CONTEXTUAL AUDIO IN HAPTIC GRAPH BROWSING

David K. McGookin Stephen A. Brewster

The MultiVis Project
Department of Computing Science
University of Glasgow, G12 8QQ
{mcgookdk,stephen}@gla.ac.uk
www.multivis.org

ABSTRACT

This paper presents a "think-aloud" study investigating the ability of visually impaired participants to make comparisons between haptic and audio line graphs. Graphs with two data series were presented. One data series was explored with a PHANTOM haptic device, whilst the other was sonified using one of two data - sound mappings. The results show that participants can make comparisons between the two lines. However, there is some cross-modal interference which makes it difficult to extract detailed information about the data series presented in audio.

1. INTRODUCTION

Access to graphical data is a significant problem for visually impaired people. Whilst textual data can be easily accessed using screen reader software such as JAWS (www.freedomscientific.com), data such as graphs and tables presents a significant problem. Several researchers have proposed and evaluated the use of sonifications to provide access to this data with systems such as the Sonification Sandbox [1] allowing easy access to data for visually impaired people.

Interaction with such sonifications however loses the spatial relationships that exist in the visual representations of graphs. For example, where a particular bar is in a bar graph, or where (rather than when) a turning point in a line graph occurs. This loss of a common vocabulary to describe the graphs may make it difficult for sighted and visually impaired colleagues to collaborate using such information [2]. Additionally sonifications can be difficult to directly manipulate in the way that is available with visual interfaces. Cohen's work in filtears [3] is a notable exception; however it is only recently that the interactive control of sonification has received greater prominence [4].

One solution to this lack of spatial awareness is to incorporate the addition of haptic (or touch) interaction alongside the use of sonification. A haptic device such as SensAble technologie's PHANTOM (see Figure 1), which uses motors and optical sensors to provide dynamic information through the sense of touch, can provide a control mechanism for interactive sonification control as well as provide a fixed reference frame device to provide spatial interaction with a sonification [5]. Work by Yu and Brewster [6] has shown that haptic line graphs can be successfully interpreted by visually impaired users. There remain problems however, since devices such as the PHANTOM only allow one point of contact with the data. If a line graph were to contain two data series, only one could be browsed at a time making it difficult to determine the relationship between the two data series; a relationship that would be visually obvious. Whilst it is possible to present multiple data series via sonification, it is difficult to combine multiple simultaneously sonified data series without those data series interfering with each other making understanding difficult [7].

Figure 1: A picture of a SensAble Technologies' PHANTOM Omni Haptic Device. The use of motors and optical sensors allows the user to "feel" virtual objects.

Since the communications bandwidth of the auditory and haptic systems is limited in comparison to the visual system, the most appropriate approach to present information to visually impaired people would be to exploit both the haptic and auditory systems in conjunction to communicate data. Doing so allows the advantages of each modality to be used to improve data access [6]. Using such a system, one data series could be communicated through the haptic modality, whilst the other could be sonified. However when browsing line graphs with multiple data series users would need to make comparisons between the data communicated using different modalities. With such cross-modal presentation however there are problems that may arise when data are presented to different modalities [8]. In this paper we present the results of an initial study investigating the cross-modal presentation of haptic line graphs to visually impaired people. In the following section we review existing research on cross-modal presentation before introducing our study and the results obtained.

2. BACKGROUND AND CONTEXT

The investigation of cross-modal integration and equivalence is not new. Flowers *et al.* [9] have found a high correlation between participants ability to make judgments on data which was presented via auditory and visual graphs, although the auditory and visual representations were not simultaneously presented. McGee [8] has performed several experiments and identified that audio can influence the perception of a haptic texture, making it appear to the user as more or less rough.

