

TECHNICAL RESEARCH REPORT

'I Hear the Pattern' ñ Interactive Sonification of Geographical Data Patterns (2004)

by Haixia Zhao, Catherine Plaisant, Ben Shneiderman

TR 2005-12



ISR develops, applies and teaches advanced methodologies of design and analysis to solve complex, hierarchical, heterogeneous and dynamic problems of engineering technology and systems for industry and government.

ISR is a permanent institute of the University of Maryland, within the Glenn L. Martin Institute of Technology/A. James Clark School of Engineering. It is a National Science Foundation Engineering Research Center.

Web site <http://www.isr.umd.edu>

“I Hear the Pattern” – Interactive Sonification of Geographical Data Patterns

Haixia Zhao, Catherine Plaisant, Ben Shneiderman
Human-Computer Interaction Laboratory
University of Maryland
{haixia, Plaisant, ben}@cs.umd.edu

ABSTRACT

In this paper we describe our investigation of using interactive sonification (non-speech sound) to present the geographical distribution pattern of statistical data to vision impaired users. We first discuss the design space in the dimensions of interaction actions, data representation forms, input devices, navigation structures, and sound feedback encoding. Two interfaces have been designed according to the design space, one using a keyboard and another using a smooth surface touch tablet. A study with three blind users shows that they are able to perceive patterns of 5-category values on both familiar and unknown maps, and learn new map geography, in both interfaces.

Author Keywords

Vision impairment, sonification, auditory user interfaces, information seeking, universal usability

ACM Classification Keywords

H.5.2.a Auditory (non-speech) feedback; H.5.2.e Evaluation/methodology; H.5.2.q. User –centered design

INTRODUCTION

For people with vision impairment, audio is an important alternative or supplementary information channel. The current support for vision impaired users to access geo-referenced statistical data (e.g., the population distribution or election results on a US state map) as well as other types of data relies on screen readers to speak the data presented as tabular records. Such speech based presentation makes it hard for blind users to understand data trends, especially in the geographical context. Automatic textual summarization techniques require predefined summary templates and lack the flexibility to support exploratory data analysis in which users need to examine the data from different aspects. Our ultimate research goal is to enable vision impaired users to accomplish various problem solving and decision making tasks. Geo-referenced data analysis often involves geographical context. This short paper focuses on the challenge of presenting geographical distribution patterns of statistical data using sonification, the use of non-speech audio to convey information [1]. This work will not only help completely blind users, but also benefit low vision users and sighted users in visually overloaded situations. One such example is to draw users’ attention to small regions, e.g., Washington, D.C., that is often the statistics outlier but can

easily be visually overlooked because it is very small on a state map.

A geographical data distribution pattern involves three dimensions, where the data is presented on a 2-D map plane with the data values as the third dimension. Research has shown that users can interpret a quick sonified overview of bivariate scatterplots and 2-D line graphs containing one or two data series [2, 3]. Alty and Rigas [4] found users can recognize simple 2-D graphical shapes presented by musical pitches tracing the outlines. For data with more than two dimensions, Meijer [5] sonified images by time-multiplexed sounds. The effectiveness of the approach remains to be established.

Our work is distinct in three aspects. First, previous work typically focused on sound design but did not consider the effect of letting the user actively explore instead of passive listening. Second, instead of regular grids, geographical data scatters on a 2-D map consisting of regions with irregular shapes and sizes, imposing special design challenges. Third, our investigation expands from using only standard devices (e.g., keyboard) and sound resources (e.g., stereo MIDI) to the exploiting of other common devices (e.g., tablet) and advanced sound techniques (e.g., virtual spatial sound). Previous work involving digital map exploration combined tactile feedback and sound to teach spatial information [6]. While the use of tactile feedback is very likely to be helpful, it severely limits the universal accessibility and portability of the interfaces because of the dependence on special or expensive devices or tactile maps that are not easy to obtain.

Previously, we have investigated several design options by testing with blindfolded sighted users [9, 10]. In this paper, we systematically discuss the design space along multiple dimensions, contrasting various design options, and explain why we have chosen the designs in our two new interfaces, one using a keyboard and the other using a smooth surface touch sensitive tablet. We then report results from a recent study with three blind users on both familiar and unknown maps.

DESIGN SPACE

Interactive data sonification needs to consider both the types of user interaction actions to support and how they should be supported. In order to convey geographical data distribution patterns, the following interaction actions were provided. *Gist* is to play a short audio message that gives a quick grasp

of the overall patterns or trends in the data. It guides further explorations and often allows the detection of anomalies and outliers. *Navigation* refers to the user moving in the data, selecting and listening to portions of the data. *Details-on-demand* is to get detailed information of a data item or group.

