta}(n) = \sum_{k=-\infty}^{\infty} h_{L\theta}(k) \cdot s_{\theta}(n-k) \qquad (1)$$

$$y_{R\theta}(n) = \sum_{k=-\infty}^{\infty} h_{R\theta}(k) \cdot s_{\theta}(n-k) \qquad (2)$$

The distance between the virtual sound source represented this way and the user is indicated to the listener by making the volume of both $y_{L\theta}$ and $y_{R\theta}$ vary in an inversely proportional manner, with respect to the desired virtual distance.

IV.2 North Beacon Mode Implementation

For the implementation of the North Beacon Mode, a clicking monoaural sound was convolved with all the 12 pairs of HRIRs experimentally measured from the prototype subject, according to equations 1 and 2. The 12 pairs of resulting binaural signals were stored in memory.

For this mode of operation the 3DSRE uses only the 8-bit angle value provided by the digital compass module, which represents the heading of the subject, with respect to the North direction. Thus, in order to provide a 3-D sound beacon that will appear to emerge from the North direction, the heading angle received in the SCCU packet was inverted and rounded to one of the 12 available azimuth values (available in multiples of 30^o), to select and play the corresponding binaural pair of sounds to the user.

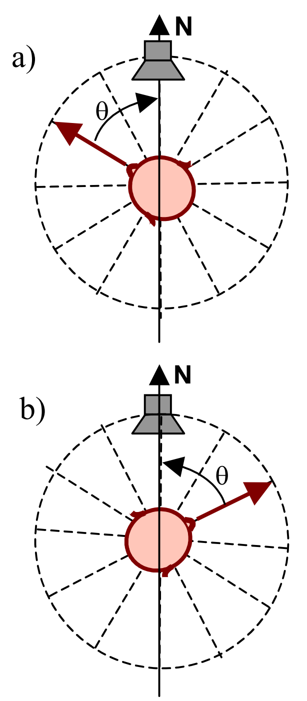

Figure 6. Examples of virtual sound placement (represented by the speaker symbol), in North Beacon Mode, when a) The heading of the subject is 60^o counterclockwise from the North direction, and b) The heading of the subject is 60^o clockwise from the North direction. Azimuth values q = 60o and q = -60o are used for the spatialization process, respectively

This 3-D sound enables the system to provide the user with an approximate direction of reference (North), upon his/her request. Figure 6 shows the intended placement of the single virtual sound source in this mode of operation, for two illustrative cases.

IV.3 Obstacle Map Mode Implementation

In this mode of operation the system creates simultaneously six virtual sound sources and renders the composite binaural sound channels to the left and right headphones, for the user to hear. Each of the six sound sources is simulated using appropriate HRTFs at a virtual location matching the direction explored by each one of the six sonar ranging modules (i.e., θ = +/- 30^o, +/-60^o and +/-120^o). The virtual distance of each source to the listener is simulated as proportional to the range information for its corresponding direction, obtained from the latest SCCU packet received by the 3DSRE.

For this mode six bursts of uncorrelated random noise were used as the input sounds for the six parallel spatialization processes, s_θ . They were convolved with the corresponding HRIR pairs, $h_{L\theta}$ and $h_{R\theta}$, as indicated by equations 1 and 2, to yield six pairs of binaural sequences $y_{L\theta}$ and $y_{R\theta}$. Each binaural sequence pair simulates a virtual sound coming from each of the directions explored by the sonar elements. These binaural sequence pairs must be amplified individually by gains $G_\theta(r_\theta)$, inversely proportional to the virtual distances, r_θ, intended for each virtual source and then mixed together for delivery through headphones. This process is implemented as two independent weighted sums of signals, each one resulting in a composite binaural component for the headphone rendering of the complete set of six virtual sound sources:

