BROWSING MODES FOR EXPLORING SONIFIED LINE GRAPHS

Lorna Brown, Stephen Brewster,
Ramesh Ramloll and Wai Yu
Glasgow Interactive Systems Group
Department of Computing Science
University of Glasgow G12 8RZ, UK
{lorna, stephen} @dcs.gla.ac.uk
www.multivis.org

Beate Riedel
Department of Psychology
University of Glasgow
G12 8QB, UK
beate@psy.gla.ac.uk

ABSTRACT

The research presented here addresses how line graphs can be made accessible to blind and visually impaired people through the use of non-speech audio. Previous research has focused on the sonification of a single data series. This paper describes a method of sonifying two data series, and the design of two modes (parallel mode and serial mode) for browsing the graphs. Quantitative evaluation showed that both browsing modes were successful, with the parallel mode being particularly effective.

Keywords

Sonification, blind, visually impaired, line graphs, sound graphs, non-speech audio.

1. INTRODUCTION

Blind and visually impaired people often face a lack of access to information due to the dominance of visual communication methods. This is particularly true if they wish to study subjects like Physics or Maths, which use graphs, tables and visualisations widely to display information. The MultiVis Project aims to provide blind people with access to this type of data through a multimodal approach using the senses available to blind people, namely hearing and touch.

Previous research has shown that non-speech audio can be used to present line graphs to blind people but has concentrated on graphs containing only one data series (i.e. one line) [4, 6]. Sighted users are able to identify relationships between multiple data series, such as intersections, and overall maximum and minimum points, with ease. As a first step towards enabling blind people to gain access to multiple data series, the research presented here was aimed at designing and evaluating a method of presenting line graphs containing two data series to blind users using non-speech audio.

Proceedings Volume 2 of the
16th British HCI Conference
London, September 2002

2. SONIFICATION OF LINE GRAPHS

2.1 Sonification of a Single Data Series

The term *sonification* refers to the use of non-speech audio to represent data [8]. Mansur [4] devised a method for line graph sonification, called *Sound-Graphs*, from which the method used here was derived. The y-axis of a graph is mapped to pitch, and the x-axis is mapped to time such that moving along the x-axis causes a musical note, representing the corresponding y-value, to be played. The higher the y-value, the higher the note sounds. SoundVis (the graph sonification software developed for this research) converts data to sound by mapping data values to MIDI (Musical Instrument Digital Interface) notes. The range of MIDI notes used is from 35 (B1) to 100 (E7) because notes out-with this range can be difficult to perceive and differentiate. The General MIDI Acoustic Grand Piano sound was used here as it has been found that harmonically complex sounds, such as the piano, can be perceived easily [1].

2.2 Sonification of Two Data Series

The aim of this research was to see whether it is possible to communicate information about line graphs containing two data series to blind people effectively using sound. The data series were sonified as described in Section 2.1. Using stereo panning, one data series was assigned to each ear through headphones. This perceptual separation helped the user to distinguish between the two series.

3. BROWSING MODES

3.1 Design of Browsing Modes

Interpreting a line graph involves more than just looking at it, or listening to it, once. A person may wish to browse back and forward through a graph to explore points of interest. Two browsing methods were designed and tested to evaluate which is the more effective. These are *parallel* and *serial* modes.

In the parallel mode users hear the two data series simultaneously (one in each ear). Thus, at any given point on the x-axis they hear the note representing the y-value for both the left and right series. In the serial mode, users listen to one data series at a time. Thus, at a given x-coordinate, they will hear a single note representing the y-value in one ear and can switch to the other ear to hear the corresponding y-value for the second series, allowing comparisons to be made.

3.2 SoundVis Interface

Blind people generally interact with computers using the computer keyboard. For this reason, SoundVis uses the numeric keypad for graph browsing. The key functions are defined in Figure 1 and were derived from those used by Ramloll [5] for browsing tables in sound. Holding down a key causes its function to be repeated, so a rapid overview of the graph can be heard by holding down the 2 key. The use of a footswitch for switching between modes reduces the complexity of the keyboard interface.

