SONIFICATION OF GEO-REFERENDED DATA FOR AUDITORY INFORMATION SEEKING: DESIGN PRINCIPLE AND PILOT STUDY

Haixia Zhao[1], Catherine Plaisant, Ben Shneiderman[1]   *Ramani Duraiswami*

Department of Computer Science[1]
& Human Computer Interaction Laboratory,
University of Maryland,
College Park, MD 20742, USA
{haixia,plaisant,ben}@cs.umd.edu

Perceptual Interfaces and Reality
Laboratory, UMIACS,
University of Maryland,
College Park, MD 20742, USA
{ramani}@umiacs.umd.edu

ABSTRACT

We present an Auditory Information Seeking Principle (AISP) (gist, navigate, filter, and details-on-demand) modeled after the visual information seeking mantra [1]. We propose that data sonification designs should conform to this principle. We also present some design challenges imposed by human auditory perception characteristics. To improve blind access to geo-referenced statistical data, we developed two preliminary sonifications adhering to the above AISP, an enhanced table and a spatial choropleth map. Our pilot study shows people can recognize geographic data distribution patterns on a real map with 51 geographic regions, in both designs. The study also shows evidence that AISP conforms to people's information seeking strategies. Future work is discussed, including the improvement of the choropleth map design.

1. INTRODUCTION

For people with vision-impairment, auditory information is an important alternative or supplementary information channel. Sonification is the use of nonspeech audio to convey information [2]. Effective data sonification can help vision-impaired users to explore data collections for problem solving and decision making. As a result, it promotes equal working opportunities for people with vision-impairment.

One of the guiding principles for the contemporary research on visual information seeking has been "overview first, zoom and filter, then details-on-demand" [1]. If information seeking in the auditory mode follows the same pattern then the collaboration between visual users and auditory users might become easier. In this paper we propose an *Auditory Information Seeking Principle (AISP)*:

Gist: quick grasp of the overall data trends and patterns from a short auditory message.

Navigate: fly through the data collection and closely examine portions of interest.

Filter: seek data items satisfying certain criteria.

Details-on-demand: obtain details of groups or an individual items for comparison.

However, many new design challenges arise because human auditory perception is very different from visual perception. Furthermore, appropriate interaction methods must be designed to fit the characteristics of the input devices that are suitable for vision-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.

The AISP was used to guide our geo-referenced statistical data sonification design which in turn is a case study to validate the principle. The current support for vision-impaired users to access geo-referenced data relies on screen readers to linearly speak the geographic region names and the data that are presented as table records, often in an alphabetic order. Examples include FedStats, the USA government statistical data gateway [3], and Corda's [4] accommodation of vision-impaired users by automatically converting maps/graphs to descriptive text. Such linear textual presentation makes it hard for blind users to locate a specific data item and understand data patterns, especially in their geographical context. There are many possible ways to improve vision-impaired users' access to such data collections. Ramloll et al. [5] found that using nonspeech sound in 2-D numerical tables significantly improved vision-impaired users' data comprehension and decreased the subjective workload.

In our effort to solve this problem, we worked with a blind design partner and developed two preliminary sonifications - an enhanced table and a spatial choropleth map. These two sonifications follow two design guidelines: (1) conform to the AISP; and (2) use techniques that do not require special equipment, and are therefore easily accessible for the general public. For example, to avoid the need for special tactile devices, haptic perception was not used as a supplementary information channel. We investigated the use of spatial sound in the sonification, but did not deploy spatial sound techniques based on speaker arrays or assume the availability of individual head related transfer function (HRTF). Instead, the spatial sound was based on the KEMAR mannequin HRTF from the CIPIC HRTF database [6] which is widely used as a "generic" HRTF.

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 quick glance at the geographic distribution pattern of the data often gives users very valuable information. Can we achieve a similar effect in the auditory mode? The geographic distribution pattern of geo-referenced data involves three dimensions, where the data points scatter on a 2-D map plane with the values of the data points as the third dimension. Research has shown that users can interpret a quick sonified overview of 2-D line graphs containing a single data series [7], and possibly two data series [8, 9, 10].

However, 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 [11] aims to let vision-impaired users "see" with hearing. It translates an arbitrary image into a time-multiplexed sound representation

that is the superposition of the sound of multiple image pixels. The effectiveness of this approach remains to be established. Wang and Ben-Arie [12] found that people can recognize simple shapes on binary images of 9 x 13 resolutions where the pixels are raster-scanned slowly. Jeong [13] shows that people can locate the minimum/maximum value on a simplified choropleth map with up to nine geographic regions, with the values presented as different sound volumes.

