

Listening to Maps: User Evaluation of Interactive Sonifications of Geo-referenced Data

Haixia Zhao, Benjamin K. Smith, Kent Norman, Catherine Plaisant, Ben Shneiderman

Abstract—In this paper, we summarize the Auditory Information Seeking Principle (AISP) (gist, navigate, filter, and details-on-demand). To improve blind access to geo-referenced statistical data, we developed several interactive sonifications, adhering to the above AISP. Two user studies are presented. In the first user study with nine sighted subjects, a preliminary map design is compared with an enhanced table design. The study shows subjects can recognize geographic data distribution patterns on a real map with 51 geographic regions, in both designs. The map-based design was strongly preferred. The study also shows evidence that AISP conforms to people's information seeking strategies. Based on the observations from the first user study, a second user study was conducted with forty-eight sighted subjects comparing four map designs. The effects of using sound to encode vertical geographic positions and two map navigation methods were compared. The result is presented and future work is discussed.

Index Terms—Auditory (non-speech) feedback, evaluation, interaction style, sound, user interfaces

1 INTRODUCTION

Auditory information is an important information channel for the visually impaired. Sonification is the use of nonspeech audio to convey information [1]. Effective data sonification can be used to promote equal working opportunities for people with vision impairment by helping them explore data collections for problem solving and decision making.

For example, the data collected by the U. S. Census are an important source of information for government and industry alike. The current support for visually impaired users to access geo-referenced data relies on screen readers to linearly speak the geographic region names and data that are presented as table records, often in alphabetical order. Examples include FedStats, the U. S. government statistical data gateway [2], and Corda's [3] accommodation of visually impaired users by automatically converting maps and graphs to descriptive text. Such linear textual presentation makes it difficult for visually impaired users to locate specific data and understand data patterns in geographical context. There are many possible ways to improve visually impaired users' access to such data collections. Ramloll et al. [4] found that using nonspeech sound in 2-D numerical tables decreased subjective workload and enhanced data comprehension.

In our effort to solve this problem, we worked with a

blind design partner and developed several sonifications. Two user studies have been conducted to investigate different sound designs and interaction methods. The sonifications have synchronized visual and auditory presentations and follow two design guidelines: (1) conform to an *Auditory Information Seeking Principle* (AISP) [5] (also summarized in section 2); and (2) use techniques that do not require special equipment, and are therefore easily accessible to the general public.

Geo-referenced data analysis often involves geographical context information. In the visual mode, a picture is often said to be worth a thousand words. A glance at the geographic distribution pattern of data often gives users valuable information. Our goal is to achieve a similar effect in the auditory mode. The geographical distribution pattern of geo-referenced data involves three dimensions, two for the geographical location of a data point on a map, and a third for the numerical value. Previous work has shown that users can interpret a quick sonified overview of 2-D line graphs containing a single data series [6], and possibly two data series [7, 8, 9].

There have been few observations about the ability to recognize data distribution patterns with more than two dimensions in the auditory mode. Meijer's work [10] aims to let visually impaired users "see" with hearing. The effectiveness of this approach remains to be established. Wang and Ben-Arie [11] found that people can recognize simple shapes on binary images of 9 x 13 resolutions where the pixels are raster-scanned slowly. Jeong [12] showed that people could locate the minimum/maximum value on a simplified choropleth map with up to nine regions, with the values presented as different sound volumes.

We conducted two usability studies to investigate different sonification designs. There were three goals for our

-
- H. Zhao & B. Shneiderman are with the Human Computer Interaction Laboratory & Department of Computer Science, University of Maryland, College Park, MD 20742. E-mail: haixia@cs.umd.edu & ben@cs.umd.edu
 - B.K. Smith & K. Norman are with the Department of Psychology, University of Maryland, College Park, MD 20742. E-mail: bsmith@lap.umd.edu & kent_norman@lap.umd.edu
 - C. Plaisant is with the Human Computer Interaction Laboratory, University of Maryland, College Park, MD 20742. E-mail: plaisant@cs.umd.edu.

studies: First, we wanted to validate the applicability of the AISP. Second, we wanted to investigate several sound designs and interaction methods, as possible components of an effective sonification design for geo-referenced data. Finally, we wanted to test people's ability to perceive patterns in sonified data.

