Preface:

The Use of Non-Speech Audio at the Interface

INTRODUCTION

This book is directed at system designers and researchers in the field of HCI. It is devoted to the thesis that human-computer interaction can be significantly enhanced through the improved use of the audio channel. Our focus is narrow, and deals specifically with one especially neglected aspect of sound: the use of non-speech audio to communicate information from the computer to the user.

There are significant potential benefits to be reaped by developing our capabilities in the use of sound. Examples of where sound is already used extensively are process control, flight management systems, and user interfaces for the visually impaired. Even the sounds of the disk drive and printer on personal computers convey information that provides useful feedback as to the state of the system.

The objective of the book is to stimulate research in the field, and provide the historical, theoretical and practical background that will enable participants to "tool up" to undertake such work.

The book assumes an intermediate level of understanding of user interface design, and computer programming. However, no special background in psychoacoustics, audio or music is assumed.

Overview

We begin with an introductory overview, including a few representative case studies. This is followed by some necessary theoretical background on psychoacoustics and musical perception. The intent here is not to deliver a course on psychology; rather, it is to distill key issues from the literature and to discuss their theoretical and practical ramifications as applied to the use of nonspeech audio at the interface.

The next two sections of the book are sets of case studies, each discussing a particular class of use of nonspeech audio. The first set deals with the use of audio as applied to data analysis. The approach is comparable to the use of graphics in scientific visualization. The motivation is to utilize additional sensory modalities, in this case audio, to permit the user to apply a wider set of skills and resources to understanding complex phenomena. This is accomplished by mapping some or all attributes of the data to parameters of the sounds that are used to "display" them. Central to the success of the approach is the selection of sounds, the mapping of the parameters, and the human perceptual system. These issues are discussed, using a number of relevant case studies from the literature.

The second set of case studies deals with embedding nonspeech audio directly into the user interface. Here, the focus is on expanding the vocabulary of user interfaces from windows, icons and mice to include sound. This could be characterized as a move from graphical user interfaces (GUI's) to perceptual user interfaces (PUI's). A number of case studies are discussed. Some, not surprisingly, deal with developing user interfaces for the visually disabled. Part of the emphasis is on showing how such work is equally applicable to the so-called "able bodied" user.

A major bias in the book, and in our approach to user interfaces in general, is the notion that the way to address the complex functionality of systems is to base their operation on existing everyday skills. With respect to audio, we have all developed an extensive set of skills that help us navigate through complex spaces and heighten our awareness of our environment. This we have accomplished through a lifetime of living in the everyday world[1]. In the next section of the book, we develop this thesis in the context of a discussion of the theory of *everyday listening*.

This is followed by our third set of case studies, those that are built upon the theory of everyday listening. These examples illustrate the use of natural sounds and their power when applied to complex applications.

One of the problems in dealing with nonspeech audio is that the practical part of audio has not generally played a large part in our background, and our physical environments and computers are not well adapted to exploiting the use of sound as advocated in the book. Consequently, the next section is a type of road map to help the neophyte navigate through the complexities of putting these ideas into practice.

Finally, we conclude by summarizing what we have discussed and speculating about the future. The book has two main purposes: to advocate the use of nonspeech audio and to bring attendees up to speed so that they are equipped to begin to pursue work in the field. This final synthesis section, then, is intended to isolate gaps in our knowledge, and to identify "good" problems and areas of activity that will help advance the state-of-the-art.

---

[1] These are notions discussed more generally by Gibson (1979) and Norman (1988).

While all three authors collaborated in preparing this material, Bill Buxton had primary responsibility for Chapters 1 and 7, Buxton and Bill Gaver for Chapter 2, Sara Bly for Chapter 3, Bill Gaver and Sara Bly for 4, and Bill Gaver Chapters 5 and 6.

Acknowledgements

The authors would like to acknowledge the help of Meera Blattner, Alistair Edwards and David Lunney for their help in making this manuscript possible. We would also like to thank the Association for Computing Machinery, the American Statistical Association, Science News and Lawrence Erlbaum Associates for permission to reprint material. Last but not least, the authors would like to acknowledge the support of Xerox PARC, Rank Xerox EuroPARC and the Natural Sciences and Engineering Research Council of Canada for supporting much of the work underlying the research and preparation of this work.

The Authors

Bill Buxton has an active background in both HCI and computer music. He is Chief Scientist of Alias|Wavefront, and its parent company, SGI, and an associate professor of computer science at the University of Toronto.
buxton@aw.sgi.com
http://www.billbuxton.com
http://www.aliaswavefront.com/research

Bill Gaver is a Senior Research Fellow at the Royal College of Art in London. He gained his Ph.D. in experimental psychology for work on everyday listening, which he applied in the form of auditory icons for Apple Computer and Xerox EuroPARC. Over time, he became increasingly interested in broader issues concerning mediated social behaviour, helping to develop EuroPARC's mediaspace (an audio and video communications network) and developing several experimental systems for supporting social activities over distances.
w.gaver@rca.ac.uk
http://www.crd.rca.ac.uk/research/main.html

Sara Bly is a research scientist at Xerox PARC. She is doing research on the use of audio and video media in supporting collaborative design. In her PhD thesis, she did some pioneering work in the use of audio to present statistical data. She is continuing this work in collaboration with Buxton and Gaver.
sara_bly @acm.org