Our pilot study shows that people are able to perceive geo-referenced data distribution patterns on a real map with 51 geographic regions using both our sonification designs. Observations and user comments indicate that AISP fits users' pattern recognition strategies. Working in parallel, Perez et al. [14] suggest that the auditory tasks are Situate, Navigate, Query, Details on demand. They make arguments based on analogy with visual interfaces, and plan to implement voice interfaces for Web browsing.

In section 2, we propose our Auditory Information Seeking Principle (AISP), raise some design challenges imposed by human auditory perception characteristics. Section 3 describes two preliminary geo-referenced data sonifications, an enhanced table and a spatial choropleth map. A pilot study is also reported regarding geographic data distribution pattern recognition, and the validity of AISP. Section 4 concludes and discusses future design improvements that are especially applicable to geo-referenced data exploration.

2. AN AUDITORY INFORMATION SEEKING PRINCIPLE (AISP)

We propose data sonification designs should adhere to this Auditory Information Seeking Principle (AISP): gist, navigate, filter, and details-on-demand.

2.1. 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.

Processing a gist is a short-term memory task. Human's short-term memory (STM) has a limited capacity and decays within 10 seconds unless actively maintained by attention/rehearsal [15, 16]. A gist should be short, maybe 5-30 seconds. A gist that does not fit into the STM should include pauses to allow midpoint STM processing.

The design of a gist often involves the following questions.

(1) *What is the data-to-sound mapping?* A data item could be enumerative, numerical, etc. An instrumental sound has the attributes of timbre, pitch, duration, etc. The mapping from data items and item relations to sound attributes must ensure the easy association of the data item to the sound. Many guidelines have been derived from experiments and practices [17, 18, 19].

(2) *How to serialize the auditory presentations of multiple data items?* Serializing multiple data items into a gist is essentially part of the data-to-sound mapping, where designers choose the item relations to map to the auditory time dimension. Because auditory perception is far less synoptic than visual

perception, the degree to which multiple sounds can be superimposed is limited. How many sounds can be presented in parallel depends on the types of sounds used and how much information needs to be extracted from each sound.

(3) *What is the data granularity to be presented in the auditory message?* While a small collection of tens of data items can be serialized in a few seconds' composition, data aggregations must be used when the collection is large, i.e., more than one hundred data items.

2.2. Navigate

Navigation refers to the user flying through the data collection, selecting and listening to portions of the collection. Navigation is not just to play, stop, resume or rewind along the time dimension of the data collection's soundscape. Navigation is an interleaved process of the user initiating an action and the system giving feedback.

In visual interfaces based on graphical displays, navigation is typically done by using 2-D pointing devices to directly manipulate the visual objects in the virtual environment, and possibly change the visual appearance of the virtual environment as the result, that is displayed through the 2-D computer screen. It is enabled by the continuous, sustained feedback empowered by today's graphics display techniques and by a human's highly synoptic visual perceptual ability.

However, this typical process in visual interfaces cannot be directly applied to the auditory mode, because sounds are perceived by humans as transient time-sensitive stimuli. As a result, a mental navigation map must be constructed for navigating the data collection through the auditory interface. A basic practice is to navigate backward/forward along the time dimension of the collection's soundscape. Advanced navigation maps can be constructed by binding the auditory presentation of the data collection to virtual objects, such as a table or a geographic map. Using the virtual objects, users can efficiently locate and request auditory stimuli of a data item or a group of items. Navigation maps vary depending on the virtual objects chosen, the data and the tasks. For example, when a table is used, navigation can follow the table order. When a map is used, navigation can follow the adjacency and relative 2-D locations of the geographic regions on the map.

During navigation, a gist of the data portion of current interest is often necessary as the system feedback. The data granularity under examination may change during navigation. The finest navigation level is usually individual data items.

Navigation involves input schemes. While relative pointing devices such as a computer mouse are not suitable for visually impaired users, keyboards and some absolute pointing devices such as a touchpad calibrated to the full 2-D input space with tactile assistances (e.g., a tactile map laid on the top of a touchpad with special pressure sensing) may work.