$$w_{RHP}(n) = \sum^{\theta} G_\theta(r_\theta) y_{R\theta}(n) \quad (3)$$

$$w_{LHP}(n) = \sum^{\theta} G_\theta(r_\theta) y_{L\theta}(n) \quad (4)$$

In equations 3 and 4, w_{LHP} and w_{RHP} are the headphone left and right channels that result from the weighted accumulation of all six left and right binaural components, respectively. That is, the summations indicated in equations 3 and 4 are over the six values of azimuth explored by the sonar elements: θ = 30^o, 60^o, 120^o, -120^o, -60^o, -30^o. Figures 7 and 8 illustrate the implementation of the Obstacle Map Mode in two specific navigation scenarios: Navigating along a hallway, and entering an intersection. Notice how some of the virtual sounds will practically disappear as the subject enters an intersection, as shown in Figure 8. Making the blind individual aware of the opportunities to turn right or left at the intersection is an important functional advantage provided to the user.

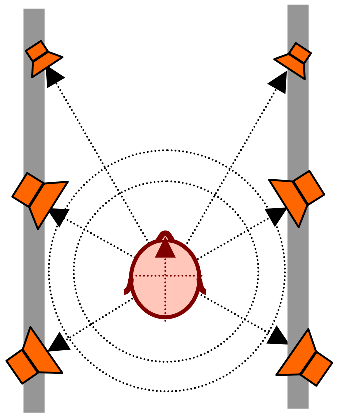

Figure 7. Distribution of the six virtual sound sources implemented in the Obstacle Map Mode, as the user moves along a hallway. The size of the speaker symbol is used to represent the volume of the corresponding 3-D virtual sound.

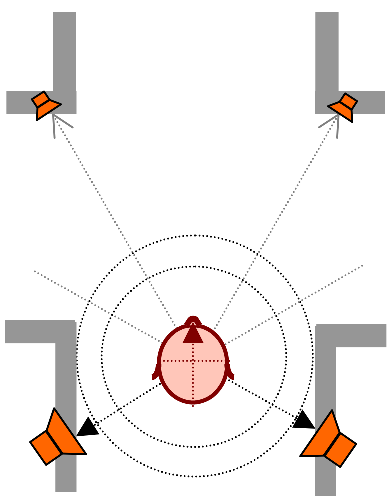

Figure 8. Distribution of the virtual sound sources implemented in the Obstacle Map Mode, as the user enters an intersection. The size of the speaker symbol is used to represent the volume of the corresponding 3-D virtual sound. The sources at +/- 60^{o} azimuth will practically disappear.

V. SYSTEM EVALUATION

A prototype of the system has been assembled for evaluation. This prototype is completely portable, requiring a 9.6 V Ni-MH, 1600 mAh rechargeable battery (Radio Shack 23-3318), to power the SCCU, implemented in a printed circuit board. The 3DSRE is implemented in a portable PC, powered by its standard rechargeable battery. The hardware associate with both of the modules (SCCU and 3DSRE) can be seen in Figure 9.

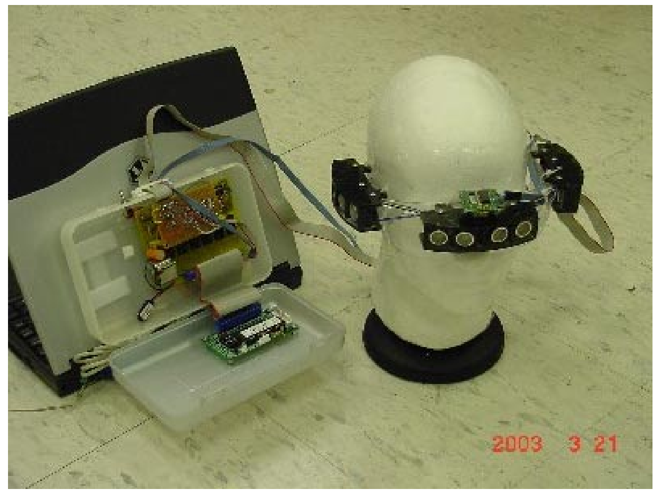

Figure 9. Hardware platforms used for the implementation of the blind navigation system.

