An Aural User Interface for Ubiquitous Computing

George Rome Borden IV
Sharp Laboratories of America
5750 NW Pacific Rim Blvd, Camas, WA 98607
borden@sharplabs.com

Abstract

Ubiquitous computing systems make it possible for people to access computing resources whenever and from wherever they want. To take advantage of this capability, the user must be able to make queries, perform computations, and generally operate the system at times and in situations that may be incompatible with restrictions imposed by current interfaces. In this paper, we present a novel user interface, discuss its qualities, and compare and contrast it with existing methods.

1. Introduction

Personal systems that offer ubiquitous access to networked data and devices are becoming more prevalent. As they begin to offer better services, people will desire to use them in evermore challenging environments. The motivation for the work presented in this paper is a desire to investigate always-on user interfaces that allow the user to perform complicated tasks and access information.

Current user interfaces are severely limited for use in a variety of situations. Visual interfaces cannot be used concurrently with other visual activities such as driving. Speech recognition interfaces may not be used while performing other speech tasks or while in noisy environments. Additionally, these interfaces often require most of the cognitive resources of the user in order to accomplish even simple tasks. There are many everyday activities where some new kind of interface is desired.

Consider, for example, mobile devices such as CD and MP3 players. Even today most such devices carry one album, and it is possible for the user to quickly memorize the location of each song. It is a simple task to move to the start of the album and click a button a number of times corresponding to the location of the song one wishes to hear. Now think about current MP3 players in which there are 1000 or more songs from various albums and artists. How does one click forward to the 500th song? Many such devices offer a visual interface to assist one in selecting songs. However, what if one is jogging or driving a car? At best the interface is rendered useless and at worst dangerous.

What is required is a user interface that allows one to accomplish a wide variety of tasks concurrent with other daily activities. In this paper we propose an interface that meets these needs and describe a demonstration system we have created that is an embodiment of our work.

2. Related Work

Gaver[1] in his paper on the SonicFinder introduced the concept of utilizing everyday sound with specific actions to provide a clear metaphor to which users can attach meaning. The specific benefits of associating sounds with user actions were later verified by Brewster, Wright and Edwards[2] in their paper evaluating the use of Earcons. These papers had a complementary visual component and were mainly directed at improving the ease of use of graphical user interfaces.

In audio only systems much work was done in phone based menu systems. The use of non-speech sounds to provide navigational cues[3], the advantages and disadvantages of various styles and designs[4], and the cognitive load on users have all been investigated[5]. Early work in general user interfaces pointed out the importance of consistency and scalability[6] and should also apply to aural UIs.

3. The proposed aural interface

The Aural UI consists of a hierarchical menu system in which individual items are either selections of a new set of items or commands. The system requires four inputs: Up, down, select (select new list or command) and exit list. The preferred input device would allow one finger on a hand to access the inputs through translation in a straight line. For example, a dial/button combination much like a Microsoft Scroll Mouse or a rocker switch/button combination. For short lists the rocker switch allows easy counting of the number of movements. The exit function could be achieved by a separate button requiring some movement outside of the line or through a double click of the button. Such an interaction prevents accidental exits, which could confuse the user.

A minimal interface for a system with short lists could actually be implemented with only two buttons, one to

Proceedings of the 6th International Symposium on Wearable Computers (ISWC'02)
0-7695-1816-8/02 $17.00 © 2002 IEEE

move down in a list and another for selection of list items and exiting a list (through double click). This emphasis on minimizing the number of buttons is related to our requirement that the user be able to operate the system with haptic feedback. We want the user to feel the button positions without looking at them and fewer buttons are likely to result in a lower error rate for this type of input.

The aural output of the system consists of aural cues and the speech reading of the names of the selection items. The audio is designed to enable an experienced user to achieve high efficiency with the system, while a novice user can operate the system with little or no instruction due to the speech reading of the items. The cues are "next item" cue, "no more items in list" cue, "entered new list" cue, "exited current list" cue and the "item selected" cue. Additionally, the "next item" cue works on a varying frequency to allow the user to know both how many items are in the list and their current location in the list. The sounds have been selected to represent everyday events[1]. Doors opening and closing represent enter and exiting lists, tapping represents scrolling through a list.

Initial demonstrations of the system have shown behaviors that illustrate the scalability of the Aural UI. For example, rapidly scrolling up until the "no more items" cue is played in order to access items that are remembered to be at the top of the list or rapidly counting the "next item" cues to access the third item in a list.

4. Implementation

The proposed user interface engine reads an XML document to generate the menu structure. Menus can then be accessed via references to the menu ID and additional menus can be dynamically added.

Once the menu structure has been set, the application enters an event loop that calls the main event routine of the engine. This routine will return any events that the user has caused through interacting with the interface. The engine handles all movement within the menu hierarchy, therefore the application need merely respond to events that actually perform some operation. Anyone who has ever coded to a Windows window, an old Macintosh OS or the current Palm OS should find this event model familiar.

5. Discussion and Conclusion

We demonstrate a novel user interface simulating a mobile phone with MP3 playback capability. It is an interesting demonstration for users of existing mobile devices, as they are able to easily contrast our system with their daily experience. We are confident that our interface can operate in a broader range of situations than existing interfaces. Jogging, driving and talking are a small subset of the number of activities that may be carried out concurrently with the use of the Aural UI.

6. Future work

While most of the features of the Aural UI are based on previous work with user studies, as a whole it has not undergone such rigorous examination. In a future paper we would like to perform and present the results of usability testing. Additionally, the varying frequency feature of the "next item" cue requires statistical evidence to determine its value. It is a well known phenomena that some percentage of the population is tone deaf, and it is expected that for those people this method will simply not work, and we would like to investigate alternatives. The ergonomic issues of creating a device that can be operated surreptitiously in a variety of social situations and without visual feedback are also delegated to a future paper.

7. Acknowledgements

My appreciation goes to Dr. Ibrahim Sezan for his support of this work without which it could not have been completed. I would also like to thank Vittal Rao Udipi and Atip Anontvechrucks for their work on the development of the demo system.

8. References

[1] W. W. Gaver, "The SonicFinder: An interface that uses auditory icons," *Human Computer Interaction*, vol. 4, pp. 67-94, 1989.

[2] S. A. Brewster, P. C. Wright, and A. D. N. Edwards, "An evaluation of earcons for use in auditory human-computer interfaces," presented at Conference on human factors in computing systems, Amsterdam, The Netherlands, 1993.

[3] S. A. Brewster, "Using non-speech sounds to provide navigation Cues," *ACM Transactions on Computer-Human Interaction*, vol. 5, pp. 224-259, 1998.

[4] P. Resnick and R. A. Virizi, "Relief from the audio interface blues: Expanding the spectrum of menu list and form styles," *ACM Transactions on Computer-Human Interaction*, vol. 2, pp. 145-176, 1995.

[5] B. R. Huguenard, J. Lerch, B. W. Junker, R. J. Patz, and R. E. Kass, "Working-memory failure in phone-based interaction," *ACM Transactions on Computer-Human Interaction*, vol. 4, pp. 67-102, 1997.

[6] B. Shneiderman, "Direct manipulation for comprehensible, predictable, and controllable user interfaces," presented at ACM International Workshop on Intelligent User Interfaces, New York, NY, 1997.