Key	Function
2 key (down arrow)	move one position right along x-axis and play sound corresponding to y-value
8 key (up arrow)	move one position left along x-axis and play sound corresponding to y-value
5 key	stay on same position and repeat the sound corresponding to y-value
4 key (left arrow)	select left data series (in serial mode)
6 key (right arrow)	select right data series (in serial mode)
9 key (page up)	jump to the start of the graph
3 key (page down)	jump to the end of the graph
Footswitch	switch between modes (Experiment 2 only)

Figure 1: Key Functions in SoundVis

3.3 Evaluation of Browsing Modes

Two experiments were carried out to explore whether blind users can successfully interpret graphs containing two data series. These experiments aimed to discover whether either browsing mode is more successful than the other, and whether users display a preference for one mode over the other, when carrying out certain tasks.

Each experiment required participants to do the following tasks: to find the point at which two lines intersect, and to find the overall maximum or the overall minimum point on a graph. An intersection is a point at which two lines on a graph cross one another because they share the same y-value. In terms of sound, the two data series play the same musical note. The overall maximum or minimum point is the very highest or very lowest point on a graph which, when sonified, is represented by the very highest or very lowest musical note.

4. EXPERIMENT 1: PERFORMANCE

4.1 Aim and Hypotheses

The aim of this experiment was to discover what level of performance could be achieved when users were asked to carry out tasks involving the exploration of two data series, and to identify any differences in performance between the two browsing modes. We believed that finding intersection points would be easier in parallel mode, since comparisons between two streams should be simpler when both are heard simultaneously. Conversely, we felt that it would be easier to identify maxima and minima in serial mode since this mode enables users to focus on a single tone without the distraction of hearing a different note in their other ear. The hypotheses were:

1. Intersection points will be found quicker in parallel mode than in serial mode.
2. Overall maxima/minima will be found quicker in serial mode than in parallel mode.
3. Accuracy levels for finding intersection points will be higher in parallel mode than in serial mode.
4. Accuracy levels for finding overall maxima/minima will be higher in serial mode than in parallel mode.
5. The subjective workload for intersection tasks will be lower in parallel mode than in serial mode.
6. The subjective workload for maxima/minima tasks will be lower in serial mode than in parallel mode.

4.2 Method

Following a pilot study with sighted users, this experiment was run with 16 registered blind participants (aged 16 to 60 with a variety of visual impairments) from the Royal National College for the Blind in Hereford. The experiment was divided into four sections according to task and mode: intersection points in parallel mode, intersection points in serial mode, maxima/minima in parallel mode, and maxima/minima in serial mode. The order in which participants took part in these sections, and the order in which the graphs were presented, were fully counterbalanced to offset any learning effects.

Prior to the experiment, participants were introduced to basic graph concepts through tactile exploration of a raised line graph. They were then trained in the use of the SoundVis interface through two training sessions, one in each browsing mode. In each training session, participants were provided with a sonified graph containing two data series in which they had to find the intersection point, overall maximum and overall minimum before progressing to the experiment itself. The experimenter provided guidance during the training if required. This training allowed different participants to start the experiment at similar levels of knowledge and skill.

Each section of the experiment consisted of three tasks, in each of which a sonified graph containing two data series was presented. All data series were smooth and had been generated from basic mathematical functions, e.g. sine, cosine. Participants had to explore the graph and identify one specific point (either the intersection point, overall maximum or overall minimum) within a time limit of four minutes. When the point was found, participants informed the experimenter who then pressed a key to record the current position and time. At the end of each section NASA-TLX (Task Load Index) scales were completed to assess the subjective workload experienced by participants (for further information on NASA-TLX see Hart & Staveland [3]).

4.3 Results

A significant difference in the average task completion time between the two modes was found. A two-way ANOVA showed a significant effect of mode ($F(1,60)=20.51$, $p<0.001$), and a significant interaction between mode and question type ($F(1,60)=12.79$, $p=0.001$). Post hoc Tukey tests revealed that this was due to the intersection tasks, where the average task completion time was significantly shorter in parallel than in serial mode ($T=5.73$, $p<0.001$, Figure 2). This supports Hypothesis 1 in Section 4.1. There was no significant difference in the average task completion time between modes for maxima/minima tasks.