In our first user study [5], a preliminary map-based design was compared to an enhanced table design. The study showed that subjects were able to perceive geo-referenced data distribution patterns on a real map with 51 geographic regions using both designs. Observations and user comments indicate that AISP fits users' pattern recognition strategies. Based on observations from the first user study, the second user study compared two map navigation methods and investigated the effect of using sound to encode vertical geographic positions.

In section 2, we describe the Auditory Information Seeking Principle (AISP), and summarize some of the unique design challenges imposed by human auditory perception characteristics. Section 3 describes several geo-referenced data sonification designs and two user studies. Section 4 concludes and discusses future work.

2 AN AUDITORY INFORMATION SEEKING PRINCIPLE (AISP)

One of the guiding principles for the contemporary research on visual information seeking has been "overview first, zoom and filter, then details-on-demand" [13]. If information seeking in the auditory mode follows the same pattern then the collaboration between visual users and auditory users might become easier. In an earlier paper [5] Zhao, et al. proposed an Auditory Information Seeking Principle (AISP):

1. *Gist*: A gist is a short auditory message presenting the overall trend or pattern of a data collection. It guides further explorations and often allows the detection of anomalies and outliers. Because sounds are perceived by humans as transient time-sensitive stimuli, the gist needs to be short and allows active attention/rehearsal to transform it from short-term memory to work memory. Furthermore, human auditory perception is less synoptic than visual perception. Multiple data items need to be presented in serialized ways instead of all at the same time (temporization). Resulting from the short length and low parallelism, data aggregation need to be used when the data collection is large, i.e., more than one hundred data items.
2. *Navigate*: It refers to the user flying through the data collection, selecting and listening to portions of the collection. Navigation is an iterative process of the user initiating an action and the system giving feedback about the user's current range of interest. Because sound is transient feedback, users need to tie sound to virtual objects and construct mental navigation maps in order to interact with the data items through the auditory interfaces. The input methods need to be designed using input devices suitable for visually impaired users. For example, the point-and-click method using a traditional computer mouse works well for normal-sighted users but is difficult for users with vision impairment.

3. *Filter*: Filtering out unwanted data items helps to trim a large data collection to a manipulable size, and allows the user to quickly focus on their interested items.
4. *Details-on-demand*: Users can select an item or group to get details. While sonification emphasizes the use of nonspeech sound, speech can be an effective presentation at the details-on-demand level.

The AISP was used to guide our geo-referenced statistical data sonification design which in turn is a case study to validate the principle.

3 GEO-REFERENCED DATA SONIFICATION AND TWO USER STUDIES

As with any other data sonification, the design space is wide for geo-referenced data sonification. Towards our goal of identifying the most effective designs, we have built several sonifications conforming to the AISP, and conducted two user studies. AISP was used to guide the design that in turn is a case study to verify AISP.

3.1 User Study One: Spatial Map vs. Enhanced Table

The first user study was a within-subjects experiment that investigated whether users could perceive geographic distribution patterns of a five-category data on a 51 region U. S. state map [5]. The purposes were to check the feasibility of using sonification to present data referring to real maps (not simulated or simplified maps), to investigate the validity of the AISP, and to obtain early user observations to guide future designs. Nine subjects were paid to participate in the study. Each subject used both a spatial map design and a table design enhanced with geographic location knowledge.

In the spatial map design, HRTF spatial sounds were tied to a US state map to create the effect of a virtual map surrounding the user at the center (Fig. 1). The spatial sound was based on the KEMAR mannequin HRTF from the CIPIC HRTF database [14] which is widely used as a "generic" HRTF. For each US state, a pitch of a string instrument sound was played for 200 milliseconds to indicate its geo-referenced datum, followed by a 100-millisecond piano pitch indicating its vertical position. The piano pitch was used as a supplement to the low vertical localization accuracy of non-individual HRTF spatial sounds [15]. The five value pitches were from an increasing scale of CEGCE starting from the middle C on a piano keyboard. A higher pitch indicated a higher value. The vertical position pitches ranged from about one octave below middle C to about two octaves above middle C. A higher pitch indicated a state to the North. Using a keyboard, users could start an automatic spatial sweep from West to East (Fig. 2) to listen to a 25-second gist of the data of all the states (*gist*), navigate the map to explore individual states (*navigate*), and request the spoken detail of individual states (*details-on-demand*). During a sweeping, a bell sound indicated the end of a sweep column, and three consecutive bell sounds indicated the end of the sweeping. A bell sound was also played when the navigation automatically jumps to a different sweep column.

