

EXPLORE THE MATRIX: BROWSING NUMERICAL DATA TABLES USING SOUND

Johan Kildal, Stephen A. Brewster

Glasgow Interactive Systems Group
Department of Computing Science
University of Glasgow, Glasgow, G12 8QQ, UK
{johank; stephen} @dcs.gla.ac.uk

ABSTRACT

When first approaching a two-dimensional (2D) data table, a user often wants to get a quick overview of the data before analysing them in more detail. Blind and visually impaired people cannot do this task at all using visual displays. Furthermore, speech synthesisers do not help sufficiently with browsing and pattern identification. The approach taken in this research work is to use other senses available to the users to provide efficient ways of exploring numerical data tables. Sighted users having to analyse complex data sets could also benefit from multimodally-presented data arrangements

The first steps have been taken by using audio to access information in tables. A sonification of tabular data has been designed, as well as a novel way of accessing the sonified data. Each row or column is sonified as a single piece of information and they are accessed sequentially, reducing the overall complexity of the 2D data structure. Experimental results show 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.

1. INTRODUCTION

Tables are foundational and very common data structures that can accommodate any set of raw data. Two stages can be distinguished in the process of a person exploring a 2D numeric data table:

1. *Browsing*: a table is browsed to obtain an overview of the data contained within it, in order to find features such as trends, patterns or extreme values.
2. *Detailed analysis*: data is analysed in detail, concentrating on the features of interest that were discovered in the browsing stage.

The research reported in this paper concentrates initially on the stage of browsing tabular data and it is aimed to develop tools and methods for blind and visually impaired users to overcome the difficulties they find in the completion of this task. The approach chosen is to use properties of audio to convey this information. Tabular data are represented by appropriate mappings of data features to auditory properties, employing methods similar to those of visual representations, in order to “amplify cognition” and “reveal hidden patterns” in the data [1] as quickly as possible, almost ‘at a glance’. This paper reports the first steps taken in this direction.

1.1 The Process of Browsing a Numerical Data Table

Browsing a set of data is a complex process that involves several functional components [2] which may occur in any order and number of iterations during the browsing process. In terms of table browsing, these functional components are:

- *Orientation*: answers the question ‘where am I?’, and implies knowledge of the structure of the table;
- *Comparison*: occurs throughout the entire browsing process. Values in the table are compared to each other;
- *Place marking*: consists of defining “mental landmarks” with potentially interesting pieces of information for possible later re-consideration;
- *Identification*: confirmation or rejection of the mental landmarks identified during place marking, e.g. ‘Am I still interested in this?’
- *Transition*: movement from one value to another, in anticipation of a goal. It can also be movement away after identification and rejection, or after success or exhaustion of information;
- *Resolution of anomalies*: people tend to give structure to the data as they browse, even if they do not need to, ignoring odd values that would spoil the structure that the user was expecting to find. It can be seen as a consequence of the wish to find the target.

1.2 Current tools for blind and visually impaired users

Blind and visually impaired users commonly employ Braille representations of data tables or spoken output through screen reader software to explore tabular data. However, these tools provide little help for the task of browsing data tables. Work by Stockman [3] adds sonification functionality to a popular screen reader for better exploration and analysis of spreadsheets.

One of the main limitations of these existing tools is that they are mentally demanding. The user’s short term memory becomes saturated before the browsing task can be completed, as the number of units of information a person can maintain in the short-term memory at any one time is limited to between 5 and 9 [4]. As a result, navigational moves produce an overload of speech feedback (or Braille information) and there is a lack of the overall picture of the data structure [5].

Another limitation of these tools is that they are time consuming. Preparing Braille versions of data tables takes a long time (and may not be possible for the blind person to do without assistance). Having values read out in speech or reading rows of the table in a refreshable Braille display is slow and does not convey overview information.

1.3 Data representation

Graphical visualisation techniques are employed by sighted users to obtain information of useful features embedded in data sets. Blind and visually impaired users need enhanced or alternative representations through other sensory channels to be able to extract information, and this information can potentially be represented in sound, similar to how it is represented in visual graphs [1].

2. SONIFICATION OF NUMERICAL DATA TABLES

2.1. Sonification Strategy

Unlike in the sonification strategy proposed by Ramloll *et al.* [5] where a table was navigated in sound cell by cell, in the sonification strategy chosen for this research all cells in a row are grouped in a single (more complex) unit of sonification, so that all the information in that row is sonified at one time. Thus, the sonified table becomes an array of sonified rows, where the number of units of information is much reduced. As a consequence, browsing speed is a lot faster, producing a musical correlation of sounds that facilitate the identification of patterns and shapes, e.g. it is easy to hear if one cell stands out from the rest in a quick succession of sounds. An additional view of the table can be taken by sonifying each one of the columns one after another (see Figure 1).