2.3. 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. In visualization, dynamic query

via sliders, buttons, or other control widgets coupled to rapid (less than 100 miliseconds) display update is the goal [20]. In sonification, different goals need to be established because such a short time period is usually not enough to present a gist of changes.

2.4. 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.

3. GEO-REFERENCED DATA SONIFICATION AND A PILOT STUDY

As for any other data sonification, the design space is wide for geo-referenced data sonification. As a step towards identifying the most effective designs, we built two preliminary sonifications by working with a blind design partner, and conducted a pilot study. The pilot study serves multiple purposes: 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.

3.1. Two sonification designs for the pilot study

In the pilot study, we observed users' ability to perceive geographic data distribution patterns on a real map (a USA state map) in the auditory mode. The study was conducted as a controlled experiment between a table-based design and a map-based design. We expect that map-based designs are better than table-based designs for tasks involving geographic knowledge. The data in the study is simulated, and categorized into five value ranges.

3.1.1. Enhanced table

In the enhanced table, all sounds are stereo without panning. The data is mapped to sound as follows: Five value categories are mapped to five string pitches from an increasing scale of CEGCE starting from the middle C on a piano keyboard. A lower pitch indicates a lower value. The value pitch of each USA state plays for 200 milliseconds.

Gist: The gist is a serialization of all the USA states that lasts for about 30 seconds, with three continuous bells indicating the end. There are 50 USA states plus the District of Columbia. In the paper, we refer to them as 51 states. For each state, the state name is spoken while the string value pitch is played. The serialization goes from the first state record to the last in the table.

The order of states in the table is a sweeping from west to east and north to south according to the states' geographic locations on the real map, and is exactly the same as the sweeping order in the spatial choropleth map design (please see next section for details). While a traditional table is often alphabetically ordered by the state names, we expect it would be

hard even for users with excellent knowledge of the map to catch the data distribution pattern because users have to mentally switch between locating a state from the state name and processing the state's value. Meanwhile, we expect that a table-based design can be enhanced by adding extra geographic information about a state in the form of a data column, such as west region, middle region, and east region. The enhanced table design in our pilot study uses the same sweeping order as in the map-based design to maximize geographic location cues and minimize users' mental load. We believe such enhancement makes the comparison to map-based designs more meaningful.

Navigate: Users can use a keyboard to navigate the soundscape in the table order, such as moving up/down the table to play one state on one key stroke, jumping to the first/last state, and so on. The playback tempo depends on how fast the users press the keys. Table 1 lists all the key functions.

	Enhanced table	Spatial choropleth map
<Up>	Go to & play previous state in table	Go to & play previous state in the sweeping order
<Down>	Go to & play next state in table	Go to & play next state in the sweeping order
<Pg Up>	Go to & play first state in table	Go to & play north-most state in current sweeping column
<Pg Dn>	Go to & play last state in table	Go to & play south-most state in current sweeping column
<Left>	N/A	Go to & play nearest state to the left of current state
<Right>	N/A	Go to & play nearest state to the right of current state
<Home>	N/A	Go to & play state in first sweeping column that is nearest to current state
<End>	N/A	Go to & play state in last sweeping column that is nearest to current state
<Enter>	Play gist of all states	
<R>	Play sub-gist starting from current state	
<Space>	Request detail of current state	
<L>	Play audio legend	
<Any key>	Stop playing the gist, set current state to be the state just played in the gist	

Table 1: User key controls in the enhanced table and the spatial choropleth map.

Details-on-demand: Anytime users press the space key, the name and value of the current state will be spoken.

3.1.2. Spatial choropleth map

In the spatial choropleth map design, all sounds, including the speech but excluding the bell sounds, are spatial sounds synthesized using KEMAR mannequin HRTF. Spatial sounds
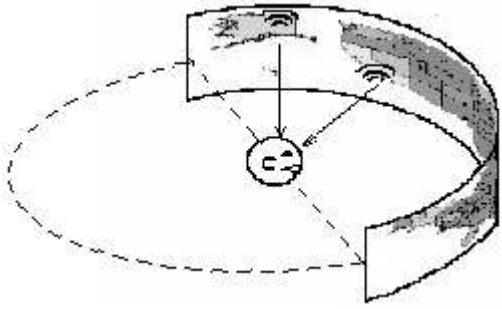
are tied to the map to create the effect of a virtual half-cylinder shaped map surrounding the user located at the center (Azimuth range: -90º ~ + 90º, elevation range: -31º ~ +63º) (Figure 1). Since non-individual HRTF has shown to result in poor elevation perception [21], we use piano pitches to indicate the elevations of states. The piano pitches range from about one octave below middle C to about two octaves above middle C. Lower pitches indicate states to the south. Five value categories are mapped to five string pitches, same to the enhanced table.

