Using an Audio Interface to Assist Users Who are Visually Impaired with Steering Tasks

Robert F. Cohen, Valerie Haven, Jessica A. Lanzoni, Arthur Meacham, Joelle Skaff, Michael Wissell University of Massachusetts Boston 100 Morrissey Boulevard Boston, MA 02125 {rfc, vhaven, jlanzoni, ameacham, jskaff, mwissell}@cs.umb.edu

ABSTRACT

In this paper we describe the latest results in our on-going study of techniques to present relational graphs to users with visual impairments. Our work tests the effectiveness of the PLUMB software package, which uses audio feedback and the pen-based Tablet PC interface to relay graphs and diagrams to users with visual impairments. Our study included human trials with ten participants without usable vision, in which we evaluated the users' ability to perform steering tasks under varying conditions.

Categories and Subject Descriptors

H.5.2 [User Interfaces]: Auditory (non-speech) feedback

General Terms

Human Factors

Keywords

Accessibility, Audio, Graph, Sonification.

1. INTRODUCTION

1.1 PLUMB

This study is part of our on-going work developing novel techniques to display graphs and relational information to users who are blind. To this end, we have built a software system called exPLoring graphs at UMB (PLUMB) [5][6] that displays a drawn graph on a Tablet PC and uses auditory cues to help a blind user navigate the graph.

Our work found its original motivation in our experience that talented students with visual impairments in computer science are often hampered in their studies, and even forced out of the field by the highly visual approach used to present so many key concepts. Much material, such as data structures, is typically taught by first presenting an example through a diagram and then moving to an algorithm and code. Students with visual impairments are at a severe disadvantage because they have

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ASSETS'06, October 22-25, 2006, Portland, Oregon, USA.

Copyright 2006 ACM 1-59593-290-9/06/0010...\$5.00.

limited opportunities to make use of these motivating examples to develop an intuitive grasp of the concept [5].

Of course, graphs have a wide range of uses beyond the teaching of computer science. Graphs contain vertices and edges, each can be attributed with some text or data. Graphs take many forms, and find use in an immense variety of applications, making their understanding critical to success in both technical pursuits and in day-to-day life.

Outside of computer science education, our work could be used to satisfy other needs of both blind and sighted persons. There are many educational uses unrelated to CS, including the use of auditory events to assist students with visual impairments in experiencing complex geometric shapes. And outside of education, our system could be used to help persons with visual impairments read street maps, navigate buildings and learn fire escape routes. Through our continued research and development, we anticipate finding many presently undiscovered applications for individuals who are either sighted or blind. (We are currently working on collaborative note-taking / research tools.)

Our end goal is to design a set of tools to assist users with visual impairments in understanding relational data. We are currently developing a system that meets the following criteria:

- The system should allow active exploration of drawn graphs by users with visual impairments.
- The interface should be usable by both users who are blind or sighted.
- Where possible, the system should be implemented on widely available hardware and software.

Ultimately, we hope that our research will enable persons with visual impairments to become the producers of such documents as well as consumers. Instead of being limited to only reading graphs, maps, or other visual data, our user interface has the potential to enable a user with a visual impairment to also produce graphs and other such relational data as well as view the work in progress.

As previously mentioned, we have implemented PLUMB for the Tablet PC, a hardware format which has emerged as a popular consumer product. We and others feel that the tablet computer represents an outstanding opportunity to provide users who are visually impaired with a friendly interface for exploring graphical information [6]. The unit's compact size, widespread availability and multimedia capability could provide an effective means for

people who are visually impaired to experience otherwise inaccessible applications with little or no hardware modification.

The most common visual representation of a graph has edges drawn as circles and edges as lines connecting their endpoints. With PLUMB, we aim to convey a drawn graph to a user who is visually impaired or blind. The fundamental behavior of the application is to generate an appropriate audio cue when the user moves the stylus over an element of the graph. Figure 1 uses HSB graphics to illustrate the sound produced by PLUMB to represent a two-edge graph. To indicate progress along an edge, we have experimented with variations of pitch, the use of musical scales, and the insertion of audio effects. We currently use a continuous musical tone to indicate contact with an edge, and modify the tone with a vibrato effect to indicate proximity to either end of the edge, increasing the intensity of the vibrato as the user approaches a vertex. The changes in intensity of the vibrato are indicated by changes of hue in the HSB color model, and variations in loudness are depicted with changes in color saturation.



Figure 1 MIDI guidance illustrated by HSB color

To help users follow the edges, variations in the loudness of the musical tone are used to represent distance from the central axis of an edge, and are here depicted by variation in saturation.