We have tested the use of the system by four fully-sighted individuals that employed it to navigate the hallways of the third floor of the FIU Engineering Center (FIU-EC), while blindfolded.

Subjects were placed at the South end of the FIU-EC 3rd floor, and asked to progress to the North as far as they could go in an interval of 10 minutes. The starting spot where the timed trials initiated was selected such that there is not a straight path to the North end of the third floor from there. Reaching the North end of the third floor would require at least eight turns, and it would represent an *effective displacement North* (along a hypothetical straight path) of 400 feet.

Subjects wore a blindfold that completely blocked their vision and placed the headgear housing the sonar and digital compass modules on top of the blindfold. Both the portable PC implementing the 3DSRE and the custom hardware implementing the SCCU were placed in a bag, which was carried by the subjects on one shoulder. Subjects were instructed to hold the push button used to switch system modes of operation in one hand, and allowed to extend the other hand in front of them for safety. Subjects listened to the binaural left and right channels produced by the 3DSRE through AIWA HP-

M046 stereophonic headphones connected to the headphone output of the portable computer (Figure 10). It should be stressed that the subjects were instructed not to use their free hand for groping to aid their navigation, but only to prevent a potential head-on collision. Similarly, it should be emphasized that subjects did not use other navigational aides (such as a long cane), and relied exclusively on the system for navigation.

Figure 10. Typical subject preparation for the blind navigation tests.

All four participants were familiar with the general layout of the third floor of the FIU-EC, and each one of them was allowed a training period of 10 to 15 minutes with the system prior to the test.

During the timed trials subjects were followed by one of the authors, who videotaped their navigation, maintaining a distance of no less than 10 feet from them. The videotape for each subject was played back and viewed to approximate the trajectory of the subject through linear segments in an Autocad® blueprint of the FIU-EC third floor. This enabled the assessment of the total trajectory traveled by each subject (along a real path) in the 10 minutes allotted and the total *effective displacement to the North* achieved (as if measured over a hypothetical straight South-to-North trajectory).

Table I shows both measurements for each subject and the corresponding average values.

TABLE I. Experimental Results.

Subject No.	Trajectory Traveled (ft.)	Effective Displacement North (ft.)
1	212	203
2	362	308
3	417	336
4	386	367
Average	*344.3*	*303.5*

The values in Table I indicate that, in an average, subjects were able to navigate, avoiding obstacles and finding the points where turns were required, at an approximate rate of (344.3 ft / 600 sec) = 0.57 ft/second.

Another important aspect of the experimental results is the deviation of the actual path that each of the subjects took from an ideal path. The actual paths observed deviate from the ideal path because of meandering due to hesitation and, in some cases, making a wrong turn and then having to correct the course. Table II compares the actual path lengths measured for each subject (taken from Table I), to the length of an ideal path (making only the correct turns along the way), *for the same effective displacement towards the North that each subject achieved*. This table also includes the ratio of the length on the ideal path to the length on the actual path. This ratio reflects a measure of navigational efficiency, with its maximum possible value of 1.0 representing the most efficient navigation, along the ideal path, which does not contain any deviations or wrong turns.

TABLE II. Navigation Efficiency

Subject No.	Actual Path Length (ft)	Ideal Path Length (ft)	Efficiency Ratio
1	212	209	0.986
2	362	325	0.898
3	417	353	0.847
4	386	383	0.992
Average	*344.3*	*317.5*	*0.931*

VI. CONCLUSIONS

Our research has resulted in the design, implementation and preliminary evaluation of a multimedia system that uses 3-D audio to render navigational information to a user in a non-visual, intuitive and non-overwhelming fashion.

Our development has been cleanly divided in two architectural and functional components. The first one is the Information Gathering and Pre-processing stage, in which the Sonar and Compass Control Unit (SCCU) interacts with six sonar modules and a digital compass to continuously update estimations of the distance from the current user position to surrounding obstacles in six directions around the user, and also estimates the heading of the subject (with respect to the North). The second component is the 3-D Sound Rendering of those pieces of information, which is carried out in one of two operational modes: “Obstacle Map” and “North Beacon”, by the 3-D Sound Rendering Engine (3DSRE), which can be chosen by the user.