Each action involves the user issuing a command and the system giving an auditory feedback. A visual interface provides sustained display that allows the user to directly manipulate any part of the data. In auditory display, information is presented over time. The system needs to help the user explore, construct and maintain a mental representation of the data space. The following design dimensions were considered in our designs: *data representation forms*, *input devices*, *navigation structures*, and *auditory feedback encoding*.

The *data representation form* should reflect the data relations that are most helpful to the task needs. While vision impaired users typically access data in a linear or tabular form, research has shown that they are able to learn, interpret, and benefit from other data representations. For geographical data distribution patterns, a map is a natural form to show the locations and adjacencies of geographical regions. Although such geographical knowledge can be added to a table as columns, the subjects in our previous study [10] strongly preferred map over table and did 12% better after exploring a map than exploring a geographical knowledge enhanced table.

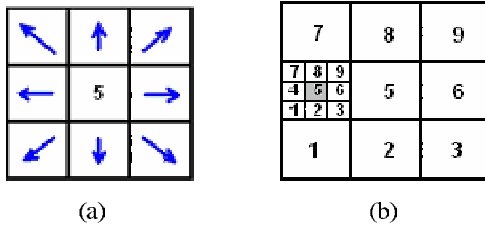


Figure 1: Keyboard as (a) a relative or (b) a low resolution absolute pointing device.

An *input device* transforms the user’s interaction intentions to computer understandable commands. Unlike speech-based input, some physical input devices may provide kinesthetic feedback that may help with user orientation and mental image construction of the data space. A keyboard is a standard input device available on all computers and is widely used by blind users. The arrow keys on a standard keyboard are natural means for relative movements in the left, right, up, and down directions. The numerical keypad on a standard keyboard allows relative movements in 8 directions (Figure 1(a)). Through remapping, a keyboard can be used as a low resolution 2-D absolute pointing device. For example, the 3 x 3 grid of a numerical keypad can be mapped to access nine fixed ranges of the 2-D space. Kamel & Landay used it in a drawing tool [7] and found that users can track grid recursion for 3 levels, which gives a resolution of 27 x 27 grid cells.

A tablet is not a standard input device, but it has become common in recent years. A 14” touch-sensitive tablet can be purchased at a retail price of less than \$150. A touch tablet

provides high resolution 2-D absolute pointing and continuous movements by fingers. The kinesthetic feedback associated with arm and finger movements, combined with the tablet frame as the position reference, may help user orientation during the navigation and mental image construction of the data space. When resources are available, a general grid with some subtle tactile dots can be placed on the tablet as position and direction landmarks to increase the users’ location and direction awareness. Specific tactile maps could also be made to add additional tactile feedback.

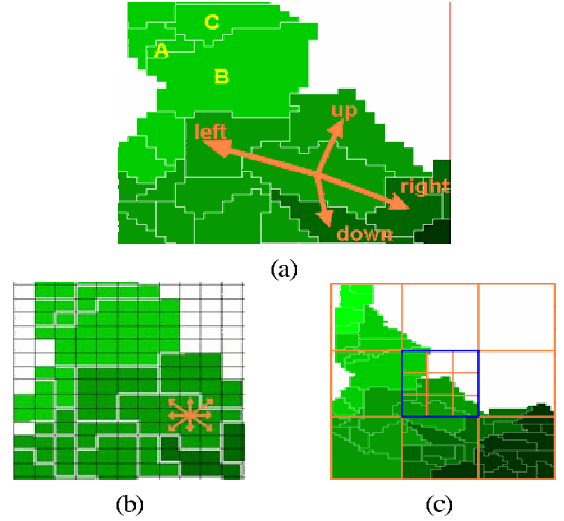


Figure 2: (a) Relative movements by states. (b) Relative movements by mosaic cells. (c) Explore by ranges and recursive zooming

A *Navigation Structure* defines the paths by which the user can navigate the data. On a map presentation, relative movements using a keyboard to go from one region to a neighbor region (Figure 2(a)) tells the user about region adjacency relation. But it does not convey region shapes, sizes, or absolute locations. The blindfolded sighted subjects in our previous studies reported that they only had weak location awareness by using this navigation method. Furthermore, it is a challenge to define a good adjacency navigation path for a map with irregularly shaped regions. A movement may not exactly end up in the direction the user expects. Reversibility of movements can also be a problem. Figure 2(a) shows three regions A, B, and C. Moving up from B usually goes to C. But if the user starts from A, moves down to B, then moves up back, he/she will end up in C although the user may expect to go back to A. To tackle some of the problems, we have also tested cell-by-cell movements on a mosaic version of the map (Figure 2(b)). However, it did not improve users’ pattern recognition or location awareness. Instead, it was much less preferred by the subjects because it required more keystrokes to move around [10].