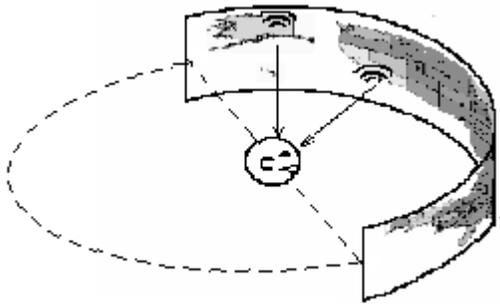


Fig. 1: Spatial sound tied to the map creates the effect of a virtual half-cylinder shaped map surrounding the user located at the center. The illustration does not reflect the real spatial parameters.

In the enhanced table, the states were ordered conforming to their occurrences in the spatial map sweeping. Users could start an automatic sweeping following the table order (*gist*), navigate the states following the sweep order (*navigate*), and request state details (*details-on-demand*). For each state, a pitch of the same string instrument sound was played for 200 milliseconds to indicate its geo-referenced datum. The state name was spoken at the same time. All sounds came from the center.

The results were that subjects were able to perceive the general pattern type after only listening to a 25-second gist once (the overall accuracy was 56% for both the map and the table). After exploring for about 110 seconds, the general pattern type recognition accuracy increased to 78% for the table and 89% for the map. Subjects were also able to grasp the details of the patterns with an accuracy of 67% for the table and 75% for the map. The subjects strongly preferred the map design to the table, although no statistically significant difference was found in terms of performance. The study also showed evidence that AISP conforms to the subjects' information seeking strategies. Our experience with our blind design partner suggests that similar results would be obtained from visually impaired users. For more details about this user study, please see [5].

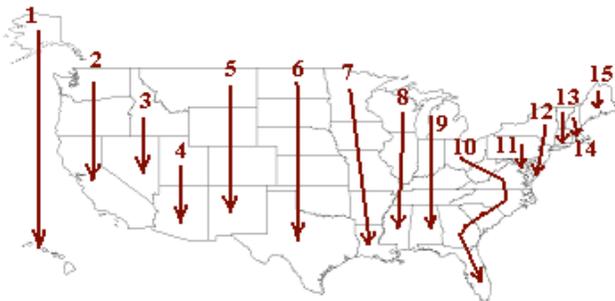


Fig. 2: The sweeping order (in both user studies)

3.2 User Study Two: Comparing Four Map-based Interface Designs

We expected the map-based design could be significantly improved. The design of our second user study is based on two observations from the first user study:

1. Some subjects reported the sound that indicates the state vertical position distracted them from the value.
2. Irregular state shapes and sizes impose difficulty on defining a good state-by-state navigation matrix. It

often causes the actual navigation direction to drift away the subjects' expectation, possibly causing misinterpretation.

3.2.1 Experimental Design

The second user study was a 2 x 2 between-subjects experiment. There were two factors.

The first factor was the presence or absence of a vertical position sound (VPS). There were two treatments. In the treatment without the VPS, for each state, only a 200-millisecond string pitch was played to indicate the state's geo-referenced data value. In the treatment with the VPS, for each state, a 100-millisecond piano pitch followed the value pitch to indicate the vertical position of the state. The ranges of the value pitches and VPS pitches were the same as those in the first user study.

The second factor was the navigation method (NM). There were two navigation methods, namely, the column and mosaic interfaces. The column interface used state-by-state navigation (See Fig. 2(a)) similar to the navigation method in the first user study. The differences were that there was no automatic jumping between adjacent sweep columns when the top or bottom of a column was reached, and the control keys were remapped using only the keys in the number pad plus the four arrow keys on a standard keyboard to control the interface. Table 1 lists all the controls available to the users.

The mosaic navigation method was based on a map created by laying a grid over the regular map. To generate the map for the mosaic navigation, an even-sized grid was laid on the map. All the states were ordered by size, starting with the smallest state. Each state was then assigned one cell closest to the state's center, to ensure that even a small state gets one cell. After that, all the unassigned cells were assigned to the state that had the maximum count of pixels in the cell. The automatic assignment was finally manually adjusted to smooth out the state borders.

Using the number pad, the users navigated the cells in eight directions. When the users moved from a cell to another cell in a different state, the sound(s) of that state were played. When the movements were within one state, no sound was played. The purpose of this was to allow the users to sense the size of a state. When a movement crossed the background (e.g., an ocean, which had no data values) to get to a state, a series of percussion sounds were played before the new state's sound(s) were played. The number of percussion sounds played equaled the count of background cells crossed. All the key controls were on the number pad plus the four arrow keys. Table 1 lists all the controls available to the users.