Whilst work has been undertaken on cross-modal perception, our work is novel in that no previous work has attempted to present different data to different modalities and ask users to make comparisons between them. This is important as if users can successfully integrate information from the two modalities, such interpretation can be exploited to provide improved access to graphs. This work is also important as prior research indicates that the interaction between modalities is not clear and must be investigated on a case by case basis [10]. As McGee [8] states *"When a multi sensory percept is processed, the level at which the combined stimuli are integrated into the whole piece of information can vary. Sensory stimuli can be ignored or attended to; used alongside each other to reinforce a mental representation generated by one of the stimuli; combined in an additive way to produce a greater effect than either single stimuli; or processed against each other in competition to produce a cognitive conflict.*". Whilst it is reasonable that cross-modal presentation can improve access to information for visually impaired people, it is unclear what the underlying perceptual and cognitive outcome will be. It is therefore important to investigate the issue of cross modal integration of line graph data to identify whether the two modalities interfere with each other, or complement each other providing information that would not be available by one modality alone.

3. STUDY OUTLINE

In order to investigate the issues previously discussed, a study involving seven participants from the Royal National College for the Blind (RNCB) in Hereford, UK was carried out. All participants were visually impaired and were paid £15 for their participation. Each participant was asked to browse and describe line graphs, each containing two data series (See Figure 2). In all cases one data series was browsed using a PHANTOM haptic device *"haptic line"*, whilst the other was simultaneously sonified *"audio line"*. As the user moved the PHANTOM device along the haptic line, a MIDI note corresponding to the position on the audio line, at that x position, was played to the user. This allowed the user to interactively control the speed of presentation of the audio line. The MIDI note presented depended on the data to sound mapping used. Two different mappings were used, each of which provided contextual information about the audio line in relation to the haptic line, but providing varying amounts of detail.

Whilst the mappings do not represent all of the possible data - sound mappings that can be applied to the audio line, they do represent two points in the complexity of sonifications that could be presented in the scenario described above. As such they provide a good foundation to start investigating cross-modal haptic and audio interaction in line graphs.

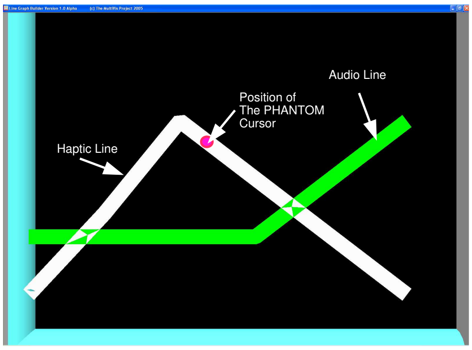


Figure 2: Illustration of the graphs that were browsed during the study, and how they were presented to the users.

3.1. Two Pitch Mapping

In this system the audio line was represented using two different pitches. If the audio line was *below* the haptic line, a low note (General MIDI note 35) was played on the general MIDI piano (MIDI patch number 1). If the audio line was *above* the haptic line, a high note (General MIDI note 100) was played also on the same instrument. These notes being at the extremities suggested by Brown and Brewster [11]. With this mapping the user is provided with only low pass information about the audio line and its relationship with the haptic line. However, salient information about the relationship between the two lines can still be determined, such as intersections and relative positions of the two lines.

3.2. Standard Pitch Mapping with Two Timbres

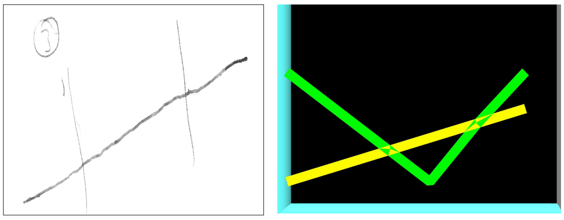
This sonification provides more detailed information about the audio line. It uses a standard mapping between y-axis value and pitch according to the equation derived in Brown and Brewster [11] (see Equation 1), which maps the y-axis value on a uniform scale to the MIDI notes 35-100.

$$note = \left(\frac{yAxisValue}{(yAxisValue_{max} - yAxisValue_{min})} * 65\right) + 35 \quad (1)$$

To provide information about the relative positions of the lines to each other, two timbres were used. When the audio line was *below* the haptic line a French horn timbre (general MIDI patch 60) was used. When the audio line was *above* the haptic line a piano timbre (General MIDI patch 1) was used. These patches allowing all of the notes used to be played well.