Figure 2 A screenshot of PLUMB. The displayed graph corresponds to the Europe map shown in Figure 3.

A sample of the PLUMB visual display is shown in Figure 2, with the visible graph components corresponding to the locations of the underlying sound components. A representation of the same graph with associated text fields is shown in Figure 3.

While using the program, the user "draws" on the tablet screen with the stylus to receive audio feedback and explore the graph. When the stylus enters or exits a vertex, a sound is played to communicate the event. Additionally, upon entering the vertex, speech is used to announce the name of the element, and possibly give a brief description. To obtain more detailed information about a particular element, the user triggers a right mouse-click event, at which point PLUMB will give a more-detailed spoken description of the element, as well as a list of other elements to which the selected element is connected. If the mouse is clicked while the pen is over an open area, general information about the graph is provided.



Figure 3: The European city graph displayed by PLUMB in Figure 2.

To build the application, we opted to write the software using C# and Microsoft's .NET framework. We made this choice primarily because this development platform has extensive support and SDKs for Tablet PC, which are not available for other tools such as Java. To generate sound, we used DirectX, the Microsoft Speech API, and the MIDI for .NET v2.0.4 library.

Since the graphs we present to the user must eventually be stored as files, we chose to represent our data as GXL documents [8]. GXL is an acronym for Graph eXchange Language, an XMLbased language designed specifically for representing graph structures. The GXL documents hold descriptions of the vertices and edges, as well as supporting information required to re-create a graph. This descriptive information is represented as tags which are parsed by the application, and used to build and lay-out the graph, as well as to populate it with the data necessary to generate appropriate sound and speech.

2. PRIOR WORK

Other work has been done in sonification of relational data (see, for example, [3]). We categorize previous systems as using either passive or active exploration. With passive exploration, the user is presented with a representation of the entire diagram at one time, with limited user input (e.g. AudioGraf [10] and the case tool of Bleckhorn and Evans [2]). Such tools are beneficial for selective algebraic expressions and mathematical data but offer no assistance with other potential applications for PLUMB. With active exploration systems, the user can explore the diagram (e.g.

TeDub [12] for exploring UML diagrams and Kekulé [4] for exploring molecular structures). Several interfaces utilize keystrokes and commands as a means for active exploration (e.g. Mynatt [12]). van den Doel et al [14] developed a system to convey simple geometric shapes using both auditory and haptic feedback. Much like PLUMB, users explored the shapes on a tablet using the accompanying stylus pen; however, the participants were sighted but blindfolded. After comparing the users' ability to either replicate the shapes on a separate sheet of paper or select the previously explored shape from a group of eighteen shapes, the results indicated that users were relatively proficient in both.

Fitts' Law [8] is a vastly studied and applied model of human gestures that we have found to be quite relevant to our analysis. To model the act of pointing, Fitts made predictions based on the amount of time it takes to reach an ending point from a beginning point. Instead of the trend of completion times being completely linear, Fitts determined that the trend in times of completing such actions actually displayed exponential growth.

Related to Fitts' work, and also quite useful to our analysis, is the Accot-Zhai Steering Law [1][11] which extends the model to two dimensions. Accot and Zhai modeled how users accomplish the task of moving a pointing device along one axis, while staying within a certain distance along the other access (like steering through a tunnel). The measure of difficulty they propose is simply the tunnel's length divided by the tunnel's width. We measured a similar task using PLUMB: the time and accuracy for participants to trace an edge from start to finish. Our goal is to determine how users who are blind solve steering tasks using an audio interface.

3. METHODS

3.1 Early Trials

This work was influenced by earlier trials performed using PLUMB. At the time, we conducted tests with six different visually-impaired computer users, two of whom are IT professionals, and one of whom is a computer science undergraduate. While our sample size was small, our aim was to get a general feel for the effectiveness of the system.

We presented the subjects with graphs representing common concepts such as airline flight maps or organizational charts, and then asked the subjects to perform simple tasks and answer questions about the contents of the graphs. We hoped to discover whether the users would find the program easy to use and understand, and whether the tablet computer was a viable format for this kind of application. In addition to observing as users interacted with the system, we also generated log-files which recorded data for all pen movements and graph events from each session. These logs have allowed us to study the behavior of the users, as well as to gather statistics about their interaction with the program.

These early trials were broken into four distinct stages, which presented the participants with graphs of increasing complexity. The first stage presented a graph consisting of a single vertex and the final stage consisted of the relatively complex graph shown in Figure 2 and 3.

