Non-Visual Overviews of Complex Data Sets

Johan Kildal and Stephen A. Brewster
Glasgow Interactive Systems Group
Department of Computing Science
University of Glasgow
Glasgow, G12 8QQ, UK
{johank, stephen}@dcs.gla.ac.uk
www.multivis.org

CHI 2006, April 22–27, 2006, Montréal, Québec, Canada.
ACM 1-59593-298-4/06/0004.

Abstract
This paper describes the design and preliminary testing of an interface to obtain overview information from complex numerical data tables non-visually, which is something that cannot be done with currently available accessibility tools for the blind and visually impaired users. A sonification technique that hides detail in the data and highlights its main features without doing any computations to the data, is combined with a graphics tablet for focus+context interactive navigation, in an interface called TableVis. Results from its evaluation suggest that this technique can deliver better scores than speech in time to answer overview questions, correctness of the answers and subjective workload.

Keywords
Sonification, tablet, data tables, overview, accessibility.

ACM Classification Keywords
H.5.2. [Information Interfaces and Presentation]: User Interfaces---auditory (non-speech) feedback.

Introduction
A natural way of approaching a set of information previously unknown to a user starts with gathering general information about the data set (metadata) and obtaining an overview, before any other more detailed analy-

sis of the information can be performed. Shneiderman has put forward his Visual Information Seeking mantra [3] stressing the need to obtain an overview first, then zooming and filtering, and getting details only on demand. Zhao *et al.* [4] have extended this idea to the auditory domain with the Auditory Information-Seeking Principle: gist, navigate, filter, and details-on-demand.

While inspecting a data set visually is quite an efficient way of obtaining overview information quickly, particularly after having constructed appropriate graphical visualisations of the data, blind and visually impaired (VI) users have far greater difficulties in accomplishing this basic task. Current accessibility tools only provide access to fully detailed unitary portions of information sequentially, and do not support mechanisms for summarising information or highlighting the most salient features. Sequential access to fully detailed information leads very quickly to saturation of the user's working memory, and imposes great demand in terms of mental workload as well as being slow. As an example, accessing information in a large spreadsheet is normally done using a screen reader that speaks all the information in the document while the user navigates sequentially across the grid. As the user retrieves the information from each cell, this has to be remembered and compared against the information in every other cell, while also keeping track of the position of cells in the whole table. All this requires large numbers of comparisons and remembering many values and locations. Reading Braille versions of spreadsheet tables presents similar problems and limitations. This inability to obtain a quick overview of data imposes serious limitations for blind and VI users undertaking education in technical subjects and collaborating with sighted colleagues in professional environments.

Non-visual overview of numerical data tables

This paper focuses on the problem of obtaining non-visually a quick overview of two-dimensional numerical data tables. These are data structures that, while being very common in everyday life (and particularly in scientific, technical and economic subjects), provide enough complexity, even with relatively small data sets, to reveal the difficulties described in the introduction. Tables are also a very general data structure that underlie every data visualisation, and bridge with the raw data sets that form those [1]. For this reason, they are a key area of study and solutions developed for tables can be extended to other data structures. Generalisation of these solutions to tables of higher dimensionality should be relatively straightforward. Recent work by Zhao *et al.* [5] sonifying geo-referenced data (in which regions of a geographic map are assimilated to a 2-dimensional table) has been considered in this research.

Requirements capture

Extensive requirements capture was conducted to find out the needs of the VI community, the techniques they currently use to obtain overviews and how it could be done better.

A number of British institutions for blind people were visited, including The Royal National College for the Blind, in Hereford, and The Royal Blind School, in Edinburgh. We conducted interviews with teachers and instructors of blind and VI children and adults, observed lessons of mathematics and algebra with blind children, and conducted focus groups with blind computer users. Advanced blind computer users were also interviewed and several pilots were run to test new accessibility, data exploration and interaction techniques.

This process confirmed the need for new techniques to browse and obtain overviews of complex data sets in general, and of numerical data tables in particular. In terms of design requirements, the system to be designed would have to avoid overloading short term memory, facilitate comparisons between subsets of the data set, support the focus and context paradigm, provide support for collaboration between sighted and VI users working on the same data sets, and offer a high degree of control in the interaction. From these requirements, an interactive table browsing system, called TableVis, was designed.

