Sonification of Range Information for 3-D Space Perception

E. Milios, B. Kapralos and S. Stergiopoulos {eem, billk}@cs.yorku.ca

Department of Computer Science, York University, North York, Canada M3J 1P3

Summary: We present a device that allows 3-D space perception by sonification of range information obtained via a point laser range sensor. The laser range sensor is worn by the user, who scans space by pointing the laser beam in different directions. The resulting stream of range measurements is then converted to an auditory signal whose frequency or amplitude varies with the range. Our device differs from existing navigation aids for the visually impaired. Such devices use sonar ranging whose primary purpose is to detect obstacles for navigation, a task to which sonar is well suited due to its wide beam width. In contrast, the purpose of our device is to allow users to perceive the details of 3D space that surrounds them; a task to which sonar is ill suited, due to artifacts generated by multiple reflections. Preliminary trials demonstrate that the user is able to accurately detect corners and depth discontinuities with ease and to perceive the size of the surrounding space.

INTRODUCTION

The visually impaired rely on non-visual senses, primarily hearing, to help them locate and identify objects within their immediate and distant environment. Although all of the non-visual senses are able to convey information pertaining to an object (i.e. texture, temperature, and size), only the auditory system is capable of providing significant distance cues, which may be coded through intensity, echoes, reverberation, doppler shift, etc. The Sonic Pathfinder [2] and SonicGuide [3] are examples of Electronic Travel Aids (ETAs) which sonify range information allowing object detection/avoidance by the visually impaired user. With both ETAs, range is mapped directly to pitch. In addition, both ETAs provide a directional cue. With the SonicGuide, sounds associated with objects to the left or right of the user will be louder in the respective ear. In the case of the SonicPathfinder, sounds associated with objects to the left or right of the user will only be heard in the respective ear.

To determine the distance to an object, most ETAs emit an ultrasonic sound and measure the time it takes for the echo to return. Due to the wide panoramic field of view, small range and reflection artifacts inherent with sonar, the use of sonar based devices is limited to localizing objects within an immediate and uncluttered environment [1]. They fail to provide sufficient resolution and artifact-free range data necessary for the perception of 3D space.

Unlike sonar, laser ranging allows for a greater range. Its narrower beam and shorter wavelength combine to detect finer detail necessary for shape and pattern perception. A laser ranging device has provided range measurements for an autonomous robot [4].

Rather than object detection, which sonar is better suited for, we decided to examine the sonification of infrared range measurements in order to perceive shapes and patterns (e.g. detection of corners, stairs, and depth discontinuities) in the user's environment. The ultimate goal of our work is to develop a device that will enable the perception of 3D space by the visually impaired user.

DESCRIPTION

Figure 1 below illustrates the range data sonification process. The users scan their environment and obtain eight range measurements per second (one measurement every 125ms), with the laser range finder (LRF). Measurements are mapped to a particular feature in the audio domain (see below). One feature we experimented with is pitch (frequency of the audio signal), in the form of MIDI (Musical Instrument Digital Interface) notes. Another feature experimented with is loudness (velocity of the MIDI note). Finally, the MIDI note is output for 90ms (e.g. there is a "small gap" of silence between successive notes), using one of the 128 instruments available with the Quick Time Music Architecture (QTMA) software synthesizer.

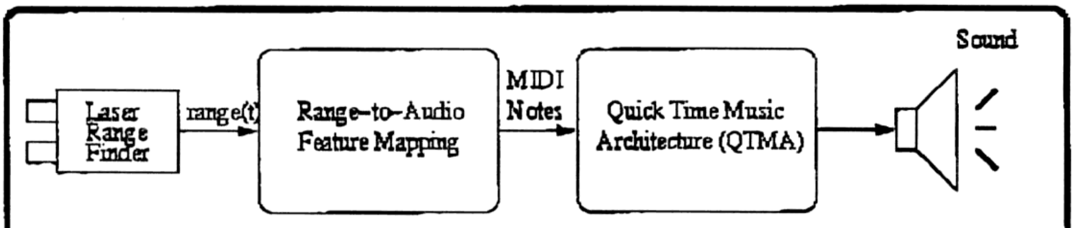

Figure 1
Sonification of Infrared Range Data

Two modes of operation have been defined. The *proportional* mode (range to audio feature mapping) provides the user with a general knowledge regarding both the close-by and distant surroundings through the detection of objects, doorways, depth discontinuities, etc. The *derivative* mode (temporal derivative of range to audio feature mapping) is used to obtain the greater detail necessary for shape and pattern detection.

THE MAPPINGS

We now elaborate on the various mappings used, summarized in table 1.

