Key words:
psychoacoustics, blind, electronic travel aid, sound code

Michał BUJACZ*, Paweł STRUMIŁŁO*

STEREOPHONIC REPRESENTATION OF VIRTUAL 3D SCENES – A SIMULATED MOBILTY AID FOR THE BLIND

A computer program for the simulation of an Electronic Travel Aid (ETA) for the blind is presented in the paper. The program incorporates different sound coding methods and serves as an experimental bench for the ETA project conducted at the Medical Electronics Division, Technical University of Łódź. The developed software, implemented in C++ and OpenGL, was tested in a number of trials with ten sighted volunteers from a broad range of ages (14-52). After preliminary training they participated in a series of sessions aimed at testing different space representation coding schemes. Obtained results have provided valuable insights into the concepts employed for construction of ETAs based on auditory displays.

1. THE PROBLEMS OF BLINDNESS AND SOLUTIONS OFFERED

Blindness is one of the most serious handicaps a human being can experience. Amongst the most severe problems related to blindness is the loss of independent mobility. For a blind person to regain some of this ability issues such as obstacle detection and avoidance, as well as awareness of one's own destination need to be dealt with. Only in the last 60 years, considerable amount of effort went into researching the skills and tools to aid the blind in these tasks. The latest of these tools are called electronic travel aids (ETA) and utilize modern technology in the service of the handicapped. Such an aid is currently under development at the Institute of Electronics at the Technical University of Łódź (*www.eletel.p.lodz.pl/sv*).

The statistics are that in Europe 3-4 persons in every 1000 are visually handicapped. The number of the blind in the world is predicted to increase by about 20% within the next 50 years, solely due to the aging demographic [4].

* Technical University of Łódź, Institute of Electronics, Wólczańska 211/215, 90-924 Łódź

1.1. EXISTING ELECTRONIC TRAVEL AIDS

An interesting fact is that the first ever electric mobility device for the blind was invented by a Polish scientist - Kazimierz Noiszewski in 1897. It was called the Elektroftalm and utilized the photoelectric properties of Selenium cells to generate sounds of intensity proportional to the brightness of whatever it was pointed at.

The first modern ETAs appeared in the 1950s. They can be divided into two main groups: obstacle detectors and imaging systems. The first group utilizes sonar or laser sensors to detect obstacles in certain directions or proximity. The second group utilizes video or sonar input to present sound or tactile images of the environment.

The obstacle detecting canes such as the LaserCane [5], UltraCane [6] or the K-sonic torch [7] utilize sonar or laser input and vibration or audio output. These devices require little training and extend the usefulness of a cane. However, they require a lot of manual scanning to help perceive the layout of the scene.

The environmental imaging systems that generate complex audio responses require a lot of conscious effort from the user. However, after prolonged training a blind person can learn to "see" through the use of the sound code. Some examples of successfully implemented sound codes include: multiplication of an actual stereo sonar response into the audible domain (Sonicguide [7]), coding position and brightness of pixels into frequency and amplitude components of sound sweeps (The vOICe [8]), and producing sound sweeps of frequency proportional to obstacle distance (NavBelt [1]).

2. THE SOUND CODE CONCEPT AND SIMULATION SOFTWARE

2.1. SOUND CODE BASICS

Sound is a much narrower medium than vision. The primary parameters available for space representation are: the fundamental frequency (tone/pitch), amplitude (of two separate channels to create a basic stereo effect), timber (colour/instrument) and temporal envelope.

To avoid the problems with synthesising sounds that would be pleasant to hear, we decided to use the musical instrument library available for the General MIDI (GM) standard. The GM interface allowed for easy control of multiple stereo outputs and generation of sounds that were not overly irritating.

The fundamental frequency of a sound is used to represent the distance towards obstacles. The direct distance toward an object is coded through an inversely proportional musical tone, i.e., near objects produce a high pitched warning tone, while far away objects are coded with a less disturbing low tone. The range of tones is the seven middle major scales: from C1 (32.7 Hz) to C8 (4186 Hz).

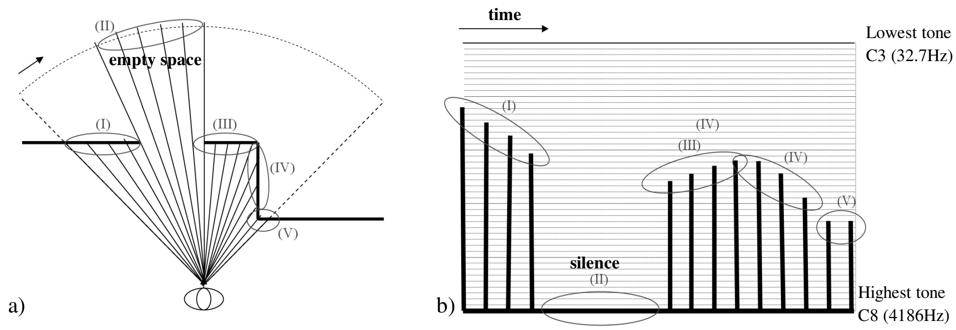

Fig. 1. The scanning concept: a) sample scene and the collection of 18 depth values corresponding to 5° scan regions. b) histogram of the sweep scan period. A user hears a sequence of tones corresponding to the depth – higher tones represent closer objects and lower tones the ones that are further away.