Whenever a border was reached and the users cannot go further in that direction, the users were reminded by a synthesized female speech that they were already at the boundary. During the sweeping and the navigation, the visual display was synchronized with the auditory presentation. This was to help the subjects understand the interfaces better during the interface explanation. During the training task and experimental tasks, the display was hidden from the users, and could be seen only by the experimenters. In the column navigation, the current state is

highlighted in a blue color. In the mosaic navigation, the current cell is always marked by a yellow dot. The dot also moves across white space when a national border or ocean is crossed.

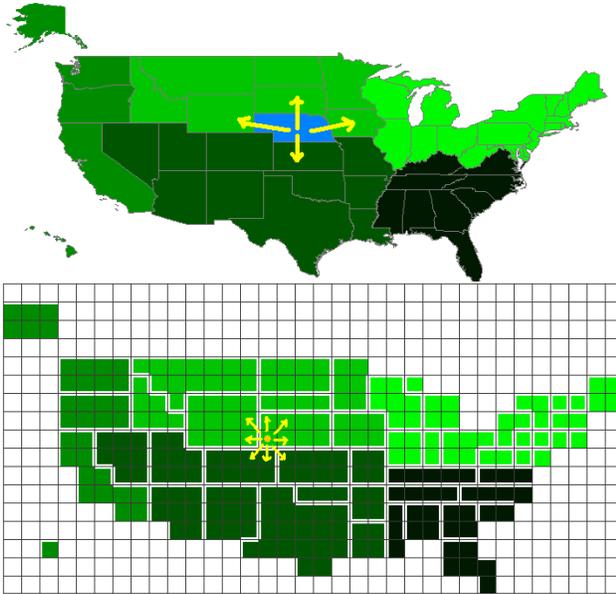


Fig. 3: (a) State-by-state column navigation. (b) cell-by-cell mosaic navigation

All four interfaces use the same sweeping order as in the first user study (See Fig. 2). Since all subjects from the first user study reported they could not tell the vertical positions of the HRTF spatial sounds used, we simply used the stereo panning (0 ~ 127) in the second study to indicate left-right sound positions. The sound to indicate the end of each sweep column and the end of a sweeping was changed from a bell sound in the first user study to a percussion sound in the second study for implementation reasons.

Forty-eight sighted subjects from introductory Psychology courses participated in the study to earn extra credit. Ages ranged from 18 to 51 with a median of 20. There were 37 female and 11 male participants. All subjects reported using computers at least one hour per week, and 44 reported using them at least 5 hours per week. Fourteen subjects reported having had professional music training for a year or more. None of these factors were significantly correlated with performance on any outcome measure at the .05 level. Subjects were randomly assigned to one of four interface conditions, and to one of six task order conditions.

First, each subject was tested on his or her ability to recognize instruments, pitches, stereo panning, and on geographic knowledge. Subjects were then taught how to use the interface they had been assigned. (Two experimenters, neither of whom knew the focal hypothesis, ran the experiment; there were no significant differences in performance for subjects run by each experimenter.) The subjects learned the sound design and interface controls while viewing the display. The display was then hidden from the subjects, and they practiced by performing a training task, with a monotonically horizontal-strip pattern, following the same procedure as in the real test.

Key	Mosaic navigation	Column navigation
8	Go to closest valid cell* above current cell, play sound**	Go to & play previous state in the current sweep column
2	Go to closest valid cell* below current cell, play sound**	Go to & play next state in the current sweep column
4	Go to closest valid cell* left to current cell, play sound**	Go to & play state in the previous sweep column that is nearest to current state
6	Go to closest valid cell* right to current cell, play sound**	Go to & play state in the next sweep column that is nearest to current state
Up, Down arrow	Go to North- (South-) most valid cell*, play sound**	Go to & play North- (South-) most state in current sweep column
Left, Right arrow	Go to West- (East-) most valid cell*, play sound**	Go to & play state in first (last) sweep column that is nearest to current state
1, 3, 7, 9	Go to closest valid cell* diagonally adjacent to current cell, play sound**	
Enter	Play gist of all states	
0	Play sub-gist starting from current state	
5	Request state name and value pitch of current state	
+	Request spoken value of current state	
Any key	Stop playing the gist, set current state to be the state just played in the gist	
* Valid cell: a cell that is part of a state		
** Play Sound: play a percussion sound for each background cell crossed, then play the target state if it is different from current state		

Table 1: User key controls in the state-by-state column navigation and the cell-by-cell mosaic navigation. Only keys from the number pad plus the four arrow keys are used.

