Proceedings of the Institute of Acoustics

AUDITORY ICONS

W. W. Gaver

Rank Xerox Cambridge EuroPARC

ABSTRACT

Auditory icons are everyday sounds meant to convey information about computer systems by analogy with everyday events. They represent one strategy for the creation of audio messages, which are nonspeech sounds that provide some or all of the output from a computer. Audio messages differ from auditory warnings in that they are seldom meant to warn people, and are instead aimed at conveying complex information in natural and intuitive ways. The purpose of this paper is to briefly review work on audio messages, with particular attention to research on auditory icons. The use of auditory icons is illustrated by the SonicFinder, which is an interface I developed at Apple Computer, Inc.

INTRODUCTION

Auditory icons are everyday sounds meant to convey information about computer systems by analogy with everyday events [1,2,3]. They represent one strategy for the creation of *audio messages*, which are nonspeech sounds that provide some or all of the output from a computer.

An example of an auditory icon is a reverberant hitting sound used to indicate the arrival of electronic mail in a system. The force of the collision might represent the size of the message, the material of the dropped object might represent the type of message, and the amount of reverberation might represent free disk space in the system. Such sounds can convey complex system information in a way that is intuitively accessible and unobtrusive.

Auditory messages are much like auditory warnings – except they are not meant to warn. Instead, they are meant to provide information that is relevant but typically noncritical. Traditionally, auditory warnings have been the most prevalent use of nonspeech sounds to represent encoded information to listeners. The primary purpose of such sounds is to alert listeners to the existence of a situation that requires their attention. Secondary goals are to use sounds that indicate exactly what that situation might be, and to design sounds that are not so annoying that they distract listeners or prevent useful communication [4]. Auditory icons, in contrast, are designed primarily to convey complex information clearly and naturally. A secondary goal is to make them unobtrusive, because if they are annoying or distracting users will simply stop using them. Only rarely are auditory icons meant to alert users to a possibly critical situation.

Research on auditory messages and warnings converges in seeking to convey encoded information via sound, information that is often fairly complex and must be readily recognized. Each kind of research is concerned with the degree to which a sound will attract listeners' attention -- for warnings, perceived urgency is at issue; for messages, subtlety is desirable. Auditory icons are an attempt to provide natural, non-annoying audio feedback in a workstation environment.

Author's present address: William W. Gaver, Rank Xerox Cambridge EuroPARC, 61 Regent Street, Cambridge CB2 1AB, England.

The purpose of this paper is to briefly review work on audio messages, with particular attention to research on auditory icons. The use of auditory icons is illustrated by the SonicFinder, which is an audio interface I developed at Apple Computer, Inc.

AUDIO MESSAGES

Work on audio messages can be roughly divided into three classes. The first is concerned with using sound to represent multidimensional data. The second focuses on the potential for sound to help the visually impaired gain access to computer facilities. The final use of sound is to complement and supplement graphics as a means of conveying complex system information to users.

Early work on audio messages focused on the use of sound as a tool to aid researchers explore multidimensional data. *Sound graphs*, a term suggested by Mansur [5], is a good descriptive name for this sort of work. Typically, each dimension of data is mapped to a dimension of sound. Each multidimensional data point, then, specifies a new sound (or the parameters of a continuous sound over time). Patterns in the data manifest as patterns in the sound, so that, for instance, all high sounds are loud and all low sounds soft. Using this strategy, Bly [6] showed that subjects could reliably distinguish members of two multidimensional, normally distributed data sets after listening to exemplars of each; Mezrich et al. [7] found that subjects could perceive correlation in many dimensional data; and Morrison and Lunney [8] showed that blind subjects could learn to recognize spectrographic data mapped to chords and series of chords.

Other work on audio messages has explored their potential for giving visually disabled people greater access to computers. Morrison and Lunney's [8] spectrographic display system was primarily aimed at enabling visually disabled students to use sophisticated chemical analysis tools. Edwards [9] developed an auditory interface to a modified direct manipulation word processor that was usable by both blind and visually impaired subjects. He used a combination of simple pitch encoding and speech to provide information about location in the system and results of menu selections. The wide relevance of such work is emphasized by Buxton's [10] observation that we are all visually impaired when faced with a cluttered graphic display.

A third use for audio messages is to provide general information to users of computer workstations. The aim is to use sounds in ways that are analogous to visual icons. Auditory icons can provide information that is redundant with that displayed graphically, as well as information that is difficult or impractical to display visually. In addition, using sound in computer interfaces can help encourage feelings of direct engagement with the model world of the computer [11].

