Strategies for designing auditory overviews

Louise Valgerður Nickerson

Introduction

Imagine getting an idea of what items are on the desktop of a computer via sound. A short presentation of each item could give you an idea of how busy it is and the sort of content is present.

Imagine planning a metro system journey and using a mobile device to listen to the state of the system. The presentation of problems could give you an indication of how much your proposed route could be affected.

Imagine playing a mobile puzzle game where audio was used to represent what is beyond your current focus. You could listen to trends and patterns giving you clues how to solve the puzzle.

The prior examples are only some possible applications of auditory interfaces, however there is little is known how to effectively develop them. The major presentation strategies are auditory icons and earcons. Auditory icons [1] are short real-world sounds and are intuitive due to their mappings. However not all concepts are easily mapped. Earcons [2] are abstract sounds that can be built into an hierarchy to encode meaning. They can present complex information but they are harder to learn. Examples of how these could be used can be found in the Mercator project [3], an interface for X Windows, and Clique [4], an task-based interface. However, their usage is poorly-definied and much relies on design knowledge rather than established convention.

I hope to gain a better understanding of how to create auditory overviews and the appropriate usage and combination of the various presentation strategies by exploring three uses of auditory overviews. A prevailing theory behind this work is Shneiderman's visual information-seeking mantra [5] as adapted to auditory information (gist, navigate, filter, details on demand) [6]. It highlights the importance of a high level views before exploring a data set. Auditory overviews have been investigated for applications such as program source code [7], census data [6] and numerical tables [8].

Prior work – desktop items

As a first step in auditory overview presentation, I sonified computer desktops, aiming to determine if it can be adequately represented by abstract sounds[9]. In evaluation, the biggest difficulty on the part of the users was not so much understanding the disparate information presented but understanding the future usage of the auditory overview. I concluded that comprehension of the information is possible but the lack of context negatively impacted the evaluation.

Resulting research methodology – The question is no longer can this be done? but rather (Q1) when does each strategy work best and how can they be combined (i.e. when are earcons preferable to auditory icons or to speech)? and (Q2) in an overview, how much information can be presented before a user becomes overloaded (i.e. how many streams of information can be presented before the interface loses coherency and it becomes noise)? I hypothesise that:

1. Earcons are best suited to high-level information; that is, layered information that gives a summary of what can be explored.

2. Auditory icons are best suited to events or simple pieces of information.

3. Speech is best suited to providing details information.

4. Circa 6 streams of information can be presented at a time before the interface loses coherency.

I propose the implementation of two complete auditory interfaces to answers the research questions and test the appropriateness of each presentation strategy.

Current work – metro system disruptions

The Underground is London's metro system and the real time disruption map [10] is part of the real time information that is published about the current state of the public transportation network. This view of the system is attractive as the visualisation already follows the principles of the info-seeking mantra [5] and it also a task with which test users would be familiar as part of their daily commute. To properly put the interface in context, I have designed the interface for use on a mobile phone as it closely matches when a sighted user would choose an auditory interface over a visual one due to lack of screen space.

Dissection of the interface – The disruption map represents lines and stations that are disrupted and provides access to the details of the problem. I have defined the following requirements: (1) sonifications of the lines should be concurrent in order to represent the relationships between the lines, (2) each line should be uniquely identifiable in either frequency space or timbre space, (3) individual stations should not be represented unless they are actively disrupted and (4) types of disruptions and severity may be represented in the sonification of the disrupted line or station.

Evaluation – In order to compare presentation strategies, each object in the interface will have an auditory icon, earcon and speech representation. Varying the combination of presentation strategy would lead to answering Q1 and varying the number of disruptions presented would lead to answering Q2.

Future work – sudoku

The next step is to increase the level of complexity of the interface. Sudoku, the popular number puzzle game, is based on filling a nine by nine grid with digits such that no number appears twice in each row, column or three by three sub-grid. This application requires varying zoom levels, ranging from the entire grid to a single cell, and corresponding overviews. The findings of the work on the disruption map will directly affect the design of an auditory sudoku. Here however, 3D sound can also be used since, as a game, it is a longer task and more reasonable to expect headset usage. A short evaluation along the lines of the disruption map evaluation is planned as well as a longer term study in order to examine learning effects.

Expected contribution to HCI

The three main expected outcomes of this research are a better understanding of auditory presentation strategies, how they can be combined in an interface and of the generation of overviews. The ability to build better auditory interfaces can support eyes-free tasks, as well as aid the visually disabled by providing more efficient and effective interfaces. Mobile and wearable computing also gain from this research in terms of usability, especially in the case of mobile phones and audio players where users already use the devices for sound purposes.

References

[1] William W. Gaver, *Everyday listening and auditory icons*, Ph.D. thesis, University of California, San Diego, 1988.

[2] Meera M. Blattner, Denise A. Sumikawa, and Robert M. Greenberg, *Human Computer Interaction*, vol. 4, chapter Earcons and icons: Their structure and common design principles, pp. 11–44, 1989.

[3] W. Keith Edwards, Elizabeth D. Mynatt, and Tom Rodriguez, "The mercator project: A nonvisual interface to the x window system," 1993, citeseer.ist.psu.edu/edwards93mercator.html.

[4] P. Parente, "Clique: A conversant, task-based audio display for GUI applications," in *Doctoral Consortium ASSETS'05*, 2005.

[5] Ben Shneiderman, "The eyes have it: A task by data type taxonomy for information visualizations," in *Proc. IEEE Conference on Visual Languages*, Boulder, CO, USA, September 1996, pp. 336–343.

[6] Haixia Zhao, Catherine Plaisant, Ben Shneiderman, and Ramani Duraiswami, "Sonification of geo-referenced data for auditory information seeking: Design principle and pilot study," Sydney, Australia, July 6-9 2004, International Conference on Auditory Display.

[7] J. Louise Finlayson and Chris Mellish, "The "audioview" – providing a glance at java source code," in *Proceedings of the 11th International Conference on Auditory Display*, Limerick, Ireland, July 6-9 2005.

[8] Johan Kildal and Stephen A. Brewster, "Providing a size-independent overview of non-visual tables," in *Proceedings of the 12th International Conference on Auditory Display*, Tony Stockman, Louise Valgerður Nickerson, Christopher Frauenberger, Alistair D. N. Edwards, and Derek Brock, Eds., June 2006.

[9] Louise Valgerður Nickerson and Tony Stockman, "Sonically exposing the desktop space to non-visual users: An experiment in overview information presentation," in *Proceedings of the 11th International Conference on Auditory Display*, Limerick, Ireland, July 6-9 2005.

[10] Transport for London, "The real time disruption information map, http://journeyplanner.tfl.gov.uk/-im/rd-t.html," .