Figure 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.

Gist: The gist is a serialized sweeping of all the 51 USA states. While various sweeping orders can be used, the current sweeping follows the fifteen sweeping columns shown in Figure 2, from north to south in each column and from column one to column fifteen. For each state, a value pitch is played for 200 milliseconds followed by an elevation pitch lasting for 100 milliseconds. A bell rings for 570 milliseconds at the end of each column. The gist lasts about 25 seconds, including a 100-millisecond pause between two adjacent columns and three continuous bells indicating the end.

Navigate: Users can use a keyboard to explore the soundscape simply following the sweeping order. For example, pressing the <Up>/<Down> key will go to and play the previous/next state along the sweeping order. Within one column, this is the same as going to the state to the north/south of the current state. When two columns are involved, a bell sound is played before jumping to the most southern state of the previous column or the most northern state of the next column. Furthermore, users can do 2-D map navigations. For example, pressing the <Left>/<Right> key will go to and play the nearest state to the left/right of the current state. Table 1 lists all the key functions.

The fact that states have irregular shapes and sizes on real maps introduces extra difficulties in both defining a good sweeping order for the gist and defining an effective 2-D map navigation grid. In the current one-state-per-key-stroke type of navigation using only four directions of up, down, left and right, it is often not easy to choose the target state when the current state is adjacent to multiple states in the direction of movement. For example, on a <Left> key stroke from the top state in column seven, which one of the top two states in column six

should become the current state? In the design for the pilot study, we based the map navigation on the sweeping columns. For example, a <Left> key stroke will take the user to the state in the previous column whose geographic center is nearest to the current state. An associated problem with this is that sometimes a series of continuous movements in one direction, such as to the left, may actually take the user in a direction shifted northern or southern from the straight left. Alternative navigation design options are discussed in section 4.

Details-on-demand: Anytime the users press the space key, the name and value of the current state will be spoken.
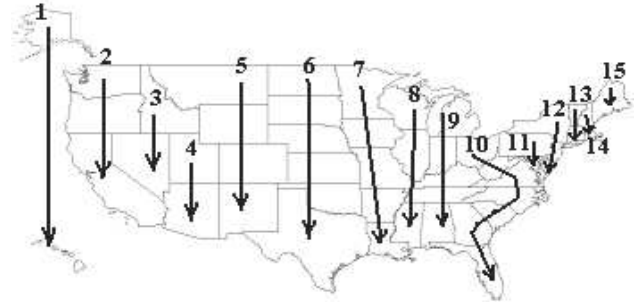

Figure 2: The sweeping order

3.2. Pilot study procedure

Nine sighted University of Maryland students/staff (2 females and 7 males, ages from 19 to 40) were paid to participate in the study. Eight are intensive computer users. Two had some music training in high school. A pre-test showed six subjects had excellent knowledge of USA state locations, while the others did not know the accurate location of some states, especially states in the middle region. The pre-test also showed all subjects can perceive the azimuth location of the spatial sound used in the choropleth map, but were often confused about the elevation location. Subjects could easily tell the timbres used and distinguish the five value pitches well.

The assignment of subjects to the two design conditions was counter-balanced. Before the test of each design, subjects learnt the sound design and interface controls, and practiced one training task following the same procedure as in the real test. Subjects performed four tasks in each design. The study lasted for a little over one hour on average.
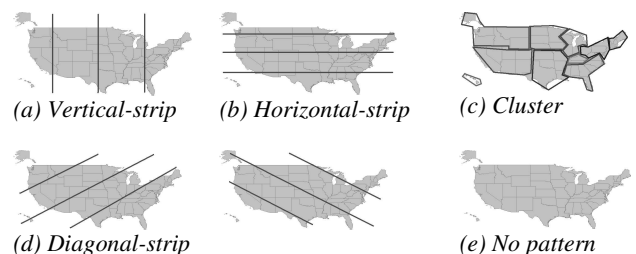