One goal of such work is to develop systems of related messages so that new messages can be created and understood on the basis of old ones. Blattner et al. [12], for instance, suggest that a hierarchical system of messages can be developed by using variations on a small set of musical motives. Each resulting musical phrase can then stand for a separate message, with similar sounding motives representing similar messages. In contrast, auditory icons use the physical parameters of sound producing events to represent attributes of system events. Although these two strategies differ considerably, they are not incompatible. Both may be used in a single system, and hybrids of the two methods may also be possible.

Auditory Icons

The notion of using everyday sounds as auditory icons is based on the observation that, in general, we listen to the world, not the sounds it makes. That is, in hearing a sound while walking down a street I am, hopefully, more likely to be aware that it is made by a large car accelerating in my direction than to notice to its pitch, loudness and timbre. Pitch, loudness, timbre and so on are perceptual attributes of the sound itself; their experience is one of *musical listening*. Size, force, bearing and so on are perceptual attributes of a sound's source; their experience is one of *everyday listening* [2].

Distinguishing the experiences of musical and everyday listening is useful in that it allows the development of a new framework for understanding sound and hearing. Such a framework is based on the correlations between dimensions of experience and dimensions of sound producing events, rather than between dimensions of experience and sound itself. Thus one can talk about the auditory perception of source attributes such as materials, forces and sizes, and not only the perception of sound qualities such as pitch, loudness, and timbre. Understanding everyday listening is a difficult task, and relatively little research has been directly concerned with how we hear events (though see [13] - [17]). Nonetheless, the framework suggested by such research can be applied to conveying information with sound.

Auditory icons, then, are everyday sounds used to convey information about computer systems by analogy with everyday events. Instead of using musical parameters of sound such as pitch, loudness and timbre to encode data, auditory icons are created by using source attributes such as material, size and force. So, for example, selecting a lengthy text file in the SonicFinder makes a sound like a large piece of wood being hit, whereas a small application sounds like a small piece of metal.

There are several advantages to this strategy of creating audio messages. Auditory icons can often represent computer events in a less arbitrary manner than can musical messages. This is because they are iconic representations of events in the model world of a computer -- that is, they are the sorts of sounds these events would make in the everyday world (mapping sound to information is discussed at some length in [1]). Because of this non-arbitrary mapping between events and sound, auditory icons can be created which are compatible with graphical icons: Things in the interface can sound like what they look like. This reinforces the mapping that leads to a consistent user model of an interface. Because of this consistency, and because auditory icons are the kinds of sounds one hears in everyday environments anyway, they are less likely to be distracting or annoying than are musical sounds.

Families of auditory icons can be created in systematic ways by using parameters of everyday sounds consistently to represent different aspects of computer events. For example, a computer object's type might be represented by the material of a sound-producing object. Then all events involving text files might make wooden sounds, for instance, while events that involve applications might make metallic sounds. Deletions might always make crashing sounds, selections could always be indicated by hitting sounds, and so forth. In such a system, selecting a file would make a hit wood sound; selecting an application would make a hit metal sound; deleting a file would make a crashing wood sound; and deleting an application would make a metallic crashing sound. In this way, a system of auditory icons can be created in which similar events make similar sounds.

THE SONICFINDER

Auditory icons are demonstrated in the SonicFinder, which is an audio interface for the Macintosh that I developed at Apple Computer, Inc. In this interface, auditory icons are associated with many common events such as selecting files, opening and scrolling windows, and deleting objects.

The SonicFinder is implemented in the form of functions called from the original Macintosh Finder code. The Finder is the application that is run automatically when the Macintosh is started. Organized by analogy with a desktop, it provides a visual representation to items of interest in the interface (e.g., files, folders, disks, etc.) and allows users to manipulate them (e.g., to move, copy, or delete files). The current SonicFinder is contained in an "init" file called *Finder Sounds* which is called by the Finder when the system is booted.

The SonicFinder uses the information that is available in the existing interface to trigger and control the playback of sounds sampled from recordings of everyday sound-producing events. Figure 1 summarizes the events in the SonicFinder that are mapped to sounds. Note that in general each auditory icon is determined by several different parameters of the event. For example, selecting an object always makes a hitting sound, the type of object is encoded by the material that is hit, and its size by the frequency of the sound -- which corresponds to the size of the struck object. In this way each auditory icon conveys several sorts of information in a systematic manner.

FINDER EVENTS	AUDITORY ICONS
Objects	
Selection	Hitting sounds
Type (file, application, folder, disk, trash)	Sound source (wood, metal, etc.)
Size	Frequency
Opening	Whooshing sound
Size of opened object	Frequency
Dragging	Scraping sound
Size	Frequency
Where (windows or desk)	Sound type (bandwidth)
Possible Drop-In?	Selection sound of disk, folder, or trashcan
Drop-In	Noise of object landing
Amount in destination	Frequency
Copying	Pouring sound
Amount completed	Frequency
Windows	
Selection	Clink
Dragging	Scraping
Growing	Clink
Window size	Frequency
Scrolling	Tick sound
Underlying surface size	Frequency
Trashcan	
Drop-in	Crash
Empty	Crunch

Figure 1. Events and the auditory icons used to represent them in the SonicFinder.