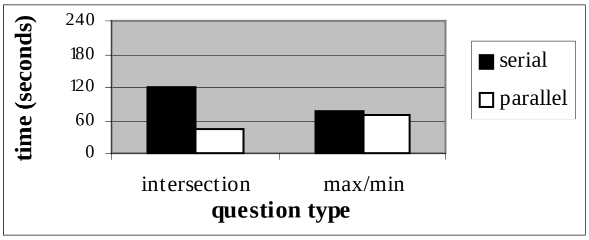

Figure 2: Average Time per Task

There was no significant difference in the percentage of correct answers between modes for either question type (Figure 3). An answer was considered correct if the exact location was identified. Overall, performance was low, as only 51% of intersections and 68% of maxima/minima were correctly identified. There was no significant difference in the subjective workload between modes for either question type.

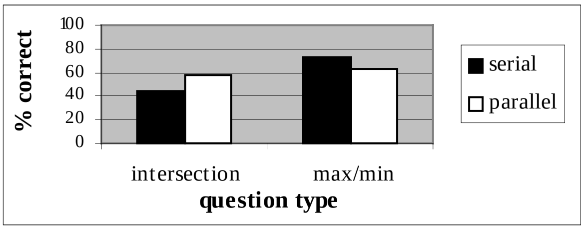

Figure 3: Percentage of Correct Answers

5. EXPERIMENT 2: PREFERENCE

5.1 Aim and Hypotheses

The aim of this experiment was to build on the previous one and investigate whether blind users display any preference for one mode over the other when allowed to switch between modes freely. We believed that participants would prefer parallel mode for intersection tasks, and serial mode for maxima/minima tasks for the reasons outlined in Section 4.1. The hypotheses were:

1. More time will be spent in parallel mode than in serial mode when looking for intersection points.
2. The final selection of intersection points will mostly occur in parallel mode.
3. More time will be spent in serial mode than in parallel mode when looking for maxima/minima.
4. The final selection of maxima/minima will mostly occur in parallel mode.

5.2 Method

14 participants from the previous experiment took part in this experiment. A short training session was carried out to refresh the concepts learned in Experiment 1, and to introduce the footswitch that would enable participants to switch between modes.

This experiment consisted of eight tasks, the order of which was randomised for each participant. In each task, participants were presented with a sonified graph containing two data series and asked to find one specific point (either the intersection point, overall maximum or overall minimum) within four minutes. Data were recorded about the number of mode changes and the amount of time spent in each mode.

5.3 Results

A two-way ANOVA analysing the percentage of time spent in each mode showed a significant effect of mode ($F(1,52)=15.21$, $p<0.001$) and a significant interaction between mode and task type ($F(1,52)=45.01$, $p<0.001$). A significant effect of mode ($F(2,78)=24.82$, $p< 0.001$) and a significant interaction ($F(2,78)=21.85$, $p<0.001$) were also found for the mode in which points were selected.

Post hoc Tukey tests revealed that participants spent significantly more time in parallel than in serial mode ($T=-7.5$, $p<0.001$, Figure 4) while looking for intersection points, and that the final selection of these points occurred significantly more often in parallel than in serial mode ($T=-6.014$, $p<0.001$, Figure 5). This preference for parallel mode supports hypotheses 1 and 2 in Section 5.1.

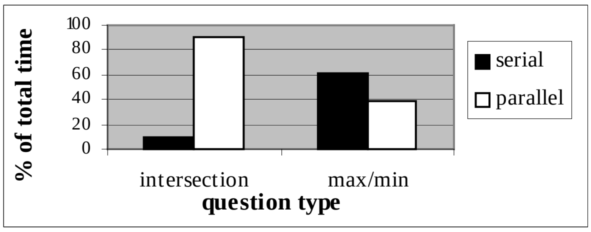

Figure 4: Mean Percentage of Time in each Mode

While there was no significant difference in the amount of time spent in each mode when looking for maxima or minima (Figure 4), the final selection of these points occurred significantly more often in serial than in parallel mode ($T=3.324$, $p=0.016$, Figure 5). This indicates some preference for the serial mode for maxima/minima tasks, and supports hypothesis 4. No support for hypothesis 3 was found.