The participants generally expressed enjoyment at using the system, and all were able to learn the interface. However, we quickly noticed that users experienced a number of common

problems while using PLUMB. This caused us to take a step back and reformulate our testing approach, and may actually cause us to reevaluate the applications for which PLUMB might be effective.

The first, and possibly the most significant observation was that random access to vertices is problematic. When a participant was presented with a graph and asked to find a specific node within that structure, the user had no context with which to seek the node, and found the process to be frustrating and tedious. This was especially true of graphs representing abstract relationships, such as the social network in Figure 4, in which there was no underlying geographic context to guide the user. Knowing where one vertex was located did not help the participant find another vertex, and the user was reduced to exploring the graph at random until encountering the target. This is not true, however, of geographic graphs in which the positions of the nodes have semantic significance, as in the train map from Figures 2 and 3. For that example, assuming the user is familiar with European geography, once that user has found the node representing London, he or she should have a pretty good idea of where to look to find Kiev

Another observation was that some users had difficulty handling the stylus. This seemed to be more of a problem for congenitally blind users than for adventitiously blind users, but we did not have enough participants in our initial trials to make that determination.

Finally, we noticed a great deal of variation in the difficulty users experienced while traversing different edges. We observed that the same participant might navigate one edge quite easily, and then struggle extensively with another. The differences appeared to be related to the variations in length, width, and angle of the different edges on the graph, but again we could not be sure without further study.

Among our questions were: Are edges of certain angles (with respect to the screen of the Tablet Personal Computer) more easily navigated than others? Is there an optimal edge width? How are edge length and tracing success correlated? With these questions in mind, we decided to take a step back and evaluate the basic components of a graph rendered with PLUMB using a variety of angles, sizes, and positions.



Figure 4: A graph with no underlying geographic context.

3.2 Objectives

The core objective of this phase of our research has been to learn how PLUMB graphs should be designed so as to be as effective as possible in communicating relational information to visually impaired users. Data-capturing software was used to record the performance of participants as they navigated simple (two-node, one-edge) graphs. We have compared performance across graphs that differ in each of the following dimensions:

- Edge width,
- Edge length,
- Edge direction,
- Edge angle,
- Screen region of presentation.

Our goal is to use the information we gather to tailor PLUMB software to create the most effective (i.e. most navigable) graph configurations possible.

3.3 Significance

We believe that PLUMB can be an important real-world tool for conveying relational information to the person with visual impairments, in particular to students with visual impairments of science and technology. In order for PLUMB graphs to be as useful as possible, they should be tailored to fit users' navigational strengths. This research is significant because it will reveal these strengths and will enable us to capitalize on them in future applications.

3.4 Participants

We conducted these trials with the assistance of ten participants with no usable vision, three of whom were congenitally blind and seven of whom were adventitiously blind. All ten were adult professionals or university students who use computers on a routine basis.

3.5 Procedure

At the beginning of each experiment, a researcher briefly described to the participant the purpose of the project and provided him or her with instructions for using PLUMB. Then, the participant was given a training exercise in which he or she explored a sample sound graph as the experimenter explained each of the sounds and informational features of the graph.

Once the preliminary training was completed, the participant was presented with a sequence of twenty-five simple graphs, each consisting only of two nodes and one edge. Several features of the graph were randomly varied from test to test. The first randomly selected feature was the width of the edge, which could be set to 20 pixels, 40 pixels, or 60 pixels across. Next was the angle of the edge, which could be vertical, horizontal, diagonally ascending, or diagonally descending. Third was the area of the screen in which the graph appeared: We divided the screen into four quadrants, and placed the graph randomly into any one of these quadrants, drawing the graph from one side of the quadrant to the other. As another option, we also sometimes used the entire screen to display a larger graph with a longer edge. Finally, we set the direction of the edge by randomly assigning one node to be a "start" vertex and one to be a "finish" vertex. To prevent repetition of the same configuration within a trial, we created a mechanism to test for previously-used graph layouts and prevent their use.

In each test within the sequence, the participant was presented with a simple sound graph as described above. The participant's hand was placed at the "start" node by the researcher¹. He or she was then asked to trace the edge to "finish" node. As the participant traversed the graph, data-capturing software recorded all significant events that took place (e.g., entry into/exit from a given node, cursor movements) and the times and locations at which they occurred. Each individual task would terminate at the moment when the participant entered the "finish" node.

After the completion of the trial, a researcher conducted a brief interview to gather qualitative information about the participant's experience using PLUMB.