(a) Vertical-strip   *(b) Horizontal-strip*   *(c) Cluster*

(d) Diagonal-strip   *(e) No pattern*
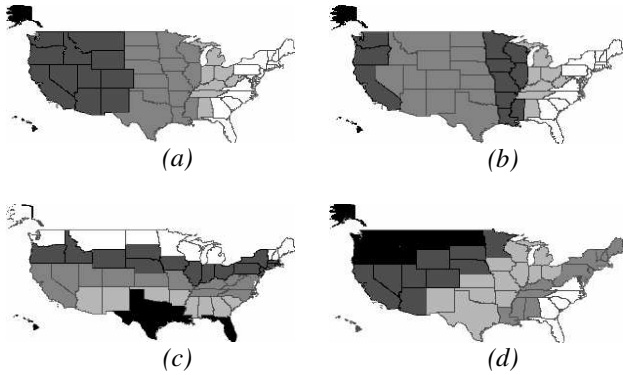
Figure 3: Pattern types

Figure 4: A sample of four visual pattern choices. A darker color presents a higher statistical value. Values are categorized into five ranges.

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 Figure 3. Second, subjects explored the sound scene by using the key controls listed in table 1 for as long as they needed or up to 3 minutes. Subjects then chose the matching pattern from four visual pattern choices, and chose their confidence level about the answer based on a 10%-break scale. Figure 4 shows a sample of such four visual pattern choices. (b) is the pattern played. (a) is considered to be similar to (b). (d) is less similar to (b) and (c) is the least similar. Third, the subjects were told the correct answer and had the chance of exploring the sound scene again for up to 1 minute. Subjects were timed for their exploration and time to choose the answer. The post-test questionnaire asked subjects about the overall experience with the two designs, their pattern recognition strategies, and so on.

3.3. Pilot study result, observation and discussion

In the visual mode, such pattern recognition tasks would take only seconds. The pilot study result shows that it is much harder in auditory mode, but novice users are still able to do it with overall good accuracy in reasonable time, as shown in Figure 5 and Figure 6. After listening to the gist once, the subjects recognized the pattern type with an overall accuracy of 56% in both designs (a random selection will yield an accuracy of 20%). After exploration, the pattern type recognition accuracy increased to 78% for the table, and 89% for the map. Even with the small number of subjects, the increase is statistically significant in map (one-tail, t=-4.0, P=0.002), and almost significant in table (one-tail, t=-1.7, P=0.060). Subjects also obtained pattern details and were able to choose the correct pattern out of a few similar choices with good accuracy (67% in table and 75% in map). Subjects spent an average of 110.7 seconds exploring the table and 110.6 seconds the map. The average task time for recognizing pattern details is 130.0 seconds for the table and 134.2 seconds for the map. After exploration, subjects recognized patterns better in the map than in the table, but no statistically significant difference was obtained.
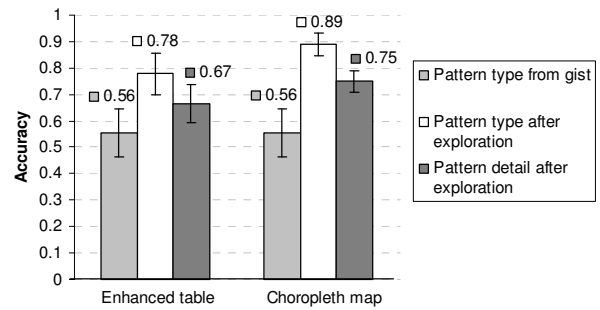

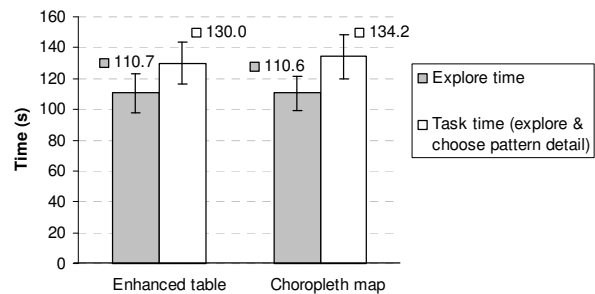*Figure 5: Overall accuracy of pattern recognition (with 9 subjects)*

Figure 6: Overall speed of pattern recognition (with 9 subjects)

Subjects used different task strategies. A common strategy they used is "the gist gives a general idea of the pattern type, I explore to confirm my guess, find where the value changes happen, try to 'visualize' the pattern, and remember the pitches of a few states/clusters/areas to eliminate the wrong patterns". The result shows subjects are able to perceive value distribution patterns on real 2-D maps using both designs. They are able to obtain the "big picture" from a short gist, which guides their exploration, which in turn confirms their initial impression and provides more detail for choosing the correct pattern. This shows evidence for the validity of the AISP.