Each subject performed 6 pattern matching tasks. There were two tasks with vertical-strip maps, (one monotonic and another interleaving), two tasks with diagonal-strip maps, (one monotonic and another interleaving), and two tasks with cluster maps. Fig. 4 and Fig. 5 illustrate the pattern types and patterns used. The task order was counter-balanced using a Latin square. The orders were set up such that no subject ever had the same general map type twice in a row. Subjects were notified that both accuracy and speed would be measured, but accuracy was more important.

The task procedure was similar to the first user study but with a few changes. Each task was carried out in three steps. First, subjects listened to the gist of the data once, and were asked whether they perceived any pattern in the data by choosing from the five pattern types as shown in Fig. 4. Subjects also chose their confidence level about the answer based on a 10%-break scale. Second, subjects explored the map by using the key controls listed in table 1 for as long as they needed or up to 3 minutes. Subjects then chose the

pattern type again, and their confidence level. Third, four maps were presented to the subjects, each with the same pattern type. One of the four maps was the actual map that the users had been exploring, and thus the general pattern was not necessarily the same as the pattern chosen by the subject. Subjects then chose the matching pattern from the four visual patterns, with their confidence level. Fig. 5 shows a sample of such four visual pattern choices. All keystrokes made by the subjects, as well as the time they took to explore the maps, were recorded. At the end of the six tasks, the subjects were given a post-test questionnaire. The entire experiment took less than an hour per subject.

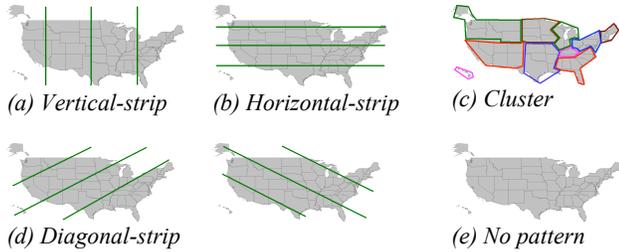


Fig. 4: Pattern types

	Monotonic	Interleaving
Vertical		
Horizontal		
Diagonal		
Cluster		

Fig. 5: Sample patterns for each pattern type

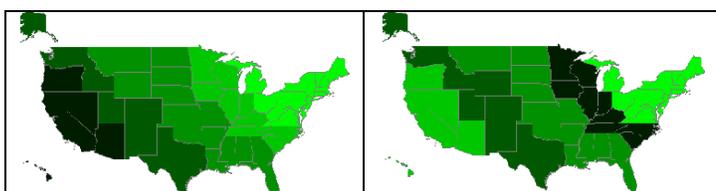


Fig. 6: A sample of four visual pattern choices. A lighter color presents a higher statistical value. Values are categorized into five ranges

3.2.2 Results

Subjects were able to perceive five-category value distribution patterns on a 51-state real map, although overall the tasks were difficult. The average pattern type recognition accuracy was 50.7% after the gist but before exploration (chance accuracy would be 20% because there were 5 choices). After exploration, the pattern type recognition accuracy increased slightly to 55.2%, and the specific pattern recognition accuracy was 48.7% (chance accuracy 25%). There are at least two likely explanations for the low accuracy for the second study compared to that in the first study. First, the tasks in the second study were more difficult. After exploration, subjects had to choose the general pattern type explicitly, rather than having their choice inferred from their selection of a specific pattern. Similarly, subjects had to choose a specific map from four similar choices of the same pattern type, rather than from a set of choices of different pattern types. Second, the differences between the subject populations possibly accounted for some of the difference. In the second study, subjects were not paid and were not given performance incentives, as subjects in the first study were.