3.3. Method

The study followed a "think aloud" methodology with participants encouraged to verbally describe each of the line series in the graph, and the relationship between them, in as much detail as they could. Participants were firstly introduced to the PHANTOM haptic device. At this stage the tasks that they would be carrying out and the general relationship between what the user felt and what they heard was explained with the aid of a raised paper diagram. Once



(A) (B)

Figure 3: Illustration of a graph drawn by a participant in the two-pitch condition (A), and (B) a screenshot of the graph.

this was explained participants proceeded to carrying out one of the conditions. Ideally, the order in which participants carried out conditions would be counterbalanced, however several of the participants found the standard pitch mapping with two timbres difficult to grasp at the same time as the general relationship between the haptic and audio lines. As such the two pitch condition was generally performed first, followed by the standard pitch mapping with two timbre condition. In all cases participant comments were recorded for later transcription. Additionally, one participant was able to provide drawings of the graphs on German film (a special paper that leaves tactile ridges when drawn on with a pen).

4. RESULTS

The comments given by participants, and notes taken by the experimenter were transcribed and analysed. In this section we outline the main findings in terms of access to general, low-pass, information in the graph, detailed information about the audio line and other useful comments made by the participants.

4.1. Low Pass Information

One of the key features of a graph containing two data series is the relationship between those data series. This can largely be categorised as the identification of crossing points (where one data series crosses the other), and the relative proportion of each data series that is higher than the other, such as might be important on a graph of one company's sales in comparison to another.

In both conditions participants were able to identify intersection points. Additionally many participants were able to correctly identify the direction of some of the crossings, i.e. whether the audio line moved from below the haptic line to above it, or vice versa. Figure 3 for example shows the sketch made by one of the participants in the two pitch mapping condition. The participant has been able to correctly identify the crossings and the direction of the first crossing. Whilst participants were able to identify this information they did not find it easy, many said that it was a difficult concept to get their head around, with one noting *"Having something to hear and something to feel and trying to compare the two I think is quite strange"*.

4.2. Cross-Modal Interactions

Whilst participants can retrieve the low pass information from both of the conditions, albeit with some effort, it was difficult for participants to extract more information from the graph. Participants found it very difficult to describe the audio line in any detail, obviously the two pitch condition does not communicate such information, but even when such information was available in the sonification participants found it difficult to extract. From comments participants made, it appears that the perception of the audio and haptic lines interfered with each other, with the haptic line taking precedence. As one participant stated when carrying out the standard pitch with two timbre condition *"I cant tell what the other (audio) line is without getting distracted by the haptic line"*. Some participants were able to discuss the shape of the audio line in cases where either the audio or haptic lines were "predictable", i.e. they were largely straight. See Figure 4 for an example. However even in such cases mistakes were still made. For example, one participant stated that the audio line of the graph shown in Figure 4, was "M" shaped. Thus in spite of getting the first part of the audio line correct, the second part appears to be incorrect. Unfortunately the data recorded in this study does not provide reasons for this. Klatzky and Lederman [10] propose that the modality that is most appropriate for the task dominates the other, however their work largely discusses real world phenomena, rather than the "artificial" scenario discussed here. In some cases where participants could not discuss the audio line in detail, the experimenter took control of the PHANTOM and moved along the haptic line, causing the audio line to be played. In such instances, participants were able discuss the shape of the audio line. This indicates that haptic information was interfering with the perception of the audio line.

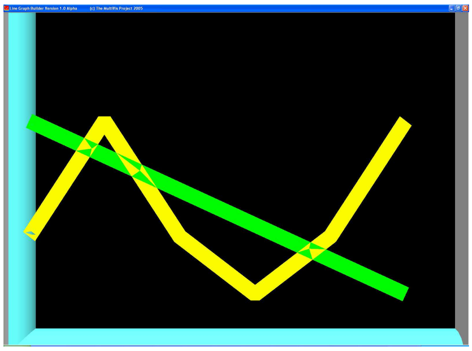


Figure 4: An example "predictable" graph that was explained in detail in the standard pitch mapping with two timbres condition. The graph is "predictable" since the haptic line is straight.

