Using our Ears:

An Introduction to the Use of Nonspeech Audio Cues

Bill Buxton

Computer Systems Research Group
University of Toronto
Toronto, Ontario
Canada M4L 1X6
tel: (416)465-9930
Fax: (416)-465-3836

Introduction

Our ears provide invaluable help in dealing with the complexities of the everyday world. We have highly developed skills in what we might call "everyday listening" and "musical listening." But while these skills are heavily relied upon in everyday tasks such as driving and crossing the street, they are virtually ignored when it comes to interacting with computers or in traditional methods of analyzing complex data.

This lack of use of the audio channel is a waste that we can ill afford. As the complexity of information presented to us by computers grows, so does our difficulty in its asimilation. Our ears can help.

In the presentation that follows, we discuss how this is so. We begin by discussing some of the different ways in which nonspeech audio can be used. We then illustrate each by way of examples, with the emphasis being on the presentation and analysis of complex data. Some underlying theory is then discussed. Finally, we give some practical pointers as to how one might proceed to incorportate sound into one's computational environment.

Readers interested in reading more about the use of nonspeech audio are referred to a special issue of the journal *Human-Computer Interaction*, that is devoted to the topic (Vol.1, No. 4, 1989).

Classes of Audio Cue

Functionally, nonspeech audio messages can be thought of as providing one of three general types of information: alarms and warnings, status and monitoring indicators, and encoded messages. Typically, different types of audio cues are used for each.

Alarms and warning messages are signals that take priority over other information. Their purpose is to interrupt any ongoing task and alert the user to something that requires immediate attention. They normally only sound in an "exception" condition. They are usually loud, easily identifiable sounds with sharp transients. An ambulance siren would be one such example.

Status and monitoring messages provide information about some ongoing task. The nature of such cues depends upon the type of task being monitored.

The key click produced when typing on a conventional keyboard is one example of how audio cues can provide status feedback for short discrete tasks. Normally in typing, the sound cue only indicates whether the key has been pressed or not. However, Monk (1986) showed that one can go beyond this. In an experimental situation, he showed how mode errors could be significantly reduced by having the pitch of the sound associated with each keystroke depend on which of two modes the system was in.

For ongoing continuous tasks, sounds providing status information are usually sustained tones or repeating patterns that are audible for the duration of the process that they are monitoring. In such cases, unlike alarms, these messages are designed to fade rapidly into the background of the operator's consciousness, so that attention can be directed to some other foreground task. They are designed to come back into the foreground only when there is a significant change in the process being monitored.

The design of this type of message exploits the fact that the human perceptual system does not remain conscious of steady-state sounds. In contrast, it is very sensitive to change. Hence, if a steady-state sound representing an ongoing background task stops, then that transition will bring the fact of a change in state to the user's attention. The sound of a washing machine turning off is one such example. In the driving example, any change in the normal background sound of the car motor is another.

Humans are capable of monitoring more than one such signal in the background, providing that the sounds are appropriately differentiated. As with alarms, however, although a number of different messages can be individually discriminated and understood, if more than one simultaneously requires attention, then it is likely that the user will become confused and performance will be affected. An actual case in which this was evident was the Three Mile Island power plant crisis. There, over 60 different auditory warning systems were activated (Sanders & McCormick, 1987, p. 155). This example illustrates that although we can recognize and simultaneously monitor a number of different audio cues, we can normally only respond to one or two at a time.

Encoded messages are used to present numerical (or quantitative) data, such as statistical information, in patterns of sound. For example, Lunney, et al. (1983) used audio motives to present spectral information of various chemicals to blind students. The complex and varying sounds used here contrast with the penetrating one or two sounds used with alarms or with the steady-state tones or patterns used in status monitoring.

The design of this class of message often exploits our capabilities of pattern matching and recognition. In some cases, such messages are much like musical melodies. The usage has a lot in common with Wagner's use of leitmotiv, Procofiev's use of motives to represent the characters in Peter and the Wolf and the sounds in the video game PACMAN.

Blattner, Sumikawa, and Greenberg (1989) is especially concerned with this type of message. The authors discuss issues concerning how to exploit our experience with music listening to encode information, based on cues such as loudness, timbre, rhythm, etc.

There has been some interesting work in using audio to encode quantitative information. Much of it (Bly, 1982a, 1982b; Mansur, 1984; Mezrich, Frysinger, & Slivjanovski, 1984; Yeung, 1980) has to do with techniques for using audio cues to aid in analyzing and presenting statistical data in what could be called a sound graph. With the exception of Mansur's work, all of these studies were motivated by the need to present more information than visual techniques could offer. In some cases, there has also been the motivation to find new ways to present information in a form accessible to the blind.