The minimum and maximum coded distances (within 1–8 m range) can be set by the user. This system feature is particularly useful for indoor vs. outdoor use. Interaural level difference (ILD) is applied to code the direction from which a sound arrives. A user can select one of 128 MIDI instruments for the sounds used within the code, as well as secondary instruments for the more complex sound coding modes.

2.2. SCANNING MODES

The representation of the scene is achieved through periodic scanning of 5°x5° sections of the scene. A single sound with a specific stereo location and frequency coded distance is attached to the nearest obstacle in a given conical region of the scene. Depending on the scanning mode used, the user hears a note or a sequence of notes in a set of directions. This auditory display concept is illustrated in Figure 1.

The sound cues are intertwined with short periods of silence. The most basic scan type works like an audio coded range-finder (Fig. 2a), with later added wide side-scan areas (Fig. 2b). A different instrument can be assigned to the secondary side-scans.

The next scanning mode utilizes a horizontal sweep (Fig. 2c). A user hears a sequence of up to 18 tones corresponding to the distance to the nearest obstacle in each 5°x5° conic region. The stereophonic balance smoothly changes from extreme right to extreme left over the course of the sweep.

As an attempt to provide 3D spatial information on the scene layout, a triple horizontal sweep was implemented (Fig. 2d). Each of the sweeps is assigned a different instrument.

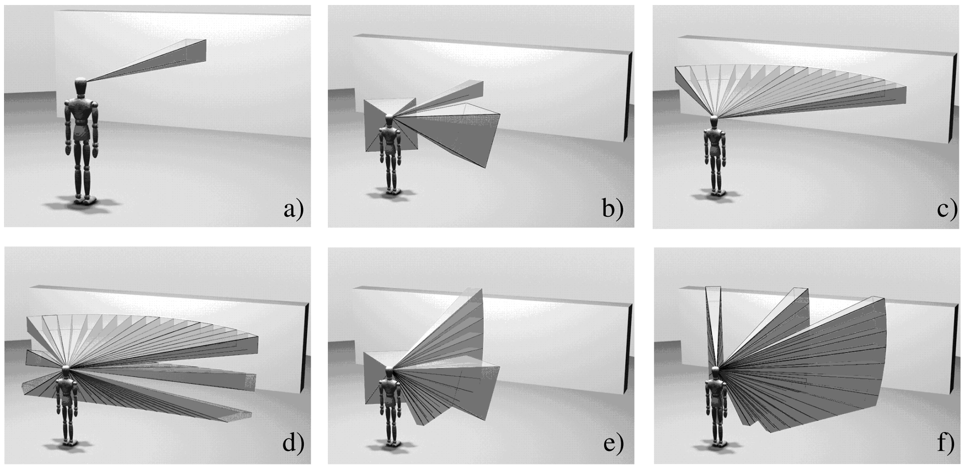

Fig. 2. Scanning modes implemented in the simulation software.

The horizontal sweeps provide information needed for scene perception, however for mobility purposes, information on low level obstacles, as well as on curbs or drop-offs is more useful. The single vertical sweep mode (Fig. 2e) is used with the secondary side-scans for better safety. The triple vertical sweep (Fig. 2f) is an attempt on covering a larger portion of the scene within one period of a scan. The side sweeps are audible only in the left or right stereo channels.

2.3. THE SOFTWARE AND OTHER SIMULATION FEATURES

The simulation software was written in C++ under Visual Studio .NET and using the OpenGL 3D graphics library. All sounds were produced by sending messages directly to the sound card's standard integrated General MIDI port.

The program displays virtual scenes composed of small 20cm x 20cm blocks. This made the code creation easier and enabled many simulation features. Various display options were available, the program could draw only the blocks within maximum range, or those that were currently sound coded, and/or those within cane and/or hand range.

In all of the scanning modes, the floor could be "turned on/off", i.e. treated similarly as other obstacles or as an empty space.

Different scenes could be loaded from simple text files, which held coordinates of every wall's two opposing corner blocks. An idealized depth map which could be constructed using stereoscopic cameras (shown in lower left corner of Fig. 3) is used in the generation of the sound code.