Most subjects reported that the left-right position cue from the spatial sound in map helped them to orientate and "picture" the data distribution pattern. Most subjects used west-east navigation and commented that it gave them the flexibility of moving in the directions they want in order to confirm the pattern, thus helping them to remember the pattern. These advantages in the map seem to have contributed to the map's higher accuracy in comparison with the table. The lack of a statistically significant difference may be due to the small numbers of subjects, and the current inefficient map design.
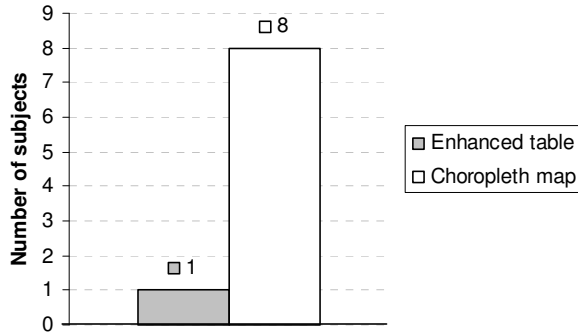
Figure 7: User preference (with 9 subjects)

Most subjects (8 out of 9, Figure 7) strongly preferred the choropleth map and commented that they could probably do better and faster in the map with more practice. Analysis of the error cases and subjects' feedback show the following reasons why the current map design does not perform as good as we expected.

(1) Subjects need time to better understand the sweeping column definition and adjacency in 2-D map navigation. The irregular state shape and size in the real map impose special difficulty in placing states in the sweeping grid, and serializing the states. The subjects were quickly shown the sweeping columns before the test and had the chances to learn more during the learning stage at the end of each task. However, due to the short experiment time, there were not enough tasks/time to ensure the enforcement of the knowledge. Sometimes the sweeping order and navigation order do not meet the subjects' intuitive expectation. For example, subjects expect the beginning of a column starts at the northern boundary of the map. The fact that the sweeping column four starts from the middle could have caused the values of those two states to be misplaced in the subjects' imaginary map. With more practice, such special short columns could be interpreted appropriately, and even be used as position anchors.

The importance of familiarizing subjects with column definitions was also confirmed by our experience with our blind design partner. Before using a tactile map to learn the sweeping columns, he could only tell west-east trends. After learning, he could describe the patterns in high details, in both the west-east and north-south directions.

(2) The elevation piano sound interferes with the value string sound. The piano sound is provided to aid north-south localization of a state, as a supplement to the fact that the 3-D sound technique does not provide robust elevation position cue. This is also expected to solve the problem in (1). But in fact, all subjects reported that two short pitches with one (piano pitch) following immediately the other (string pitch) are too overwhelming for them to process both, especially when multiple states were played continuously. This is consistent with McGookin & Brewster's result [22] about users' identification ability of concurrently presented earcons. As a result, all subjects chose to focus on the value sound and ignore the position sound when they listen to the gist, occasionally shifting quickly to a piano sound to confirm the current location. Some subjects seemed to be able to filter out the second sound, but some subjects reported that the second sound is so distracting that it caused their misinterpretation of the value pattern. All subjects reported that it is much easier to locate the current state through the between-column bell sound, the navigation direction using the arrow keys, and the general west-east, north-south sweeping order they already know.

Comparatively, the verbal presentation of state names in the table is not rated as being distracting as the piano sound in the map. Some subjects reported that it is difficult to switch between state names and value sound frequently, but they also commented that there was no need to pay attention to every state name. Instead, they followed the west-east, north-south order in which the states are played, searching for some specific states as their position anchors. The anchor states usually are the states the subjects are familiar with, and/or the north-boundary/south-boundary states that indicate the sweeping column change. With the accurate position cue provided by the anchor states, the subjects then placed other states in between following the west-east, north-east sweeping order.

The role of state names is clearly demonstrated by subjects' keystroke patterns during the exploration. They often pressed arrow keys fast to skip the playing of state names, and slowed down for state names once they encountered a change of the value pitches or they expected a change in columns. This is also demonstrated in the map where the requests for state names mostly happened upon change of value pitches and change of columns.