4.3. Other Comments

In addition to the comments which directly relate to comparing the data represented through the haptic and audio modalities, participants also discussed other features of the interaction between sound and haptics that they used to help complete the task. One participant commented that although they found it easy to slip off of the haptic line, the absence of audio alerted them to this fact. The audio acting as a reassurance that the user was still connected to the haptic line. In addition users discussed that they liked the use of the PHANTOM as a control mechanism, allowing them to move faster or slower through the sonifications as they wished; such control being useful if the user is trying to find the exact sport of an intersection (moving slowly), or trying to get a quick overview (moving more quickly).

5. CONCLUSION

From the comments made by participants during the "think-aloud" study, we can consider that undesirable cross-modal interactions occurred when participants tried to retrieve detailed information from the audio line when using the more detailed sonification (standard pitch mapping with two timbre conditions). Even when trying to concentrate solely on the audio line, participants still described that the haptic line was interfering. Participants could however determine where the crossing points were, as well as the direction of those crossing points, indicating that at least some contextual information could be extracted from the audio line. As such the two pitch condition, which only provides such information, is preferable to the standard pitch mapping with two timbres.

The information provided by the sonifications in this study may be useful in haptic graph displays where only one data series can be browsed at a time. The use of contextual audio providing information about other data series in the graph, even if only low pass information is available, may help the user to construct a more effective "mental picture" [12] of the relationship between the two data series. Since users can gain low pass information via different modalities it would be useful to identify to what extent such information assists in creating such a picture.

6. ACKNOWLEDGEMENTS

We would like to thank all participants and staff at the RNCB Hereford. This work is supported by EPSRC grant GR/S86150/01.

7. REFERENCES

[1] B. N. Walker and J. T. Cothran, "Sonification sandbox: A graphical toolkit for auditory graphs," in *Proceedings of ICAD 2003*, Boston, Massachusetts, 2003, pp. 161–163, ICAD.

[2] F. Winberg and S. O. Hellstrom, "Qualitative aspects of auditory direct manipulation a case study of the towers of hanoi," in *Proceedings of ICAD 2001*, Espoo, Finland, 2001, vol. 1, pp. 16–20, ICAD.

[3] M. Cohen and L. F. Ludwig, "Multidimensional audio window management," *International Journal of Man-Machine Studies*, vol. 34, pp. 319–336, 1991.

[4] T. Herman and A. Hunt, "An introduction to interactive sonification," *IEEE Multimedia*, vol. 12, no. 2, pp. 20–24, 2005.

[5] C. Sjostrom, "Using haptics in computer interfaces for blind people," in *CHI 2001*, Seattle,Washington, 2001, vol. 1, pp. 245–246, ACM Press.

[6] W. Yu and S. A. Brewster, "Evaluation of multimodal graphs for blind people," *Universal Access in the Information Society*, vol. 2, no. 2, pp. 105–124, 2003.

[7] J. H. Flowers, "Thirteen years of reflection on auditory graphing: Promises, pitfalls, and potential new directions," in *First International Symposium on Auditory Graphs (AGS2005)*, Limerick, Ireland, 2005, pp. 1–5.

[8] M. R. McGee, *Investigating a multimodal solution for improving force feedback generated textures*, Ph.D. thesis, University of Glasgow, 2002.

[9] J. H. Flowers and T. A. Hauer, "Musical versus visual graphs: Cross-modal equivalence in perception of time series data," *Human Factors*, vol. 37, no. 3, pp. 553–569, 1995.

[10] R. J. Klatzky and S. J. Lederman, "Touch," in *Handbook of Psychology: Volume 4 Experimental Psychology*, Alice F Healy and Robert W Proctor, Eds., vol. 4 of *Handbook of Psychology*, pp. 147–176. John Wiley & Sons, Hoboken, Hew Jersey, 2003.

[11] L. M. Brown, S. A. Brewster, R. Ramloll, M. Burton, and B. Ridel, "Design guidelines for audio presentation of graphs and tables," in *Proceedings of ICAD 2003*, Boston, Massachusetts, 2003, pp. 284–287, ICAD.

[12] S. Pinker, "A theory of graph comprehension," in *Artificial Intelligence and the Future of Testing*, Roy Freedle, Ed., pp. 73–126. Laurence Erlbaum Associates, Hillsdale, New Jersey, 1990.