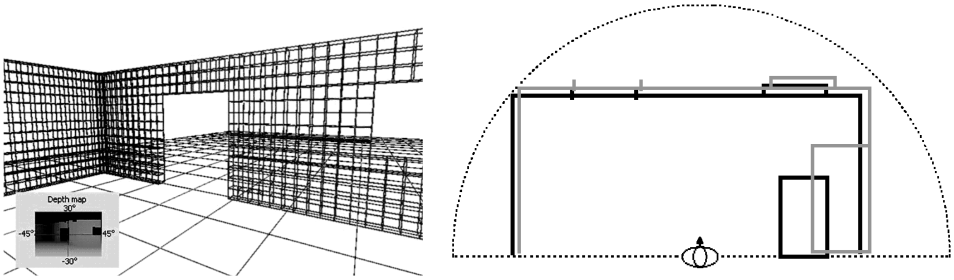

Fig. 3. A 3D scene used in the trials (left) and one of the trial result drawings (right). The grey line is the layout drawn by the test participant, black - the actual scene layout, dashed - device range.

3. THE EXPERIMENTAL TRIALS

The tests were conducted on a standard home PC. Each participant chose from three types of headphones: small earplugs, medium open-air, or large leather closed-air.

There were 10 volunteers: five male, five female of ages 14 to 52. Each of them was trained for two to four hours in the use of the simulation software and the various scanning modes. The training consisted of stages in which different program features were explained and scanning modes demonstrated in the order of increasing complexity.

The volunteers participated in three types of trials: mobility, orientation (observation of a scene's layout) and a combination of both. They were free to choose two scan modes they thought most suitable for the tasks. All of them chose the horizontal sweep as the best choice for observing scene layout. Eight out of ten chose the single vertical sweep for safely and quick movement.

The mobility trials consisted of a simple labyrinth of a series of parallel walls which were perpendicular to a user's direction of movement. The trial volunteers walked through three very similar scenes, first "by touch" - seeing only blocks within cane range, then using the two chosen scanning modes, and finally with full visibility of one of the scenes. The time taken to walk through the scenes was recorded for later comparison.

The orientation trials consisted of two simple scenes in which the users were not to move, only look around from a single location. After the allotted time of three minutes each person drew the perceived scene on a special form.

The combination trial consisted of a single task: walking into a doorway in an alleyway given a simple map of the scene layout. The users were warned that there will be obstacles present in the scene that were not marked on the map. After the first two runs, the users were to draw the unmarked obstacles on the map.

4. RESULTS AND CONCLUSIONS

The mobility trial times varied greatly among the ten participants, mainly due to their different experience with moving around virtual 3D scenes. Almost all participants agreed that with prolonged use they could learn to trust the audio code.

In the orientation trials almost all participants were able to draw the scenes they observed with much detail (see Figure 3).

Users had no problems in navigating the scene in the combination trials. All participants correctly identified the additional obstacles - a street sign and a car.

The experiments showed that the designed scene representation methods are potentially useful and provide valuable insight for future sound code development. It was the opinion of the volunteers that simpler scene representation modes were much more useful than the complex ones and the horizontal sweep seemed most instinctive to almost all participants, although being too slow for mobility purposes.

REFERENCES

[1] BORENSTEIN J., *The NavBelt: A Computerized Travel Aid for the Blind,* University of Michigan website, *http://www-personal.engin.umich.edu/~johannb/navbelt.htm*

[2] BREGMAN S., *Auditory Scene Analysis: the Perceptual Organization of Sound*, A Bradford Book, The MIT Press, Cambridge, Massachusetts, 1990

[3] BUJACZ M., *Stereophonic representation of virtual 3D scenes – a simulated mobility aid for the blind,* Master's Thesis, Technical University of Lodz, 2005

[4] European Blind Union website, *http://www.euroblind.org/*

[5] Nurion-Raycal: Electronic Travel Aids for the Blind website, *http://www.lasercane.com/*

[6] SOUND Foresight Ltd website, *http://www.soundforesight.co.uk/*

[7] *The 'K' Sonar Cane - a history*, The BAT 'K' Sonar-Cane website, *http://www.batforblind.co.nz/*

[8] The vOICe – Seeing with Sound website, *http://www.seeingwithsound.com/*

STEREOFONICZNA REPREZENTACJA WIRTUALNYCH SCEN 3D - SYMULACJA URZĄDZENIA WSPOMAGAJĄCEGO NIEWIDOMYCH

Zaprezentowany program komputerowy symuluje pracę elektronicznego urządzenia wspomagającego niewidomych, demonstrując różne metody przedstawiania trójwymiarowych scen za pomocą stereofonicznego kodu dźwiękowego. Program został przetestowany na grupie dziesięciu ochotników, którzy po kilkugodzinnym szkoleniu wzięli udział w serii testów sprawdzających skuteczność, z jaką mogą odtworzyć udźwiękowione sceny za pomocą rysunków, oraz poruszać się pośród niewidocznych przeszkód. Rezultaty testów potwierdziły użyteczność zaproponowanych metod i dostarczyły wniosków pomocnych w realizowaniu projektu konstrukcji urządzenia wspomagającego niewidomych w Instytucie Elektroniki Politechniki Łódzkiej.