Describing an auditory interface in writing is difficult, and listing all possible events and the sounds they make is tedious and relatively uninformative. A example of a interaction with this interface seems more useful. Figure 2 shows a typical interaction with the SonicFinder, in which a file is selected, dragged to the wastebasket and thrown away.

In Figure 2A, the file is selected by the user and makes a noise like something being tapped. Because it is a text file, it sounds like wood -- if it had been an application, it would have sounded like metal; a folder would make a sharper paper-like sound; disks make a hollow metal sound (like a container being tapped); and the wastebasket a different hollow metal sound. In the Finder, applications and text files have widely varying icons and are easily confused. In the SonicFinder, they are are readily differentiated by the sounds they make.

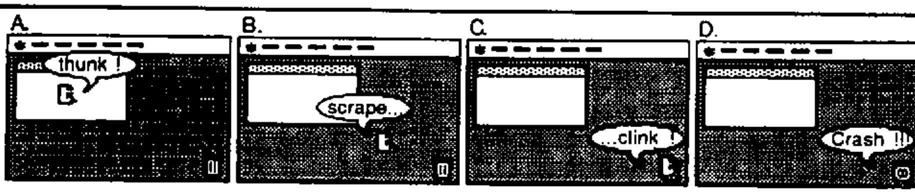

Figure 2. A typical interaction with the SonicFinder.

The sound made by a selected object also depends on its size. Large objects make lower sounds than small objects, and empty disks make lower sounds than full ones (as is typically true of objects in the everyday world). Size is not usually displayed in the Finder, and in this example is only available from the auditory icon.

After the file is selected, it is dragged towards the wastebasket (Figure 2B). A simple scraping sound whose frequency depends on the size of the dragged object is played continuously while in this mode. The bandwidth of the sound is changed while dragging to indicate whether the object is over the window in which it was selected, another window, or the desktop. This represents a partial solution to the problem of indicating when a drag will result in a move or a copy.

When the file is moved over the wastebasket (Figure 2C), the dragging sound stops and the wastebasket's selection sound is played. More generally, when an object is dragged over a container into which it may be dropped (i.e., a folder, disk, or the wastebasket), that object makes its sound and the scraping sound stops. This sort of auditory confirmation that a target has been hit is one of the more obviously useful features of the SonicFinder.

Finally, when the file is dropped in the wastebasket (Figure 2D), the sound of shattering dishes provides satisfying feedback that it has been marked for deletion. When the deletion actually occurs, a crunching sound is played to indicate the destruction of the object.

This sample interaction illustrates a number of aspects of the SonicFinder. First, many sounds accompany the use of this interface. This results in the creation of a fairly constant auditory environment in which no single sound is particularly incongruous or distracting. Second, the sounds seem to fit well with the events they represent. As suggested before, this seems due to the direct nature of the mapping between the sounds and the events they represent. Finally, in this interface auditory icons convey information that the graphic portion of the interface either does not display (e.g., about file size), displays less effectively (e.g., about dragging over a container), or in a less satisfying way (e.g., about deleting an object).

The sound made by copying is another example of an auditory icon which is sometimes more useful than its graphical counterpart (see Figure 3). When a file is copied, the sound of pouring water accompanies the event. The frequency of the sound is increased to indicate progress by analogy with the sound a container makes as it is being filled. In current versions of the Finder, progress is also indicated graphically by a "dial bar," a horizontal rectangle which is progressively filled with a pattern to indicate the percentage completed. Thus the auditory icon is redundant with the visual display. But the graphic indicator requires the user to attend to the screen, while the sound does not. During lengthy copy operations, the advantage of using sound to represent progress is obvious and substantial.

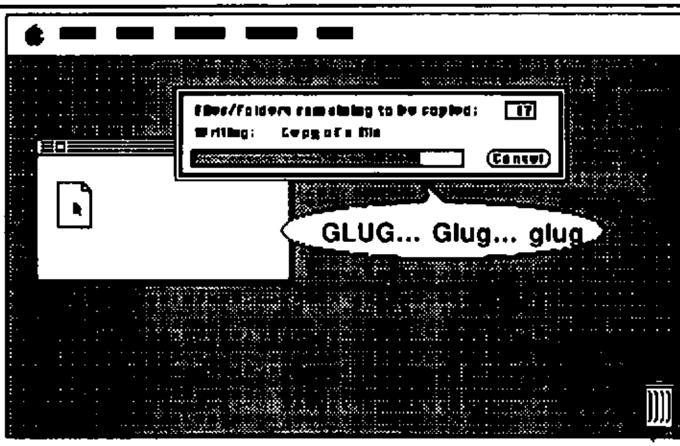

Figure 3. Copying makes a sound like pouring water; the frequency of the sound increases to indicate progress.

