

InterSon: Interactive Sonification for Geo-referenced Data Exploration for the Vision Impaired

Abstract. InterSon is an interactive sonification tool that allows vision impaired users to explore complex geo-referenced statistical data for fact finding, problem solving and decision making. Examples include maps of population density, crime rates or housing prices. The integrated use of sounds and speech allows users to hear the overall distribution of values on maps and to explore the map to get more details. Users can use the standard computer keyboard, or take advantage of special devices such as a touchpad when they are available. Synchronized auditory and visual displays allow the use of residual vision and facilitate collaboration with sighted colleagues. The prototype was developed at the University of XXXX and is being evaluated with vision impaired users at the University of XXXX.

1 Introduction

Audio is an important information channel for people with vision impairment. The current support for vision impaired users to access geo-referenced statistical data (e.g., population distribution or election results by regions), as well as other types of numerical data sets, relies on screen readers, such as JAWS or Window Eyes, to speak the data presented as tabular records. While speech can accurately describe information, the presentation tends to be very long. Additionally, the reading of data organized in tables does not allow listeners to perceive geographic patterns that can easily be perceived on maps. We believe that effective data sonification (i.e. non-speech audio) can help vision impaired users to explore data collections typically presented on maps and promote equal working opportunities.

Ramloll et al. [4] found that using non-speech sound significantly improved vision impaired users' comprehension of 2-D numerical tables. Research in [1, 2] showed that users can interpret a quick sonified overview of bivariate scatterplots and 2-D line graphs with one or two data series. Several data sonification toolkits [3, 5] have emerged to allow sighted researchers to load data arrays and experiment with data-to-sound mapping designs. While some allowed basic movements in the data, previous work typically lack the support for task-oriented user interaction with the data. Existing toolkits were designed for sighted sonification researchers and developers but not suitable as a data exploration tool for vision impaired users. If the use of maps by blind users for learning real world geography is heavily studied, hardly anyone is addressing the problem of providing universal access to maps presenting numerical data [8, 9].

InterSon (*Interactive Sonification*) enables users to perform map exploration tasks through a sequence of information seeking actions within and across multiple data

views that are tightly coupled. It combines the strengths of both speech and non-speech and lets users actively adjust the auditory feedback detail level according to their information needs. Using equivalent and synchronized visual and auditory displays, InterSon allows the use of residual vision and potentially improves collaborations with sighted colleagues.

2 Description of the interface

The design of InterSon is guided by our Action by Component taxonomy [6] and builds on earlier designs and experimentations [7]. InterSon provides two data views – a table and a map. The table shows multiple statistical attributes simultaneously. Each row corresponds to a region and columns to attributes. Like in most table viewers rows can be sorted, allowing quick locating of low or high values. The map (Fig 1a) shows the geographical distribution of one statistical attribute and can be used to answer geography oriented questions. In each view, users can perform the following actions: start an automatic sweep to obtain a quick *overview* of the data patterns, *navigate* the data collection to examine portions of interest, request *details* of a data item, *select* interesting data items for later examination, and to seek data items satisfying query criterions by *searching* and *filtering*. Users can switch back and forth between the two views which are tightly coupled so that actions in one view are reflected in the other (e.g. selecting, zooming, filtering).

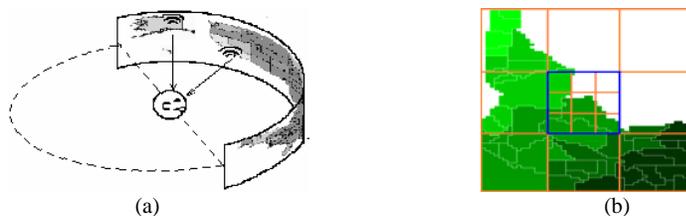


Fig 1. (a) A virtual auditory map that users hear.
(b) The recursive 3x3 grid used to navigate the map using the keyboard

InterSon can be used with a standard computer keyboard (maximizing portability and universal access) or with a standard touchpad, allowing blind users to point to and explore the display in a direct manipulation fashion. With the touchpad, users drag their fingers or press individual spots on the smooth surface touchpad and hear a sound whose pitch indicates the value for the region at the finger position (e.g. high pitch for high values, low pitch for low values). Stereo sounds provide some complementary direction information. The sound feedback stops when the finger lifts off. The touchpad is always calibrated so that the current map range is mapped to the entire surface of the touchpad. While users can easily navigate the table view with arrow keys in a cell-to-cell mode or a column/row mode, using a keyboard to navigate maps with irregularly shaped and sized regions brings special design challenges. For the keyboard interface, we allow a combination of techniques to navigate the map. Users can use arrow keys to move either from one region to its neighboring regions,

or move from one grid cell to its neighboring cells in the mosaic map option. We also take advantage of the good knowledge that blind users have with the numerical keypad. The map is divided into 3 x 3 ranges and users use a 3x3 numerical keypad to activate a spatial sweep of the regions in each of the nine map ranges (Fig 1(b)). For example hitting '1' will play the sounds for all the regions in the lower left of the map, using a consistent sweeping order (left to right then down). Users can zoom into any of the ranges, within which they can recursively explore using the 3x3 pattern or use arrow keys to move around.