Sonification technique

Kildal & Brewster's table sonification technique [2] was taken as a starting point for the design of TableVis. With this technique, instead of sonifying each cell in a two-dimensional table, a whole row or column is sonified as a single piece of complex information. Thus, a table can be explored by rows or by columns, and two simplified and complementary views of the same data table are obtained, each one of them containing the complete overview information: an array of rows and an array of columns. To construct the sonification of each row or column, all the values in it are mapped to pitch, dividing the complete range of values in the table to a range of 66 MIDI values, where higher numeric values are mapped to higher pitches. These sounds are played in rapid succession (left to right or top to bottom, depending on the case), hiding complexity of detail and highlighting the more salient features in it. Due to properties of hearing, each sonified row or column is perceived as a single (although more complex) unit of information. As a result, the total number of units of information in the table is reduced from mxn (where m and n are the number of rows and columns) to only m or n depending on the direction of navigation. Thus, the number of units of information becomes a lot more manageable for the user's short term memory. Kildal & Brewster [2] found that an overview of the data in a table is obtained faster using this sonification technique than using screen reading software, without a reduction in accuracy

When sonifying a row or column in this fashion (using MIDI-generated piano sounds), it was observed that the shorter the duration of each sound (i.e. the faster the sonification), the less detailed information is perceived by the user. In the extreme case, when sounds are perceived to be simultaneous, like a chord, the pitch perceived is a blend of all the pitches in that dissonant chord. Prototypes showed that, although little could be told about the precise values forming these chords, carrying out comparisons between them and judging relative average pitches was easy, as was spotting outliers. However, as the duration of each sound was increased (slower sonification), more detail about the structure of each row or column was revealed, but comparing them became more difficult. The user could access local details on demand in speech (coordinates and value) at any time. No computations were done with the values in the cells before rendering them in sound.

Physical input device

A graphics tablet was chosen to control TableVis, as it offers a number of advantages. It is an absolute positioning pointing device. The borders of the working area of the tablet can be felt with both hands (a physical frame was built to identify the borders more easily, as seen in figure 1). The users can move the pen over the area and use their sense of proprioception to judge the position of the pen in relation to the borders, being able to jump to specific regions of the working area directly.

The table to be explored is presented on the working surface of the tablet, scaled to fill it completely. Each cell remains in the same position throughout the exploration. Users can access any region of the table directly, and the distance of the pen relative to the borders informs them of which area of the table is being explored at any time, providing contextual information.

Interaction

TableVis offers three modes of exploration. In the *cells* mode, the cell being pointed at on the tablet is sonified, facilitating 'freehand' exploration of the data table. In *rows* and *columns* modes, a complete row or column is sonified when the pen is pointed at it. These two last modes facilitate obtaining quick overviews of the complete data set. It is enough to draw a line across the whole table, horizontally in columns mode or vertically in rows mode, to generate the data sonification interactively.

An additional input device (a Griffin PowerMate USB rotary knob with a push-button function, also shown in figure 1) was used to modify the sonification speed during exploration and to shift between exploration modes. These functions can be integrated in the built-in controls of some graphics tablet models.

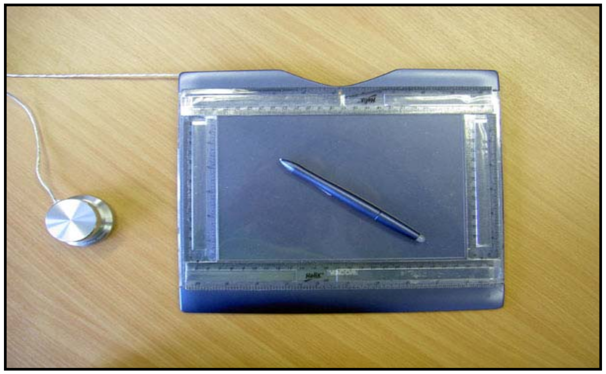

figure 1. Graphics tablet with physical frame delimiting the active area, and a rotary knob with push-button function at the side

Experimental design

A two-condition within-subjects repeated-measures experiment was used to evaluate the TableVis prototype and to compare its performance against an analogous speech-based interface.

The *sound* condition included all the functionality of TableVis as described above, but the cells mode was not made available in order to test the technique used in the rows and columns modes to obtain overview information. The default sonification speed was set to the maximum, with around 10ms sound duration per cell, but the user could modify this during exploration.

In the *speech* condition, the interface design was the same except in that the value in each cell was read out instead of being mapped to pitch. In the rows mode, the cells in the selected row were read from left to right (top to bottom in the columns mode) at a speech rate that was set by users before the experiment, to match the speech rate they normally use with screen reader applications. Using the rotary knob, the users could select the number of digits of precision with which values were spoken, to optimise speed of speech and access overview-relevant information.

	Mon	Tue	Wed	Thur	Fri	Sat	Sun
midnight	68697	70660	71119	71180	71728	68456	70930
1 a.m.	69756	69300	70243	72907	73121	71418	71831
2 a.m.	71986	72264	70900	72872	72358	72186	71940
3 a.m.	72057	73230	76630	75575	75963	72850	73160
4 a.m.	73818	74185	73562	76049	74033	73336	75009
5 a.m.	79294	79247	79223	77668	74573	74960	75830
6 a.m.	80757	79494	81236	79310	75880	81014	75803
7 a.m.	77357	78664	78403	79621	78679	82557	77338
8 a.m.	81236	83197	83780	79208	82885	83667	81231
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

figure 2. An excerpt from a data table, with data about hourly visits to a website in a week

Sets of data tables were produced (see figure 2), all with the same metadata: 7 columns (days of the week) and 24 rows (hours of the day) with data representing the number of visitors in an internationally popular website (to ensure visits are equally possible at any time of the day). A maximum of 2 min. was allowed to explore each table and answer its overview question, which could be of two possible kinds, each one requiring obtaining a different kind of overview information:

1. Which time of the day / day of the week gets more / less visits?
2. In which quadrant of the table (top-left, bottom-left, top-right or bottom-right) are values highest/lowest?