88	88	75	38	31	63	45	73	30	71
44	108	48	47	57	89	70	95	54	24
69	62	49	64	23	21	41	61	37	84
76	23	54	33	52	31	71	64	35	66
54	27	71	48	76	55	89	97	65	33
76	58	21	62	71	35	64	22	45	53
26	27	99	36	99	54	81	69	28	74
57	47	62	48	96	79	86	83	25	23
32	77	58	95	53	55	66	91	89	65
34	52	31	37	68	73	54	62	44	69

Fig. 1. Browsing the table row by row (left to right) or column by column (top to bottom)

The values in a row or column are not sonified simultaneously, which would produce a generally dissonant chord of high musical complexity. Instead, values in one row (or column) are sonified in very rapid succession, from left to right (or from top to bottom). This approach takes advantage of the high temporal resolution of sound, a difference of 20ms being enough to be able to tell which of two sounds is heard first [6].

For a good compromise between browsing speed and clarity, the duration of the sound corresponding to one cell was chosen to be 60ms, avoiding overlapping between sounds of neighbouring cells. The sonification is paused for a time gap of 200ms in the transition from one row or column to the next, for a better separation between chunks of information.

2.2. Value Mapping

Like in the work by Ramloll *et al.* [5], values in the table are mapped to pitch, with frequencies in the range 27.5Hz – 4.786kHz (21-108 MIDI values). Higher values are represented by higher pitches and *vice versa*, which is accepted as natural and intuitive [7]. The sounds are generated using MIDI, and the

instrument chosen is the piano, because of its rapid attack and wide pitch range.

2.3. Position Mapping

When browsing the table row by row, sounds in each row are panned in stereo from left to right equidistant from each other, using the same method as Ramloll *et al.* [5]. Following the same metaphor (mapping the values in the audio space around the user according to how they are located in the table), when listening to the table column by column the sounds are panned so that all the values of the leftmost column are played on the left channel, all the values of the rightmost column on the right channel and the rest of the columns at regular distances between left and right.

2.4. Navigation Control

The user can choose the direction of browsing, either row by row or column by column. When the sonification is started, the whole table is played in sound, one row at a time (top to bottom) or one column at a time (left to right), until the end is reached or the user stops the process. Stopping and starting again reinitiates the sonification from the beginning of the table, which provides a quick mechanism for restarting the browsing. At any time, the user can shift the direction of the browsing from rows to columns and *vice versa*, always starting the sonification from the beginning of the table, which is important to provide a clear reference position in the table.

2.5. Dimensional Reduction

As represented in Figure 2, this sonification interface can be seen as a way of reducing a two-dimensional structure (the table) into two complementary one-dimensional structures (arrays of rows or columns).

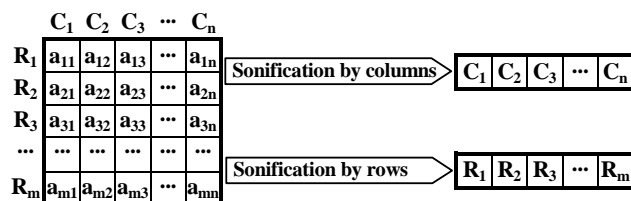


Fig. 2. A two-dimensional table becomes a one-dimensional structure after sonification. (R_x = row in position x starting from top; C_y = column in the position y counting from the left; a_{xy} = value in the cell in row x and column y)

3. EXPERIMENTAL EVALUATION

An evaluation experiment was designed and conducted to compare the performance of users carrying out table browsing tasks utilising the table sonification interface described above (the test condition) and a custom-designed speech-synthesis-based interface with specific functionality for browsing 2D numerical data tables that was implemented for this experiment (the control condition). All the implementations were done in Java.

3.1. Formulation of Hypotheses

- Hypothesis I. Less time will be needed to obtain an overview of the data in the table in the non-speech sound condition than in the speech condition.
- Hypothesis II. There will be no decrease in the accuracy of the overview obtained by browsing tables in the non-speech sound condition compared to the speech condition.
- Hypothesis III. There will be no increase in the overall subjective workload (NASA Task Load Index, TLX [8]) in the non-speech sound condition compared to the speech condition.

3.2. Implementation of a Speech Browser

Current speech browsing tools give little support for table browsing, other than just reading the values in the cells sequentially. Therefore, a more sophisticated interface was designed which would make a good comparison with the non-speech sound based browser.

In our speech browser, the values in the cells are read by a female voice and any other information concerning orientation, position or navigation aids is spoken out by a male voice. This difference in timbre makes it easier for participants to segregate the different types of information presented.