4. RESULTS

Data was analyzed across several different variables. First, we recorded the total time in seconds, denoted by t, required by the participant to traverse an edge from the start node to the end node. Second, we recorded the total number of errors produced by the participant during each traversal, denoted by e. An error was defined as the stylus exiting the region of the edge before reaching the destination node. To aid in our analysis, we have calculated a ratio consisting of the total number of errors divided by the total time of traversal for each test, a value we will denote as *B*. Additionally, we have computed the averages for the longer, full-screen graphs (e=19.93, t=39.38, B=.51), and compared these to the results of other trial types, both sorted by width and without considering width. To our surprise, we also found that edge angle and edge direction were not statistically significant factors in performance between subjects.

	Mean Time (t)	Mean Error (e)	Error /Time (B)
All	23.81	10.37	0.44
Small Width (20 Pixels)	31.22	18.72	0.59
Medium Width (40 Pixels)	21.85	7.53	0.34
Large Width (60 Pixels)	17.74	4.20	0.23

Table 1: Summary of results by edge width

Finally we investigated performance differences between participants who were congenitally blind (CB) and those who are adventitiously blind (AB). We discovered a significant difference between the overall time and total error count of those subjects born blind and those who became blind after some level of visual

¹ Placing the user's hand at the start node is not required for PLUMB. We do this to get a precise start time for measuring the task of following an edge.

experience. Interestingly, however, there was no significant difference between the error/time ratio B displayed by the two groups (see Table 2).

	Mean Time (t)	Mean Error (e)	Error /Time (B)
Congenital	31.95	13.14	0.41
Adventitiou s	20.93	9.96	0.47

Table 2: Performance of Congenitally Blind participants compared with that of Adventitiously Blind participants

DISCUSSION

These trials were conducted primarily to identify those aspects of PLUMB that are not intuitive to users. Participants were given a short while to understand the use of the devices, and then tested before they could acclimate to the system. (Three participants had used the system briefly in earlier trials.) This quick introduction allowed us to see how new users performed without any familiarity with the interface. Our goal is to create a system which requires minimal training time for a user to be able to utilize its basic functions, while creating a flexible, dynamic interface adaptable to multiple applications.

As might be expected, subjects made more errors as the edges became narrower. Also, as the edge got narrower, more time was needed to successfully reach the end node. Furthermore, the more time a participant spent on a given edge, the higher percentage of errors were made, relative to time. This relationship of errors to distance and time corresponds well with Fitts' Law [8]. Perhaps an even more appropriate tool for analysis can be found in the Accot-Zhai steering law, which states that the time T required to traverse a straight "tunnel" of constant width W and length A can

be approximated by
$$T = a + b \frac{A}{W}$$
, where a and b are

constants.

Our findings appear to follow this prediction, perhaps with factors such as adventitious blindness versus congenital blindness being represented by different constant values.

One assumption we originally made, that may no longer be valid, is that exiting an edge during traversal is in fact an error, and that a large number of exits represents difficulty in traversal. Particularly for the longer edges, we observed that some subjects employed creative strategies for traversal of an edge. One example was "tacking", which resulted in a large number of errors in a relatively short traversal time. Much like sailing into the wind, subjects would zigzag, or "tack" along the line, constantly going in and out of the edge. Though resulting in a high error rate and consequently a high B score, subjects were able to confidently navigate the graph.

The narrower edges proved most difficult for subjects during the "short" edge tests. However, the largest number of errors consistently and predictably appeared when longer edges were

encountered. This phenomenon appeared during the "full screen" edge tasks, where regardless of the width of the edge the error rates and exploration times were significantly higher than those encountered in the small quadrant tasks (see Figure 5). The high rate of error for long edges can be minimized, of course, by reducing the length of edges in such graphs as much as possible.

	Mean Time (t)	Mean Error (e)	Error /Time (B)
All edges	23.81	10.37	0.44
Full-screen edges	19.93	39.38	0.51

Figure 5: Mean performance comparisons of all trials vs. Full-Screen edges.

Interestingly, there was no appreciable bias in edge direction or orientation between subjects. (There was too much individual variation with a high standard deviation.) Within subjects, however, we found that individuals have very definite strengths and weaknesses when attempting to navigate different quadrants of the display or with particular tracing angles or directions. These differences may be explained by variations in each individual's physiology, tactics and mental representations of the display. However, the difficulties encountered are individualized enough that no angle, quadrant, or direction was found to offer significantly more difficulty than any other.



Figure 6: Error as a function of edge width.