The first question required exploring the table in one direction only, giving a precise answer (a particular row or column). With the second question, the table had to be explored in both directions (horizontal and vertical), but only one of four areas had to be identified.

A group of 6 blind and VI users was recruited for the experiment, all regular computer users with some experience using data tables. Each participant had to obtain an overview of 24 tables, 12 in each condition, with 6 questions of each type in each condition. The experiment was designed fully counterbalanced, with a training session at the beginning, and lasted 60 minutes in total. During the training session, three of the participants showed more difficulties in acquiring the skills needed to use the interface. In those cases, training was longer and there was time only to complete one condition, which was chosen to be the sound condition. For each condition, workload was measured using the NASA-TLX test.

Experimental results and conclusions

Results from the experiment (figure 3) suggest that the average time to give and answer and the overall workload were lower in the non-speech sound condition than in the speech condition, with similarly low error rates in both conditions. These results are similar to those reported by Kildal & Brewster [2] for blindfolded participants. However, no statistical significance can be calculated due to the small population used.

Analysis of cursor trace logs shows that, while participants commonly explored the complete data set before giving an answer in the sound condition, only a subset of the data set was often explored using speech. The reason for this is the slow, serial nature of the speech holding the users up. This suggests that the low error rate in the speech condition may be partially due to users guessing answers. A further study with more users is being undertaken and will investigate this.

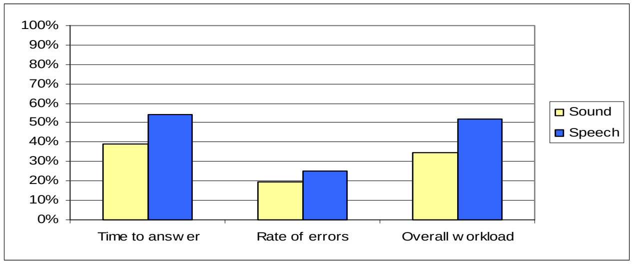

figure 3. Summary of mean scores

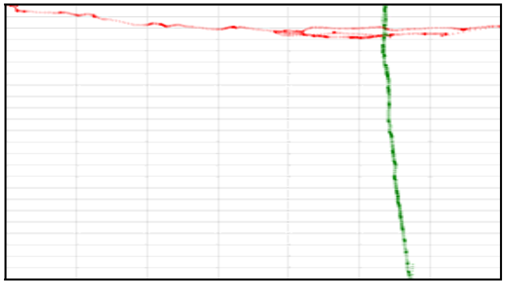

figure 4. Exploration trace showing a good exploratory strategy in the sound condition. The horizontal traces were done in columns mode (shown in red) and the vertical traces in rows mode (green)

Analysis of the logs also shows which interaction techniques were more successful in the non-speech sound condition. Those users who read details-on-demand less often obtained better timings and lower error rates than those who used this functionality more frequently, suggesting that the additional speech information is more an obstacle than an aid to gather overview information. Finally, users who developed a strategy of systematically browsing in rows mode for vertical discrimination and in columns mode for horizontal discrimination (as seen in figure 4) performed more efficiently. Several non-speech participants were able to browse all the data and answer correctly in well below 20 seconds, showing that this could be a very natural technique.

We propose this technique to obtain quick overviews of tabular numerical data, with low rates of errors and low overall workload. Results from the experiment showed that this technique could deliver higher scores than speech in time to answer overview questions and correctness of the answers, requiring lower subjective workload from the users

Acknowledgements
This research is supporter by EPSRC grant GR/S86150/01

References
[1] Card, S.K., Mackinlay, J.D. and Shneiderman, B. Information Visualization. In Card, S.K., *et al.* eds. *Readings in Information Visualization. Using Vision to Think*, Morgan Kaufmann Publishers, Inc., 1999, 1-34.
[2] Kildal, J. and Brewster, S.A., Explore the Matrix: Browsing Numerical Data Tables Using Sound. In Proc. ICAD2005, (2005).
[3] Shneiderman, B., The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In Proc. IEEE Symposium on Visual Languages, (1996), IEEE Comput. Soc. Press, 336-343.
[4] Zhao, H., Plaisant, C., Shneiderman, B. and Duraiswami, R., Sonification of Geo-Referenced Data for Auditory Information Seeking: Design Principle and Pilot Study. In Proc. International Conference on Auditory Display (ICAD), (2004).