The SonicFinder is a working interface that has been distributed internally within Apple Computer, Inc., and demonstrated to a number of audiences outside of the company. Many users prefer the SonicFinder to quiet Finders, and audience reaction is typically favourable. The ways sound works within the interface seem natural and desirable to those who encounter it.

There seem to be two kinds of advantages associated with the SonicFinder. First, it allows more flexible interactions to users, as the examples of finding the size of objects, target acquisition, and copying illustrate. Second, it seems to increase feelings of direct engagement with the model world of the computer. The addition of sounds makes that world more tangible to users, and in some cases may aid understanding of that world. Though no rigorous user testing has been performed to date on the SonicFinder, its continued use is evidence that it is a successful interface.

CONCLUSION

In this paper I have reviewed work on auditory messages, focussing on auditory icons as a viable strategy for displaying information to computer users. Auditory icons are appealing because they represent computer events in an systematic, intuitively obvious way. They can present complex information, and information that is difficult or awkward to present graphically. Finally, they can greatly increase the tangibility of the model world of the computer, as illustrated by their use in the SonicFinder.

Many possibilities for auditory icons remain that are left unexplored in the SonicFinder. In this interface, audio information is conveyed by properties of sound producing events. Other kinds of information might be conveyed by properties of a virtual auditory environment in which those events take place. For instance, reverberation, which is associated with spaciousness in the everyday world, might be used to provide information about system variables such as processor load or disk space. In addition, many of the benefits of auditory icons are likely to be most apparent in multiprocessing systems, and systems where collaborative work is being performed. The success of the SonicFinder attests to the potential of auditory icons as a means for creating auditory messages, but this interface represents only a first step in the direction of creating auditory environments for computers.

References

[1] Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. Human-Computer Interaction. 4 (1).

[2] Gaver, W. W. (1988). Everyday listening and auditory icons. Doctoral Dissertation, University of California, San Diego.

[3] Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. Human-Computer Interaction. 2, 167 - 177.

[4] Patterson, R. D. (1982). Guidelines for auditory warning systems on civil aircraft. Civil Aviation Authority. London, Paper 82017..

[5] Mansur, D. L., Blattner, M. M., & Joy, K. I. (1985). Sound-Graphs: A numerical data analysis method for the blind. Journal of Medical Systems. 9, 163 - 174.

[6] Bly, S. (1982). Presenting information in sound. Proceedings of the CHI '82 Conference on Human Factors in Computer Systems, 371 - 375. New York: ACM.

[7] Mezrich, J. J., Frysinger, S., & Slivjanovski, R. (1984). Dynamic representation of multivariate time series data. Journal of the American Statistical Association. 79, 34 - 40.

[8] Morrison, R., & Lunney, D. (1985). [contribution to panel on communicating with sound]. CHI '85 Proceedings. 118-119.

[9] Edwards, A. (1989). Soundtrack: An auditory interface for blind users. Human-Computer Interaction. 4 (1).

[10] Buxton, W. (1989). Introduction to this special issue on nonspeech audio. Human-Computer Interaction. 4 (1).

[11] Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1986). Direct manipulation interfaces. In D. A. Norman & S. W. Draper (Eds.), User centered system design: New perspectives on human-computer interaction. (pp. 87 - 124). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

[12] Blattner, M., Sumikawa, D., & Greenberg, R. (1989). Earcons and icons: Their structure and common design principles. Human-Computer Interaction. 4 (1).

[13] Ballas, J. A., & Howard, J. H., Jr. (1987). Interpreting the language of environmental sounds. Environment and Behavior. 19, 91-114.

[14] Freed, D. J., & Martens, W. L. (1986). Deriving psychophysical relations for timbre. Proceedings of the International Computer Music Conference,. Oct. 20 - 24, 1986, The Hague, The Netherlands.

[15] Jenkins, James J. (1985). Acoustic information for objects, places, and events. in Warren, W. H., & Shaw, R. E., Persistence and change: Proceedings of the first international conference on event perception. Hillsdale, NJ: Lawrence Erlbaum Associates.

[16] Vanderveer, N. J. (1979). Ecological acoustics: Human perception of environmental sounds. Dissertation Abstracts International. 40/09B, 4543. (University Microfilms No. 8004002).

[17] Warren, W. H., & Verbrugge, R. R. (1984). Auditory perception of breaking and bouncing events: A case study in ecological acoustics. Journal of Experimental Psychology: Human Perception and Performance. 10, 704 - 712.