The interface offers two navigation modes: automatic and manual. In the automatic mode, the synthesiser reads out all the values in the table from begin to end, by rows or by columns. Each row or column is announced by the male voice. In the manual mode, the user can navigate cell by cell in any direction using the arrow keys. To speed up the navigation, the user can keep navigating to neighbouring cells, truncating the speech that was being synthesised in the previous cell. The male voice warns if the user is trying to leave the limits of the table. At any point during the navigation, the user can ask the coordinates of the current active cell, or can return to the initial cell at the top left corner of the table. In any of the modes the user can modify the speech rate to go as fast as it is comfortable. Shifting between manual and automatic modes does not change the cell in focus, helping the user to avoid becoming lost by jumping to another part of the table.

3.3. The Experiment

A two-condition within-subjects repeated-measures experiment was designed to test the hypotheses formulated above. Sixteen sighted participants took part in the experiment. They carried out table browsing tasks in 10 x 10 cell tables using each one of the browser interfaces. After browsing each table, participants answered a question that could be about location in the table or about pattern identification.

Location in the Table. The participant had to select the quadrant in the table (see Figure 3) containing the highest value, the lowest value or a value that was noticeably different from all the rest of the values in the table.

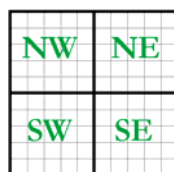


Fig. 3. Quadrants in a 10 x 10 table

Pattern Identification. The participant had to identify the general pattern of the data in the table, choosing between five possible options (see Figure 4 for examples of the types of tables used):

- Same value in the whole table;
- Values are highest (or lowest) in the centre of the table and decrease (or increase) towards the edges;
- Values in the table increase or decrease from left to right (constant values in each column);
- Values in the table increase or decrease from top to bottom (constant values in each row);
- None of the above.

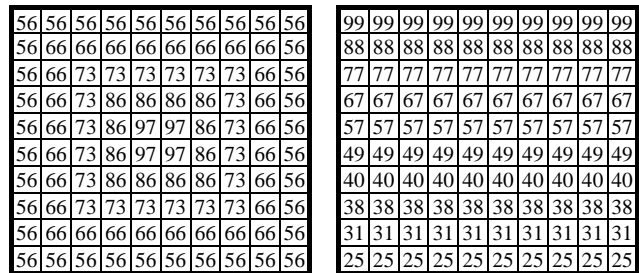


Fig. 4. Examples of some tables used in the pattern recognition task in the experiment. Left: values decrease from centre to edges Right: values decrease from top to bottom.

Each participant browsed 20 tables, 10 in each condition. Both the order in which interfaces were evaluated and the tables that were used in each condition were counterbalanced, as shown in Table 1.

	First Condition	Second Condition
Group-1	speech (table sets L1 & P1)	non-speech sound (table sets L2 & P2)
Group-2	non-speech sound (table sets L2 & P2)	speech (table sets L1 & P1)
Group-3	Speech (table sets L2 & P2)	non-speech sound (table sets L1 & P1)
Group-4	non-speech sound (table sets L1 & P1)	speech (table sets L2 & P2)

Table 1. There were four participants in each group. L1 and L2 were sets of five tables with a question about location in the table. P1 and P2 were sets of five tables with a question on pattern identification.

We used a forced-choice paradigm in the experiment, i.e. the participant saw the question to be answered and the possible answers, one of which had to be selected after browsing the table in search for the answer. Browsing time was measured, allowing up to 2min. for each table.

Before the experiment, each participant received training browsing four tables with each interface and answering questions. Data in the tables was optionally visible during training (since all the subjects were sighted). At the end of the experiment, participants completed NASA TLX scales [8] indicating the subjective workload experienced.

4. RESULTS

The arithmetic mean of the time in seconds spent browsing each table before giving an answer was, for almost all the participants, higher in the speech condition (M=68, SD=4.18)

than in the non-speech condition ($M=25$, $SD=1.83$). A two-tailed t-test (paired two sample for means) shows that this difference is significant ($T_{15}=698$, $p<0.001$). The rate of errors made in the speech condition (18.13%) and in the non-speech sound condition (25%) were not significantly different ($T_{15}=1.84$, $p=0.085$). Finally, the average overall workload in the speech condition ($M=7.36$, $SD=0.26$) and in the non-speech sound condition ($M=8.39$, $SD=0.25$) showed no significant difference either ($T_{15}=1.10$, $p=0.29$). The results of these tests are in accordance with the hypotheses formulated in 3.1.