While examining data, we also discovered significant differences in the way congenitally blind and adventitiously blind people navigated the trials. Congenitally blind users spent much more time traversing edges and made more errors in the process (see Figure 6). Additionally, CB participants were much more sensitive to the width of the edge, losing accuracy at a higher rate than their AB counterparts as the edge became narrower. We speculate that this may be due to several factors. CB individuals may have rarely or never held a pen or stylus. Another factor may be that the mental representations used by CB participants differ from those of AB individuals who might have seen a laptop, or who may have visually read diagrams and maps at one time in their lives. Interestingly, while the total times and the total error rates for the two groups differed, the ratio of error to time was very similar between groups.

5. CONCLUSION

Through the process of human testing, we have gained a great deal of data, both quantitative and qualitative. Some of these findings may seem to be truisms: for example, longer and narrower edges are harder to follow than shorter and wider ones. However, we felt that it was important to gain a baseline understanding of the limitations of our interface before proceeding to more elaborate and ambitious goals.

We also found a few results that we did not expect, such as the discovery that there was no overall preference for one angle or position over another, despite the clear preferences that individuals demonstrated. Additionally, our assumptions about how people would use our interface were challenged by the variety of strategies we witnessed, such as "tacking".

When evaluating these results, we wonder how much more proficient users can become with extensive training. It is not surprising that congenitally blind users are slower on average than adventitiously blind users since they have less experience using a pen. It is our hope that training will reduce this gap.

Additionally, our preliminary trials of PLUMB have alerted us to the fact that this tool may be better suited for graphs with geographic information than for graphs with purely abstract information. While we will be continuing to develop PLUMB for the Tablet PC, we are also branching off with new projects to address this limitation through the use of other input methods, such as joysticks and keyboard shortcuts.

6. ACKNOWLEDGMENTS

We would like to thank Emily Higgins for her ongoing help, support, and useful ideas.

7. REFERENCES

- Accot, J., Zhai, S, Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1997, pages 2295-302.
- [2] Blenkhorn P., Evan D. G. Using Speech and Touch to Enable Blind People to Access Schematic Diagrams. Academic Press, 1998, vol. 21, no. 1, pp. 17-29.
- [3] Brewster, S.A. Theses: Providing a structured method for integrating non-speech audio into human-computer interfaces. 1994. Available at http://www.dcs.gla.ac.uk/~stephen/publications.shtml#4.

- [4] Brown A., Pettifer S., Stevens R., Evaluation of a Non-Visual Molecular Browser. In Proceedings of ASSETS'04 (Atlanta, GA USA, October 18-20, 2004), 2004, pp. 40-47.
- [5] R.F. Cohen, A. Meacham, and J. Skaff. Teaching Graphs to Visually Impaired Students Using an Active Auditory Interface, *Proceedings of the 2006 Technical Symposium* on Computer Science Education (SIGCSE 2006), pages 279-282 (2006).
- [6] R.F. Cohen, R. Yu, A. Meacham, and J. Skaff., PLUMB: Displaying Graphs to the Blind Using an Active Auditory Interface, Proceedings of the 7th international ACM SIGACCESS Conference on Computers and Accessibility (ASSETS'05), pages 182-183 (2005).
- [7] *Cornell News*, Available at http://www.news.cornell. edu/releases/Jan05/Wong.software.to.html
- [8] Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47, 381-391 (1954).
- [9] Holt R. C., Winter A., Schürr A., GXL: Towards a Standard Exchange Format. In Proceedings 7th Working Conference on Reverse Engineering (WCRE 2000). Science of Computer Programming, Elsevier Science, 2005, pp. 34. Available at http://www.gupro.de/GXL/.
- [10] Kennel A.R. Audiograf: a diagram-reader for the blind. In Proceedings of the second annual ACM conference on Assistive technologies (Vancouver, British Columbia, Canada). ACM Press, New York, NY, 1996, 51-56.
- [11] Kulikov, S., MacKenzie, I. S., Stuerzlinger, W. Measuring the effective parameters of steering motions. Extended Abstracts of the ACM Conference on Human Factors in Computing Systems. ACM Press, New York, NY, 2005, 1569-1572.
- [12] Mynatt, E., Auditory Presentation of Graphical User Interfaces, in Auditory Display Sonification, Audification, and Auditory Interfaces, G. Kramer Ed., 1994 pp. 533-555.
- [13] The TeDUB-Project (Technical Drawings Understanding for the Blind). Available at http://www.kinf.wiai.unibamberg.de/research/projects/TeBUB.
- [14] van den Doel, K., Smilek, D., Bodnar, A., Chita, C., Corbett, R., Nekrasovski, D., and McGrenere, J. Geometric Shape Detection with Soundview. In *Proc. Tenth Int. Conf.* on Auditory Display (ICAD 2004), Sydney, Australia, July 2004, pp. 1-8.