No statistically significant difference was found in the performance across the four interfaces. However, we found that subjects' performance was significantly affected by four correlational factors:

1. Geographic knowledge. Subjects' knowledge of United States geography was positively correlated with performance on identifying both the general ($r = .31, p < .05$) and specific ($r = .36, p < .05$) patterns after the exploration period.
2. Pitch differentiation ability. All subjects were able to distinguish between three pitches on our pretest, but seven subjects needed a second try to get the answer correct. These subjects generally did worse on all outcome measures, and significantly worse on the general pattern type after exploration ($r = .45, p < .05$).
3. Task strategy. The post-test questionnaire asked subjects to describe what strategies they used. We compared the strategies people used by identifying common words and phrases in the descriptions. Some strategies appear to have been more effective than others. Subjects who reported listening for changes did particularly well on the two outcome measures taken after the exploration period. (For the general pattern type after exploration, $r = .36, p < .05$; for the specific pattern type, $r = .31, p < .05$) Subjects who reported trying to visualize the map did particularly well identifying the general pattern type after the gist ($r = .49, p < .05$). Subjects who reported paying attention

to the piano sound (which indicated the vertical position of the states) did significantly worse identifying specific maps ($r=.30$, $p<.05$). The most common strategies reported were moving around to find particular states, and visualizing the map.

The tasks varied in difficulty, and each had only a single pattern type, so there is no way to separate the difficulty of identifying patterns or maps from the difficulty of the specific tasks given to the subjects. However, we can report that the easiest tasks were the vertical patterned maps, and the most difficult map was a diagonal one. The only task for which identification of the general pattern type improved significantly after the exploration period was the vertical, monotonic map ($t(47) = 2.72$, $p<.05$). See Table 2 for details.

The post-test questionnaire asked subjects to rate the difficulty of the task, and the difficulty of the interface. Subjects in all conditions found the task hard (2.8 on a 7-point Likert scale, with very difficult as 1 and very easy as 7.) Subjects using the column navigation interface found it easier to use than subjects using the mosaic navigation interface. On the same scale, column navigation users rated the interface 6.0, mosaic users rated it 5.2. This difference was significant ($F(1, 44)=5.23$, $p<.05$).

Two questions on the post-test questionnaire asked about the stereo sound. The first asked whether the stereo sound helped users locate states. The responses were just above neutral (4.5) on a 7-point Likert scale from very distracting (1), to very helpful (7). The second question asked whether the stereo sound helped subjects picture the data distribution. Subjects that did not have the vertical position sound gave higher responses (5.2) to this question than subjects than had the vertical position sound (4.4) ($F(1,44)=5.21$, $p<.05$).

	Horiz (train)	Cluster 1	Cluster 2	Diag Int	Diag Mon	Vert Int	Vert Mon
Gen1	0.44	0.44	0.54	0.50	0.50	0.63	0.44
Gen2	0.42	0.50	0.50	0.44	0.52	0.67	0.69
Spec	0.42	0.52	0.52	0.29	0.46	0.63	0.50

Table 2: Average accuracy for different general pattern types. Gen1 indicates identification of the general pattern type after the gist. Gen2 indicates identification of the general pattern type after exploration. Spec indicates identification of the specific map. Int represents interleaving patterns; Mon represents monotonic patterns. Horiz, Diag and Vert represent horizontal, diagonal and vertical patterns, respectively. Chance performance levels for Gen1 and Gen2 are .20; for Spec chance performance is .25. The horizontal pattern was the training task and was always first. The other three patterns are averages of two tasks and were counterbalanced throughout the rest of the experiment.

Two questions asked whether the vertical position sound was helpful for locating states or picturing the data distribution (only subjects that were in one of these conditions were asked this question). The average responses were very close to neutral, 3.8 to the first and 4.2 to the second, on the same scale as that for the stereo sound.

A question asked subjects how good of a sense they had of where they were on the map. On a 7-point scale, subjects gave a response of 4.6, which was just better than “Some sense,” (4) but considerably lower than “Good Sense” (7).

There were no significant differences between conditions for this question.

Finally, subjects were asked about the tempo of the sounds. The average response was 4.5 on a 7-point scale, close to “The right tempo” (4), on the side of “too fast”.

3.2.3 Discussion

Because there were no significant differences in performance between the four interface conditions, we must be cautious in drawing conclusions. We can say, however, that the vertical position sound seems to have been unhelpful at best, and subjects that reported paying attention to it did worse than those that did not. It also seems to have taken away from utilization of the stereo sound. There does not appear to be any advantage to using a vertical position sound, and combined with the added complexity of such a sound, we do not intend to use it in future sonifications.

The column navigation was liked better than the mosaic navigation, although there were no significant differences between the interfaces in terms of performance. We don’t know exactly what it was about the column navigation that was preferred, but it was somewhat simpler to use, with fewer keystrokes necessary, and feedback given after every keystroke.