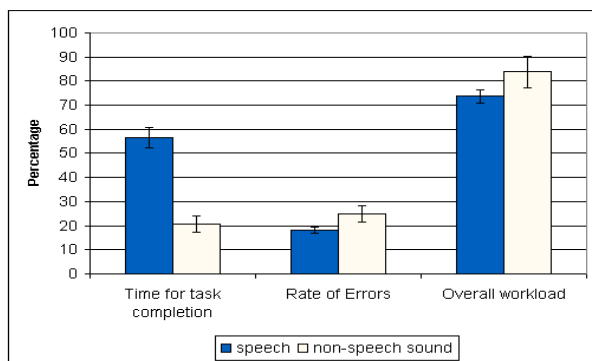


Fig. 5. Experiment results summary, showing standard deviations

5. DISCUSSION

The results of the experiment suggest that less time is needed to get an overview of the numerical data in a table using the non-speech sound based browser than using the speech based browser. The results also suggest that there are no significant variations between both conditions in the accuracy of the overview obtained or in the subjective workload experienced.

It can be expected that the difference in browsing speed between both conditions would be bigger if the values in the tables were of a higher complexity (e.g. high numeric values with several decimal positions) than the values used for the experiment (integers between 21 and 108 with no decimal positions). The sonification would be equally fast in both cases whereas having those values read in speech would take longer and would saturate the working memory earlier.

One limitation of this sonification technique is that the range of midi sounds that can be used is limited to 88 musical tones (Brown et al. [9] suggest that out of the maximum MIDI range of 0-127, the highest and lowest values should not be used in order to avoid problems with some sounds being indistinguishable). For this reason, the higher the range of numeric values in the table the lower the resolution of the sonification, possibly representing different values with the same midi sound. As a consequence, minor variations in the values will not be reflected in the sonification. This is not necessarily a problem as it will mask small variations in the data and further enhance any outlying values. Browsing data sets with a very wide range of values is an area where further research is needed but we have shown that for smaller ranges the technique works well.

6. CONCLUSIONS AND FUTURE WORK

The promising results obtained in the experiment encourage further development of these techniques and testing them with

visually impaired users. Other properties of sound such as timbre or rhythm could be used to map additional information features in the data, and appropriate exploratory procedures can be developed to combine the various mappings in a coherent way. The use of 3D sound might help obtain better information about the position of values in tables with two or even more dimensions, by mapping them to locations in the space around the listener. Additional browsing functionality such as more control over the navigation or place-marking could be used to develop and perfect new exploratory procedures. Further functionality can also provide increased resolution of the sonification locally for tables with wide ranges of values.

The results presented in this paper show that mapping numeric values in a 2D-table to non-speech sounds and presenting them grouped in rows or columns is an efficient method to allow users to browse tabular numerical data non-visually in order to obtain a quick overview – tasks which are particularly difficult using current technologies – and a basis is provided for future research in this direction.

7. REFERENCES

- [1] S. K. Card, J. D. Mackinlay, and B. Shneiderman, "Information Visualization," in *Readings in Information Visualization. Using Vision to Think, The Morgan Kaufmann Series in Interactive Technologies*, S. K. Card, J. D. Mackinlay, and B. Shneiderman, Eds.: Morgan Kaufmann Publishers, Inc., 1999, pp. 1-34.
- [2] B. H. Kwasnik, "A Descriptive Study of the Functional Components of Browsing," in *Proceedings of the IFIP TC2/WG2.7 Working Conference on Engineering for Human-Computer Interaction: North-Holland*, 1992, pp. 191-203.
- [3] T. Stockman, "The Design and Evaluation of Auditory Access to Spreadsheets," presented at Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia, 2004.
- [4] G. A. Miller, "The Magical Number Seven Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *The Psychological Review*, vol. 63, pp. 81-97, 1956.
- [5] R. Ramloll, S. Brewster, W. Yu, and B. Riedel, "Using non-speech sounds to improve access to 2D tabular numerical information for visually impaired users," presented at People and Computers XV - Interaction without Frontiers. Joint Proceedings of HCI 2001 and IHM 2001, 10-14 Sept. 2001, Lille, France, 2001.
- [6] J. Pierce, "Hearing in Time and Space," in *Music, Cognition and Computerized Sound*, P. R. Cook, Ed.: MIT Press, 2001, pp. 89-103.
- [7] S. Roffler and R. Butler, "Factors that influence the localization of sound in the vertical plane," *Journal of the Acoustical Society of America*, vol. 43, pp. 1260-1266, 1967.
- [8] S. Hart and L. Staveland, "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research," in *Human Mental Workload*, P. Hancock and N. Meshkati, Eds.: Elsevier Science Publishers B.V., 1988, pp. 139-183.
- [9] L. M. Brown, S. A. Brewster, R. Ramloll, M. Burton, and B. Riedel, "Design Guidelines for Audio Presentation of Graphs and Tables.," presented at Proceedings of ICAD 2003 Workshop on Auditory Displays In Assistive Technologies, Boston, MA, 2003.