We have implemented both components of the design using custom-made circuitry in a printed circuit board, a standard OOPIC board and a portable PC. The complete system is portable and operated by two rechargeable batteries.

We have tested the system in 10-minute trials in which blindfolded subjects were instructed to attempt to reach as far North as possible, through the network of hallways in the third floor of our FIU-EC building. The four subjects that participated in the tests did not use any other form of guidance and were instructed to rely exclusively on the system for navigation. Subjects were successful in navigation through the hallways, avoiding collisions with the walls, and traveling at an average pace of 0.57 ft/sec. Furthermore, subjects were generally successful in orienting themselves and maintaining a course due North whenever possible, minimizing unwanted deviations and wrong turns, to reach an average navigational efficiency of 93.1%, as defined by the ratio of the length of the trajectory actually traveled over the length of an ideal trajectory (rid of any deviation) connecting the same starting and ending points.

We believe that the results obtained are encouraging, prompting us to proceed towards the enhancement and miniaturization of this cost-effective system, and to explore its potential interaction with other traditional resources for blind navigation, such as the long cane.

VII. ACKNOWLEDGMENTS

This research was supported by NSF grant EIA-9906600 and ONR grant N00014-99-1-0952.

VIII. REFERENCES

[1] Acroname Inc. "The Devantech SRF04 Ultrasonic Range Finder", product literature, 2001.

[2] AuSIM3D, HeadZap: HRTF measurement system manual. AuSIM, Inc., 4962 El Camino Real, Suite 101, Los Altos, CA 94022, 2000.

[3] Cheng C. I. and Wakefield G. H., "Introduction to Head-Related Transfer Functions (HRTF): Representations of HRTFs in Time, Frequency and Space" Journal of audio engineering society Vol. 49. No. 4, pp. 231 – 249, April 2001.

[4] Choudhury M., and Barreto A., "Omnidirectional Range-Finding Sonar System for Human Navigation." Proceedings of the 2002 Florida Conference on Recent Advances in Robotics (FCRAR 2002), Miami, FL., May 23-24, 2002. (CD-ROM format).

[5] Dodson, A H; Moon, G V; Moore, T; Jones, D; *Guiding Blind Pedestrians with a Personal Navigation System*, Journal of Navigation, Vol 52, No 3, ISSN 0373-4633, pp 330-341, September 1999.

[6] Kuhn G. F., "Model for the Interaural Time Differences in the Azimuthal Plane", J. Acoust. Soc. Am., vol. 62, No. 1, pp. 157-166, 1977.

[7] Loomis J.M., Golledge R.G., and Klatzky R.L., "Navigation System For the Blind: Auditory Display Modes and Guidance", Presence, Vol. 7, No. 2, April 1998, Pp. 193 – 203.

[8] Microchip Inc., "PIC16C63A/65B/73B/74B Datasheet", document No. DS30605C, 2000.

[9] Savage innovations Inc., "Object-Oriented Programmable Integrated Circuit – programmer's manual", 1999-2001.

[10] Shoval S. , Borenstein J., and Koren, Y. *Auditory Guidance with the Navbelt—A Computerized Travel Aid for the Blind*. IEEE Transactions on Systems, Man and Cybernetics. Vol. 28, No. 3, 1998, pp. 459 – 467.

[11] Shoval S. , Borenstein J., and Koren, Y. *The Navbelt—A Computerized Travel Aid for the Blind Based on Mobile Robotics Technology,* IEEE Transactions on Biomedical Engineering, Vol. 45, No. 11, 1998, pp. 1376 – 1386.

[12] Ordóñez C., Barreto A. and Gupta N., "2-D Auditory Mapping of Virtual Environment Using Real-Time Head Related Transfer Functions", Proceedings of IEEE SoutheastCon 2002, April 2002, Columbia, South Carolina, pp. 364-369.