Some of the difficulty subjects had identifying patterns and maps may have been due to the experimental conditions, and not the interface itself. Subjects in our experiments had a few minutes to learn the interface, but might have benefited from more learning time. Visually impaired users would likely spend considerable amounts of time learning and using such systems, so more training time is reasonable. It might also be useful to suggest certain strategies to users during training, such as visualizing the map and trying to listen for changes, as these strategies were helpful for other users.

4. CONCLUSION AND FUTURE WORK

The two user studies have revealed certain strengths and weaknesses of our sonification interfaces. The observations will help us improve both user training and the interface design. Our future work includes two main directions:

First, we plan to replicate the studies with visually impaired users to compare with the observations obtained with sighted users. This will require the development of new outcome measures that do not depend on visual displays.

Second, the interface can be improved in many ways. For data-to-sound mapping, we believe the temporization (sweeping order) could be improved. In the second user study, the easiest patterns to recognize were those that had gradient changes that conformed to the sweeping order. We plan to investigate this relation further, and to provide multiple sweeping orders for the users to explore patterns.

We believe the interaction design could be improved by using absolute pointing methods. To avoid the need for special external devices, both of the current map navigation methods are based on a standard keyboard. Every keystroke results in a relative, incremental change to the exploration position on the map. The questionnaire showed that

subjects had a fairly weak sense of where they were on the map during the navigation. Many studies have shown that, in both the real world and virtual environments, motor (e.g., vestibular) information is used together with sensory (e.g., visual) information to construct a mental spatial representation (e.g. [16]). Zheng et al. [17] investigated how users' navigation devices and modes of operation affect their ability to develop an accurate mental spatial representation of a virtual environment. Absolute pointing mode was found to be better than relative mode.

There are various ways to provide absolute pointing methods for map navigation, using either a standard keyboard or a special external device. For example, the keyboard key layout can be mapped to positions on the map. Each key represents a position on the map. Of course, the pointing resolution is limited. A touchpad calibrated to the full map range can be used to provide continuous movement on the map. It can be further enhanced by a tactile grid or map laid on top of it. Using a tactile map requires the users to have access to special devices such as tactile embossers, and restrict the flexibility of switching maps (e.g., zooming into a state on a state map brings up the county details of the state, requiring a different tactile map). Using a generic tactile grid can be a middle-point solution that provides some tactile positioning cue while reducing the requirement of special devices.

ACKNOWLEDGMENT

This material is based upon work supported in part by the National Science Foundation under Grant No. EIA 0129978 (see also <http://ils.unc.edu/govstat/>) and Grant No. ITR/AITS 0205271, the US Census Bureau, and the National Center for Health Statistics. We thank our blind design partner Ahmad Zaghal, and thank Dmitry N. Zotkin for providing the spatial sound technique. We thank Michael Kelehan and Cristina Frascchetti for helping to run the second experiment. We thank Jeffery K. Smith for helping to design the second experiment. We also thank Barbara Shinn-Cunningham and John Flowers for their comments and suggestions.