4. CONCLUSION AND FUTURE WORK

We have proposed an Auditory Information Seeking Principle (AISP) that could be used to guide the design of data sonification. Our pilot study with geo-referenced data sonification shows users are able to perceive five-category value distribution patterns referring to a real map with fifty-one geographic regions in the auditory mode. Evidence has been obtained from the study that AISP fits users' pattern recognition strategies, indicating the validity of AISP. The two preliminary sonifications - the enhanced table and the spatial choropleth map, are both effective in conveying geo-referenced data patterns.

4.1. Improve the map-based design

Our pilot study did not show the choropleth map was statistically significantly better than the enhanced table. However, the current choropleth map design is inefficient and can be significantly improved. Using better sound technology, we expect that further advances in spatial sound techniques and easier ways of measuring individual HRTF will reduce the elevation confusion and make the virtual spatial map more realistic and accessible. Use of head tracking and motion should also help. Even with the current sound technique and generic

HRTF, a lot can be done by carefully designing the sonification and auditory interfaces. Here are some of the design options we want to investigate next, based on the map metaphor.

4.1.1. Serialization

Serialization is important in designing a gist to convey an overview of the whole data collection, a sub-collection, or changes resulting from user interactions. The data relation to be implied by a serialization should be appropriate to the current task needs. For tasks within geographic context, geographical locations/relations of data items often need to be presented. Our current choropleth map design only provides vertical sweeping line moving from west to east. However, other serializations can be used such as horizontal sweeping, diagonal sweeping, or spiral sweeping starting from the map center or a selected point. Furthermore, states can be grouped into clusters according to the choropleth data values and geographical locations. States within the same value-location cluster can use any serialization above. Different clusters can be ordered by their choropleth data values. Training is often necessary for users to understand serializations on real maps.

4.1.2. Navigation

Navigation not only allows users to fly through the data collection but also gives important cues of geographical locations/relations. Our ultimate goal of navigational interaction is not to rely on any special purpose external devices, such as tactile or haptic displays. The IC2D blind drawing program [23] successfully used a navigation and accurate point selection scheme based on the layout of a telephone keypad. Its grid recursion scheme recursively divides a 2-D cell space to a 3 x 3 grid and uses the nine telephone keypad numbers (Figure 8(a)) to recursively point to the centers of the cells. Based on this scheme, we can design various map navigation methods using only standard keyboards. For example, users can enter a sequence of numbers to recursively zoom into a portion of the map to examine that part of the data (Figure 8(b)). Flexible zooming is important especially when dealing with a large data collection, i.e., a USA county map with 3141 counties. Users can move from one state to another following the eight-way adjacency (Figure 8(c)). For example, from the current state marked <5>, entering number <1> will go to the state to the northwest, and entering number <6> will go to the state to the east. Of course, special care is still needed when the current state is adjacent to one same state in multiple directions (e.g., the states to the south or southwest of state <5> are one single state marked as <7 or 8>), or is adjacent to multiple states in one direction. But it gives more flexibility in defining adjacency than the left-right-up-down four-way movements. Figure 8(d) illustrates a grid of equal sized cells laid on top of a real map. Moving to a cell triggers the state that dominates that cell. The resulting effect is a Mosaic version of the map as shown in Figure 8(e). Users can use it not only to navigate the map but also acquire knowledge of approximate sizes, shapes and layouts of the states. The finer the grid is, the more accurate the Mosaic map is. Besides grid-based navigations, users could

define scanning bars with controllable orientation and length, and use the scanning bar to sweep the map (Figure 8(f)).
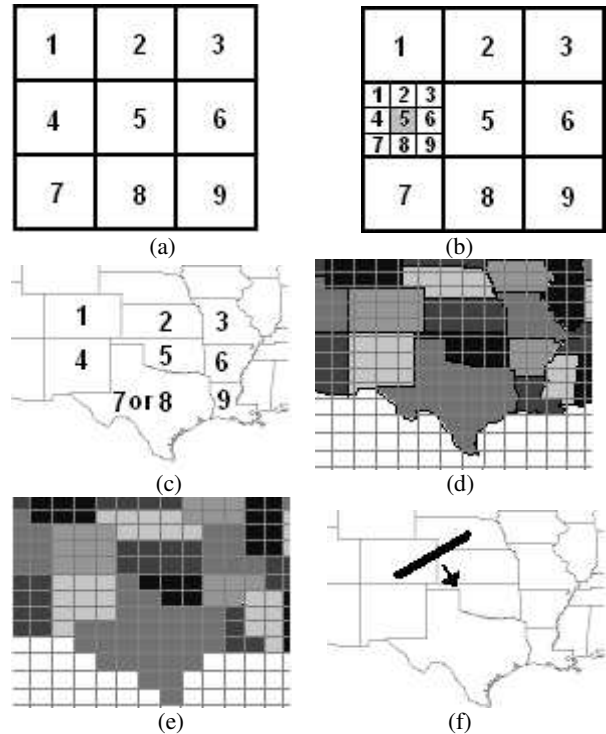