InterSon integrates the use of speech and musical sounds. Variations in pitch are mapped to variations of the value of the variable being explored, and different musical instruments are used to indicate when users are outside the map, or crossing a region border in the touchpad interface, or crossing a body of water to reach a neighboring region in the keyboard interface. Speech provides accurate and detailed information, such as region names and exact numerical values. InterSon provides a keyboard shortcut to switch among four information levels: region name only, sound only, name and sound, name and sound plus reading of the numerical value.

Interaction with an auditory interface without seeing is similar to using a command language interface. InterSon strives for a set of simple and intuitive commands that are consistent throughout the system. The application menu system can be navigated with audio, and has links to help messages. It is useful during the initial training and serves as reminder for novice users. For each system component, InterSon has a visual counterpart that is functionally equivalent to the auditory interface. The auditory and visual parts are synchronized, allowing users to choose the perceptual modes or combinations that are most suitable for them. For example, a user with residual vision can obtain general impression of the data via seeing and use audio to acquire more details. A sighted user can mostly use the visual mode but audio can be helpful in revealing items hidden by occlusions, or drawing users' attention to small map regions that are often overlooked. The synchronization, enforced at the data item and command level, allows vision impaired users to collaborate with their sighted colleagues.

A control panel allows advanced users to customize sonification parameters, such as the choices of musical instruments, sound duration, voice speed or sound volume.

3 Implementation

InterSon was designed and implemented at the University of XXXX. The GUI part is written in Java JFC/Swing and the musical sounds are typically produced through the Java MIDI Sound technique. Speech is produced by an accompanying speech server built on Microsoft Speech SDK 5.1. or by loading and playing prerecorded speech or musical sound files. InterSon gives real-time auditory response to every user action, with delays usually less than 100 milliseconds. InterSon can also use virtual spatial sound techniques by sending playing sound commands to spatial sound servers as network datagrams. We have connected InterSon to a spatial sound server developed at (XXXXXX) that simulates real world sounds using Head Related Transfer Functions.

4 Evaluation

The current and earlier versions of InterSon were tested by two blind users using real data sets and task scenarios. Two controlled experiments with sighted users provided insights into the strengths and weaknesses of early designs [7]. During the summer of 2005, in depth case studies with vision impaired users will be conducted, and a between subject experiment will compare the performance of the keyboard-only interface and the touchpad interface with 15 users with congenital blindness, 15 with acquired blindness and 15 sighted but blindfolded users. Our hypotheses are that touchpad user will have higher satisfaction and give more correct answers to questions about shape, size and density of areas; but they also be slower at finding areas with specific values and may miss some areas entirely. We also hypothesize that sighted users will perform better than users with acquired blindness or congenital blindness. Blind subjects have more experience analyzing environmental sounds [10], but it is not clear whether this enhanced awareness extends to other auditory domain, such as listening to abstract sounds, and they have less experienced dealing with spatial information. One of the challenges with this experiment is to evaluate how well blind users perceive the pattern of data on the map, i.e. general trends, location of clusters and exceptions. We will test what subjects have learned about the data pattern with a multiple choice test administered with printed tactile maps. We hope that both interfaces will do equally well at recognizing patterns but observations and log analysis will allow us to compare user strategies, and inform future interface refinements.

References

1. Brown, L., Brewster, S. Drawing by ear: interpreting sonified line graphs. *Proc. ICAD 2003*
2. Flowers, J.H., Buhman, D.C., and Turnage, K.D. Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. *Human Factors*, 39, 3 (1997), 340-350
3. Pauletto, S., Hunt, A. A toolkit for interactive sonification. *Proc. ICAD 2004*
4. Ramloll, R., Yu, W., Riedel, B., and Brewster, S.A. Using non-speech sounds to improve access to 2D tabular numerical information for visually impaired users. *Proc. BCS IHM-HCI 2001*.
5. Walker, B. N. and Cothran, J. T. Sonification Sandbox: a graphical toolkit for auditory graphs. *Proc. ICAD 2003*.
6. XXXXXX. XXXXXXXXXXX, *Proc. ACM SIGCHI Extended Abstracts, 2005*
7. XXXXXXX. XXXXXXXXXXX., to appear in *IEEE Multimedia*
- 8 Landau, S., and Gourgey, K., (2001) "Development of a talking tactile tablet", *Information Technology and Disabilities*, VII(2).
9. Parente, P., and Bishop, G., (2003) "BATS: the blind audio tactile mapping system", Proceedings of ACM South Eastern Conference, Savannah, GA,).
10. Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R.J. & Belin, P. (2004) Pitch discrimination in the early blind. *Nature*, 430, 309