PROPORTIONAL MODE

TABLE 1: A summary of the range to audio mappings used.

Audio Domain	Mode			
	Proportional		Derivative	
	Range to Audio	Punctuation	Frequency Piecewise Constant	Frequency Corresponds to Musical Scale
Velocity (V)	$V = a / range^b$ where: a, b are constant	Low frequency drum cymbal to indicate depth discontinuity	$V = a * (\Delta range)^b$ where: a, b are constant	

Frequency (F)	if range $\geq$ min_range:
	$F = kc^{-(range-min\_range)*a}$
	else
	F = 440Hz
	where:
	a, c, k are constant
			No. Change: F_1
			Pos. Change: F_2
			Neg. Change: F_3
			where:
			F_1 = 2093Hz
			F_2 = 3951Hz
			F_3 = 261Hz
				No. Change: MIDI note 54
				Positive Change: MIDI notes 25 – 49
				Negative Change: MIDI notes 61 - 84

While in the proportional mode, range measurements are mapped inversely to pitch (frequency of the MIDI note), using a logarithmic mapping. In addition to producing the most favorable results, the logarithmic mapping allowed for a change in pitch throughout the entire distance range. Best results have been achieved when using the following constant values: k = 4000Hz, c = 2.718 (the base of the natural logarithms), a = 0.15 and r = 0.3m. Furthermore, range measurements were also mapped to loudness (velocity of the MIDI note), using a modified version of the inverse square law (a = 75 and b = 0.3).

Depth discontinuity, or, a sudden large change in depth, is defined as the difference between the present and previous range measurements greater than a pre-defined value. Such information was conveyed to the user utilizing one of the instruments of the QTMA. Best results were achieved using a low frequency drum cymbal.

DERIVATIVE MODE

While in the derivative mode, the change between consecutive range measurements is mapped to the audio domain. Such a mapping provides greater information allowing for the detection of shapes/patters, details of which are not necessarily noticeable in the proportional mode.

Two forms of this mapping were experimented with. Both mappings considered whether the change in range measurements is positive or negative. A positive change indicates the location being measured is farther away from the user relative to the last location (i.e. the current measurement is greater than the previous measurement). A negative change indicates that the location being measured is closer to the user (i.e. the current measurement is less than the previous measurement). With the first mapping *(frequency piecewise constant)*, MIDI note 96 (F_1 = 2093Hz), at a low velocity was output when there was no change between consecutive measurements. MIDI note 107 (F_2 = 3951Hz), was output when the change was negative, while MIDI note 60 (F_3 = 261Hz), was output for a positive change.

The second mapping (*frequency corresponds to the musical scale*), takes advantage of most peoples familiarity with the notes of the musical scale [2]. MIDI note 54 at a low velocity was output when there was no change between consecutive range measurements. Positive changes in range (e.g. moving farther) were mapped to the two octaves beginning with MIDI note 25 and ending with MIDI note 48. Negative changes in range were mapped to the two octaves beginning with MIDI note 61 and ending with MIDI note 84. The octave containing MIDI notes 49 – 60 separates the two regions allowing the user to easily discriminate as to which region an output note belongs to.

With both mappings, the change in range measurements was mapped inversely to the velocity of the MIDI note using a modified version of the inverse square law. Best results were achieved using a value of $a = 100$ and $b = 0.45$.

DISCUSSION

Informal lab surveys suggest that the output produced using the QTMA is both pleasant to the ear and effective in conveying information. Several subjects were able to quickly locate doorways and other depth discontinuities while in the proportional mode, with only basic instructions and no training. In the derivative mode, subjects were able to detect corners, flat level surfaces and depth discontinuities. Some training was required however.

ACKNOWLEDMENTS

The work was supported by a research grant from the Natural Science and Engineering Council of Canada (NSERC) and a York Science Summer Research scholarship to B. Kapralos. We thank Prof. Patrick Dymond for suggesting the use of MIDI, Prof. Laurence. Harris for suggesting the use of derivative mappings and Robert Arrabito of DCIEM for his valuable comments and his encouragement. Greg Reid provided help with audio and signal processing facilities.

REFERENCES

[1] Easton. D. R., *Optometry and Vision Science* **69**, 3 - 14 (1992).

[2] Heyes, D. A., *Journal of Visual Impairment and Blindness* **77**, 200 – 202 (1984).

[3] Kay, L., *Journal of Visual Impairment and Blindness* 77, 12 – 16 (1984).

[4] Nickerson S. B, P. Jasiobedzki, D. Wilkes, M. Jenkin, E. Milios, J. Tsotsos, A. Jepson, O. N. Bains, Robotics *and Autonomous Systems*, April (1998).