(a)

(b)

(c)

(d)

(e)

(f)

Figure 8: (a) A 3 x 3 grid of cells mapped to the telephone keypad. (b)Recursively zoom into a portion of the 2-D space (e.g., gray area) by entering a sequence of keypad numbers (e.g., sequence 1, 5). (c) Eight-way state adjacency on a real map based on the keypad layout. (d) (e) A Mosaic map can be created for an arbitrary map by laying a grid of equal sized cells. The Mosaic map can be used as the navigation grid, and implies the knowledge of state shapes, sizes and layouts. (f) Use adjustable scanning bars to sweep the map.

When special devices are available, alternative navigations can be used. For example, vision-impaired users can do low accuracy pointing using a touchpad calibrated to the full range of the map. A tactile map laid on top of the touchpad helps navigating and locating specific states.

4.2. Auditory interface for information seeking

One of our research goals is to design an auditory interface for auditory information seeking in geo-referenced statistical data. The pilot study has shown sighted users can perceive geographic distribution pattern in auditory mode. Our experience with the blind design partner shows similar results could be obtained from vision-impaired users. We want to investigate if such designs can be used by vision-impaired users to accomplish other analysis tasks besides pattern recognition, not only on simple maps with tens of states, but even more complicated

maps with thousands of geographical regions. Such an investigation will involve many rounds of design improvements and user studies to examine user behaviors and the cognitive processes.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] 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

[2] 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

[3] FedStats, http://www.fedstats.gov, last accessed on Jan. 20th, 2004

[4] Corda Technologies Inc., http://www.corda.com, last accessed on Jan. 20th, 2004

[5] R. Ramloll, 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

[6] 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.

[7] 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.

[8] 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)

[9] 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.

[10] L. Brown, S.A.Brewster, "Drawing by ear: interpreting sonified line graphs," in *Proc. Int. Conf. Auditory Display* 2003.

[11] 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

[12] Z. Wang, and J. Ben-Arie, Conveying Visual Information with Spatial Auditory Patterns, *IEEE Transactions on Speech and Auditory Processing*, 4, 446-455, 1996

[13] W. Jeong, *Adding haptic and auditory display to visual geographic information*, PhD thesis, Florida State Univ. 2001.

[14] M.A. Perez-Quinones, R.G. Capra, Z. Shao, "The ears have it: a task by information structure taxonomy for voice access to Web pages", in *Proc. IFIP Interact* 2003.

[15] L.R. Peterson, and M.J. Peterson, "Short-term retention of individiual verbal items." in *Journal of Experimental Psychology*, 58, 193-198. 1959

[16] R.C. Atkinson, and R.M. Shiffrin, "The control of short-term memory." In *Scientific American*, 82-90. August 1971

[17] G. Kramer, "Some organizing principles for representing data with sound," in G. Kramer (Ed.), *Auditory Display, SFI Proc. Vol. XV111*, Addison-Wesley. 1994

[18] S.A. Brewster, P.C. Wright, and A.D.N. Edwards, "Experimentally derived guidelines for the creation of Earcons," in *Proceedings of HCI95*, Huddersfield, UK, 1995, pp. 155-159.

[19] B. Walker, and G. Kramer, "Mappings and metaphors in auditory displays: an experimental assessment," in *Proc. Int. Conf. Auditory Display*. 1996

[20] B. Shneiderman, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 3rd Edition, Addison Wesley Longman Inc., 1998

[21] 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.

[22] D.K. McGookin, and S.A. Brewster, "An investigation into the identification of concurrently presented earcons." In *Proc. Int. Conf. Auditory Display*, 2003

[23] H.M. Kamel, J.A. Landay, "The integrated communication 2 draw (IC2D): a drawing program for the visually impaired, in *ACM SIGCHI* 1999.