We expect that navigations based on absolute pointing may help with the problems. By recursively mapping the 3 x 3 numerical keypad on a keyboard to fixed ranges on a map (Figure 2(c)), users could gain knowledge of region absolute positions and region densities in each range. A tablet also provides the ability for continuous movements. Such

continuous movements, when combined with appropriate sound feedback, may allow the users to also perceive region sizes and shapes.

An *Auditory Feedback* conveys information about the data item(s) activated. For the gist of a geographical data pattern, both the statistical value and the 2-D location of each geographical region need to be encoded as sound attributes in the feedback. Using sound duration to present the value significantly prolongs the feedback length, and is not appropriate for a gist in which multiple regions need to be presented. Pitch is a more appropriate choice. Region locations can be mapped to sound locations. MIDI stereo panning can convey azimuth position. Virtual spatial sound synthesized with Head Related Transfer Functions (HRTF) provides relatively high perceptual resolution in the azimuth plane, but is not satisfactory in the elevation plane. None of the subjects in our previous study [9] were able to tell the elevation position of the spatial sound synthesized with a generic HRTF. Elevation perception can be improved by using individualized HRTF, but measuring individualized HRTF is a long process that requires special equipments and careful calibration. Furthermore, HRTF spatial sound is very computing intensive. While we also plan to try individualized HRTF spatial sound, currently we focus on MIDI stereo sound. Of course, elevation perception could be enhanced via extra sound encoding. For example, a pitch of an instrument that is easily distinguishable from the value instrument can be played after each value pitch to indicate the region's elevation position. Unfortunately, our previous study [10] found that such extra sound interfered with the value pattern perception.

While humans are good at selective listening, attending to multiple simultaneous sounds is difficult [8]. Proper sequencing is necessary to present multiple regions. Research has shown that sequencing that preserves spatial relations help users to construct a mental image of the 2-D representation. Regions on a map can thus be sequenced according to their positions. Figure 3(a) shows a vertical spatial sweep on a US state map. Adding a distinguishable sound between sweep columns notifies the user of column switching, provides extra cues for sound location perception, and was considered to be very helpful by subjects in our previous studies. For *details-on-demand* that involves very few regions, speech can be used to provide accurate information.

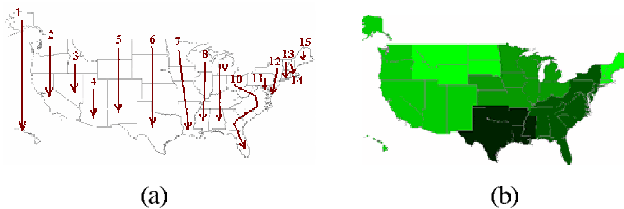


Figure 3: (a) Vertical spatial sweep. (b) A cluster pattern

For continuous exploration with a touch tablet, the sound for the region at the user's finger position can be played continuously for as long as a predefined touch pressure is sensed. A short tick sound can be played when a region

border is crossed. A background sound tells the user there is no region at the touched position.

TWO NEW INTERFACES

Based on the design space discussed, we have recently developed 2 new interfaces, one using a keyboard, and the other using a tablet. In both interfaces, a region value is presented as a string pitch, and the region azimuth location is presented by MIDI stereo panning (0~127). In the keyboard interface, users use a 3 x 3 numerical keypad to activate a vertical spatial sweep of the regions in each of the nine map ranges (Figure 2(c)). Each region value sound was played for 100ms. A 100ms percussion sound indicates the end of a sweep column. A 300ms piano sound indicates the end of the sweep. Users can zoom into a range, and recursively explore that range in the 3 x 3 style. Pressing <0> plays a vertical spatial sweep for the current zoomed map range. Optionally, region names can be spoken along the value pitches. Users press the arrow keys to navigate among adjacent regions in the current zoomed map range. In the tablet interface, users drag their fingers or press individual spots on a 14" smooth surface touch tablet to activate the region at the finger position. The value sound for a region continues until the finger lifts off or moves to another region. A Guitar Fret Noise sound represented background. A 100ms tick sound indicates that the finger has moved across a region border. In both interfaces, users can request spoken name and value of the current region.

The visual displays in both interfaces are synchronized with the auditory displays. During user navigation or automatic sweep, the boundary of the region being played was visually emphasized by 8-pixel-wide lines flashing in a pink color. This has allowed the low vision subject (see below) to track the position of the region being played.

USER EVALUATIONS

The two new interfaces were tested with two completely blind subjects and one low vision subject. All are keyboard users and never used a tablet before. Simulated 5-category data was used to create 5 types of patterns (vertical strip, horizontal strip, diagonal strip, cluster, and no pattern) on maps. An example of monotonic diagonal strip and cluster patterns is shown in Figure 2(c) and Figure 3(b) respectively. In the study, subjects explored the data, described the pattern perceived, reported their current locations on the map when inquired, and answered questions regarding the map geography and their experience with the interfaces.