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, available from <http://www.icad.org/websiteV2.0/References/nsf.html>, last accessed on Jan. 20th, 2004
- [2] FedStats, <http://www.fedstats.gov>, last accessed on Jan. 20th, 2004
- [3] Corda Technologies Inc., <http://www.corda.com>, last accessed on Jan. 20th, 2004
- [4] R. Ramlohl, W. Yu, B. Riedel, and S.A. Brewster, "Using non-speech sounds to improve access to 2D tabular numerical information for visually impaired users." in *Proc. BCS IHM-HCI 2001*, Lille, France, 515-530. 2001
- [5] H. Zhao, C. Plaisant, B. Shneiderman, R. Duraiswami, "Sonification of geo-referenced data for auditory information seeking: design principle and pilot study", to appear in *Proc. of Intl. Conf. Auditory Display*, July 6-9, Sydney, Australia, 2004.
- [6] J.H. Flowers and T. A. Hauer, "Musical versus visual graphs: cross-modal equivalence in perception of time series data." *Human Factors*, 1995, 37(3), 553-569.
- [7] 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*, 1997 v39 n3 p340(11)
- [8] T.L. Bonebright, M. A. Nees, T.T. Connerley, and G. R. McCain, "Testing the effectiveness of sonified graphs for education: a programmatic research project", In *Proc.Int. Conf. Auditory Display*, 2001, Espoo, Finland, pp. 62-66.
- [9] L. Brown, S.A.Brewster, "Drawing by ear: interpreting sonified line graphs," in *Proc. Int. Conf. Auditory Display* 2003.
- [10] P.B.L. Meijer, "An experimental system for auditory Image representations," *IEEE Transactions on Biomedical Engineering*, Vol. 39, No. 2, pp. 112-121, Feb 1992. Also as the vOICe system at <http://www.seeingwithsound.com>, last accessed on Jan. 20th, 2004
- [11] Z. Wang, and J. Ben-Arie, Conveying Visual Information with Spatial Auditory Patterns, *IEEE Transactions on Speech and Auditory Processing*, 4, 446-455, 1996
- [12] W. Jeong, *Adding haptic and auditory display to visual geographic information*, PhD thesis, Florida State Univ. 2001.
- [13] B. Shneiderman, "The eyes have it: a task by data type taxonomy for information visualization." In *Proceedings of the IEEE Symposium on Visual Languages*, Sept 3-6, pp. 336-343, 1996
- [14] V. R. Algazi, R. O. Duda, D. P. Thompson, and C. Avendano. "The CIPIC HRTF database," In *Proc. IEEEWASPAA01*, New Paltz, NY, pp. 99-102, 2001.
- [15] E.M. Wenzel, M. Arruda, D.J. Kistler, F.L. Wightman, "Localization using nonindividualized head-related transfer functions." *Journal of the Acoustical Society of America*, Vol. 94, no. 1, pp.111-123, 1993.
- [16] D.J. Simons and R.F. Wang, "Perceiving real-world viewpoint changes," *Psychological Science*, 9(4), 315-320.
- [17] X.S. Zheng, G.W. McConkie, and B. Schaeffer, "Navigational Control Effect on Representing Virtual Environments," *Human Factors and Ergonomics Society Meeting*, October 2003

Haixia Zhao is a doctoral student at the Department of Computer Science, and Human-Computer Interaction Laboratory (HCIL, <http://www.cs.umd.edu/hcil>), both at the University of Maryland. She obtained a Master of Science degree in 2002 from the University of Maryland, a Master of Science degree in 1999 from Academia Sinica, and a Bachelor of Science degree in 1996 from Zhejiang University (P.R.China). She has worked on database, data warehousing, and information management system before joining the Human-Computer Interaction Laboratory. Her current research interest is on human-computer interaction.

Benjamin K. Smith is a graduate student in the Neuroscience and Cognitive Science (NACS) program at the University of Maryland. He works for the Laboratory for Automation Psychology and Decision Processes (LAPDP, <http://www.lap.umd.edu>). He graduated from Brown University in 2000 with a Bachelor of Science degree. His research interests include human-computer interaction and cognitive psychology.

Kent Norman is a Professor of Cognitive Psychology at the University of Maryland. He is the director of the Laboratory for Automation Psychology and Decision Processes (LAPDP, <http://www.lap.umd.edu>) and is a founding member of the Human/Computer Interaction Laboratory (HCIL, <http://www.cs.umd.edu/hcil>). His research is on human/computer interaction and cognitive issues in interface design. Current research involves the design of on-line surveys and visualizations for decision making and is funded by the U.S. Bureau of the Census. He is an associate editor for the *International Journal of Human/Computer Studies*.

Catherine Plaisant is Associate Research Scientist at the Human-Computer Interaction Laboratory of the University of Maryland Institute for Advanced Computer Studies. She earned a Doctorat d'Ingénieur degree in France in 1982 and has been conducting research in the field of human-computer interaction since then. In 1987, she joined Professor Shneiderman at the University of Maryland, where she has worked in collaboration with students and members of the lab, throughout the growth of the field of human-computer interaction. Her research contributions range from focused interaction techniques to innovative visualizations validated with user studies to practical appli-

cations developed with industrial partners.

Ben Shneiderman is a Professor in the Department of Computer Science, Founding Director (1983-2000) of the Human-Computer Interaction Laboratory (<http://www.cs.umd.edu/hcil>), and Member of the Institute for Advanced Computer Studies and the Institute for Systems Research, all at the University of Maryland at College Park. He is a Fellow of the ACM and AAAS and received the ACM CHI (Computer Human Interaction) Lifetime Achievement Award. His books, research papers, and frequent lectures have made him an international leader in this emerging discipline.