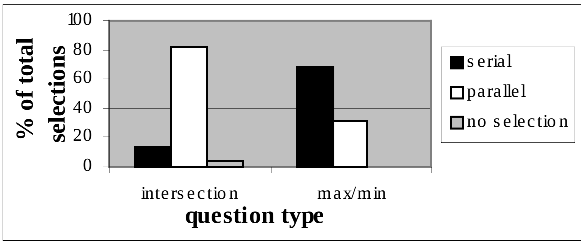

Figure 5: Percentage of Times Points Selected by Mode

6. CONCLUSIONS AND FUTURE WORK

The results reported here show that it is possible for blind people to interpret sonified graphs containing two data series, but that accuracy is low, particularly for the finding of intersections. The two browsing modes were found to be useful for different types of task. For this reason, graph sonification systems should feature both browsing modes, but default to parallel mode, since a clear preference was displayed for this mode.

A possible reason for the low accuracy in intersection tasks is that participants regularly mistook several other musical intervals (octaves, perfect 5ths, semitones and perfect 4ths) for intersection points. The theories of affinity of tones and octave equivalence explain that listeners often perceive the two notes in a perfect 5th or octave interval as similar, or even the same [7]. This problem is intensified by the harmonic complexity of the piano sound (the perfect 5th and octave are heard as harmonics of the original note) and may be solved by using sine wave tones, for which octave equivalence is less pronounced [7]. Semi-tone intervals sound close and dissonant, which may have confused some participants. Further research is required to establish why perfect fourths were mistaken for intersections.

The pilot group of sighted users achieved higher levels of accuracy for intersection tasks than the blind users, indicating that high levels of performance can be achieved. However, a direct comparison is inappropriate, due to changes in the experiment following the pilot study. It is possible that the blind participants did not have a thorough enough understanding of the graph concepts. More exhaustive training may solve these problems. A further possibility is to add some audio cues to indicate interesting points saliently (e.g. a percussive sound when an intersection occurs).

The low accuracy for finding maxima/minima may have been due to memory limitations for pitch [2]. This could be addressed by providing a *position saving tool*. This would allow users to save positions to which they could return at any time without moving through the rest of the graph.

A potential solution to all these problems would be to use a multimodal approach, combining sound and touch. This would allow users to employ their available senses to facilitate the task at hand.

Providing this tool will allow blind and visually impaired users to gain more independent access to this type of information, which is in common use in the visual world. Future research will look at additional tasks such as identifying whether graphs are in phase with one another. It is hoped that the method presented here will provide a basis from which methods for the sonification of multiple data series can be developed.

7. ACKNOWLEDGEMENTS

This research is funded by EPSRC grant GR/M44866 and ONCE, Spain. Thanks to all of the participants from the Royal National College for the Blind.

8. REFERENCES

[1] Brewster, S.A., Wright, P.C. & Edwards, A.D.N. Experimentally derived guidelines for the creation of earcons. In *Adjunct Proceedings of HCI'95* (Huddersfield, UK.), 1995, pp. 155-159.

[2] Deutsch, D. Memory and Attention in Music. In *Music and the Brain*, Critchley, M., Henson, R.A. (Ed.), William Henemann Medical Books Limited, London, 1977, pp 95-130.

[3] Hart, S., Staveland, L. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human Mental Workload*, Hancock, P., Meshkati, N. (Ed.), North Holland B.V., Amsterdam, 1988, pp 139 - 183.

[4] Mansur, D.L. *Graphs In Sound: A Numerical Data Analysis Method for the Blind*. M.Sc., University of California, 1975.

[5] Ramloll, R., Brewster, S, Yu, W., Riedel, B. Using Non-speech Sounds to Improve Access to 2D Tabular Numerical Information for Visually Impaired Users. In *Proceedings of BCS IHM-HCI* (Lille, France) Springer, 2001, pp. 515-530.

[6] Sayhun, S. *A Comparison Of Auditory And Visual Graphs For Use In Physics And Mathematics.* Ph.D., Oregon State University, 2000.

[7] Terhardt, E., *Perception of Musical Sound*, http://www.mmk.ei.tum.de/persons/ter.html

[8] Wilson, C.M., Lodha, S.K. Listen: A Data Sonification Toolkit. In *Proceedings of ICAD '96* (Palo Alto), 1996, pp. 35-40.