The study shows that the two completely blind users were able to recognize patterns on both familiar and unknown maps in both interfaces. They were also able to gain geographical knowledge about unknown maps through exploration using only a keyboard or a smooth surface touch tablet. For the low vision user, seeing the overall shape of the map and simple patterns was possible but he only perceived 3 shades of color instead of 5. While his residual vision provided an overview, he thought that "the sound is helpful in augmenting what I can see" and "it is especially helpful for small areas or when the map is small". We will

continue user testing with more blind users but this pilot study has already provided useful insights.

In our previous studies with sighted users, subjects' knowledge about the map has caused significant difference in pattern recognition performance. The blind users in this study did equally well in pattern recognition with both familiar and unknown maps. In fact, they were able to quickly learn new maps in the two interfaces and use the newly acquired knowledge to help constructing their mental images.

Different input devices and navigation methods seemed to have affected map learning. Specifically, navigations based on absolute pointing seemed to be superior to navigations based on relative movements. In previous studies, subjects used a keyboard to navigate in a relative style. On a 7-point scale where 4 is "some sense" and 7 is "good sense", they gave an average score of 4.6 for their location awareness on the map, although they had at least some pre-test knowledge about the map. In this study, the two completely blind subjects both gave a score of at least 6 even for maps they did not know before the test. This, together with their responses to location inquiries during the exploration, indicated they had good location sense. According to the subjects, using the 3 x 3 keypad to explore the map by fixed ranges gave them good sense of general region locations and distributions on the map. By zooming into one of the 9 map ranges and using arrow keys to move around within that range, they were able to explore details about the region layout in that range. For the unknown map in Figure 2(c), one of the completely blind subjects listened to each of the 9 map ranges and immediately reported that the map was probably shaped like an "h". By using a tablet and listening to the tick boundary sound, the subjects were able to describe more, including region relative sizes, and approximate shapes. We also found that the completely blind subjects were able to position themselves on the tablet with relatively high resolution and able to move quite straight in the horizontal, vertical, and diagonal directions.

Regarding the easiness of pattern recognition tasks, tablet seemed to be faster and easier than keyboard. However, one completely blind subject ranked the tablet interface very difficult for unknown maps because that "for a map that I do not know, I have to explore every single pixel on the map using the tablet. Using the keyboard, I can quickly tell which part has things and which part is blank". This same subject thought the tablet interface was less enjoyable because of the same reason. Another completely blind subject, on the other hand, really enjoyed the tablet interface, and thought "It's so cool. I just cannot resist playing with it". All subjects expressed that they would not have problems using the keyboard interface but would prefer the tablet interface

when it is available, because "it can tell more about the map" and "it is nice to be able to touch".

CONCLUSIONS

Our work indicated that blind users were able to perceive geographical data patterns on both familiar and unknown maps from interactive sounds. They were also able to learn new maps using only a keyboard or a smooth surface touch sensitive tablet when appropriate sound feedback was provided. Our work provides an example of systematical investigation of interactive auditory display design guided by a design space framework. The lessons and insights obtained could benefit non-visual interface designs for other user groups.

ACKNOWLEDGMENTS

The work is partially supported by NSF under Grant ITR/AITS 0205271, EIA 0129978 and US Census Bureau.

REFERENCES

1. G. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner, J. Neuhoﬀ, et al, Sonification report: status of the field and research agenda, 1997, <http://www.icad.org/websiteV2.0/References/nsf.html>.
2. J.H. Flowers, D.C. Buhman, and K.D. Turnage, Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples, *Human Factors*, 39, 3 (1997), 340-350.
3. L. Brown, S. A. Brewster, Drawing by ear: interpreting sonified line graphs, *Proc. ICAD 2003*
4. J. L. Alty, D. Rigas, Communicating graphical information to blind users using music: the role of context, *Proc. ACM CHI 1998*
5. P.B.L. Meijer, An experimental system for auditory Image representations, *IEEE Trans. Biomedical Engineering*, 39, 2 (1992), 112-121
6. P. Parente and G. Bishop, "BATS: the blind audio tactile mapping system", in *Proc. ACM Southeast Regional Conf. 2003*
7. H.M. Kamel, J.A. Landay, "The integrated communication 2 draw (IC2D): a drawing program for the visually impaired, in *Proc. ACM SIGCHI 1999*
8. S. Handel. *Listening: An Introduction to the Perception of Auditory Events*. MIT Press, 1989
9. H. Zhao, C. Plaisant, B. Shneiderman, and R. Duraiswami, Sonification of geo-referenced data for auditory information seeking: design principle and pilot study, *Proc. ICAD 2004*
10. H. Zhao, B.K. Smith, K. Norman, C. Plaisant, and B. Shneiderman, Interactive Sonification of Choropleth Maps: Design and Evaluation, to appear in *IEEE Multimedia Special Issue on Interactive Sonification*, Apr-Jun 2000