J Multimodal User Interfaces (2012) 5:97–109
DOI 10.1007/s12193-011-0075-3

ORIGINAL PAPER

Navigation of interactive sonifications and visualisations of time-series data using multi-touch computing

Sam Ferguson · Kirsty Beilharz · Claudia A. Calò

Received: 3 February 2011 / Accepted: 3 November 2011 / Published online: 23 November 2011

Abstract This paper discusses interaction design for interactive sonification and visualisation of data in multi-touch contexts. Interaction design for data analysis is becoming increasingly important as data becomes more openly available. We discuss how navigation issues such as zooming, selection, arrangement and playback of data relate to both the auditory and visual modality in different ways, and how they may be linked through the modality of touch and gestural interaction. For this purpose we introduce a user interface for exploring and interacting with representations of time-series data simultaneously in both the visual and auditory modalities.

Keywords Interaction · Sonification · Tabletop · Data analysis

1 Introduction

Multi-touch technologies provide an opportunity to develop new approaches to communication of statistics and data. Scientific endeavours rely on data to be able to draw conclusions, but the communication of data is often limited to simplified graphics and appendices in research papers, and sometimes to data that is made available on academic, government or similar websites. This situation is changing rapidly, however, with the adoption of emerging technologies, and the number of organisations that are seeking to allow open access to data are increasing. To keep pace with this, new technologies are rapidly being developed to visualise and sonify this data, such as the field of information visualisation and data sonification. Parallel to these developments are the rise of new ways of interfacing with computing, through multi-touch, gestural, tangible and tabletop interactions. The current research discusses interaction, and develops a taxonomy of methods for representing, interacting with and discussing time-series data.

Data is the basis for *some* of our decisions. It is still one of our best ways of understanding the world around us, and forms the central part of the scientific method. Yet many people believe that the analysis of data is a technical job, or a province of researchers only—that understanding data is either difficult, or mundane, or that data is inherently biased and corrupted, or even that data cannot be understood without an understanding of statistics. These opinions are more likely to be a result of tools, logistics and interfaces than of actual difficulty with understanding data. Terminology and data format manipulation are two essential precursors for data analysis, but they do not add to our understanding, and usually form an insurmountable impediment for those who may wish to investigate any data for themselves. The result is that data is only investigated by those who usually have a role in making decisions (such as bureaucrats or officials), their subordinates, or by researchers—and is done in a way that is not transparent and cannot easily be checked.

New technologies generally play a role in increasing the amount and ease of information transfer (the world wide web, the telegraph, radio, and the printing press are standout examples). They can also have a role in the ease of information understanding and the way in which data are transferred through representations. This research discusses interaction methods for data analysis, and we believe that a change in the method of data transfer has the potential to engage a whole new class of users in the interpretation, the

S. Ferguson (✉) · K. Beilharz · C.A. Calò
Faculty of Design, Architecture and Building, University of Technology, Sydney, Australia
e-mail: samuel.ferguson@unsw.edu.au

auditing, and the meaning-making from data. The ability to investigate different sources of data also has the possibility to change the way in which people obtain information and think about it, moving away from the reliance on various media, and towards the possibility of personal and deep interaction with data. One way of engaging users is by making the experience of data more interactive and sensory.

1.1 Sonification background

Sonification has been defined as 'the use of non-speech audio to convey information' [29]. Recently, Hermann has expanded this definition, and has defined sonification in a more systematic manner, as a sound that: reflects objective properties or relations in the input data, has a systematic and reproducible transformation to sound, and can be used with different input data [21].

Sonification has been well-researched over the last 15 years, but it remains relatively complex to use it to investigate data. This is mainly due to the difficulty of building sonifications, the lack of an agreed set of patterns for their design [1, 5, 30], and the logistics of transmitting sonifications compared with visual representations. However, the proliferation of interactive systems has opened up new possibilities for sonification [26], and many difficulties associated with the time-bound nature of sonification can be overcome through effective interaction design. Furthermore, multi-modal displays can leverage the benefits of both the auditory and visual modalities in a complementary fashion—sonification has many benefits that are due to perceptual strengths of the auditory modality, and do not exist in visual representations.

Practical data representations, however, seldom integrate sonic elements to either complement or replace visual presentations. This is not surprising, as for the day-to-day task of actually discussing or using data to make decisions visual representations have enormous advantages over auditory or haptic interfaces; they can be compared and discussed in an amazingly flexible manner, and can be embedded in documents easily. By comparison, an auditory representation is unfamiliar, and at least requires presentation using a computer with relatively specialised software and sound reproduction equipment. This immediately limits the way in which the sonification can be experienced to a subset of the formats that a visual representation can be presented—for instance, their use within a document or book (whether computer-based or not), in a lecture theatre, in a busy office, in a cafe, etc. require preparation and possibly even specific designs, while a single graphical form will generally suffice for most of these situations. For sonifications to contribute to mainstream data analysis, it is essential that sonifications add something essential to visualisations, and are not pitted directly against a more practical alternative. A synergistic cooperation across modalities is the motivation for this approach, taking into account the multi-modality of analysing data through a combination of visual and auditory display, as well as the multi-modality of representation. This is in contrast to an approach that views sonification as an alternative to visualisation, and is aimed at leveraging the strengths of each modality.

1.2 Tabletop computing

Tabletop computing has been researched extensively in recent times—studies have focused on typical computer actions, such as file system access methods [12] and personal information management [13], but have also extended to topics such as concept mapping [14]. Collins et al. [12] outlines the following design considerations especially for collaborative touch-table design (software interface, interaction and surface):

- context of use, communal location
- impact of size and orientation for small groups
- reach
- clutter
- limited input, no tactile feedback
- resilient to accidental touch and gestural ambiguity.

Research into the use of multi-touch interfaces for sonification is still in its infancy, but is expanding at a rapid rate. Tünnermann et al. have investigated sonification of multi-dimensional datasets, [39], Bovermann et al. [10] have introduced the idea of tangible data scanning as a model of interacting with data. Hermann et al. have discussed the use of tangible objects on a tabletop as a method for data exploration [22]. Much sonification research has pointed out the importance of interaction to sonification, for instance in the sonification of sound [15], for monitoring complex processes [23], and for fields as diverse as juggling [9] and recreational sports [6]. There has also been some investigation into possible methods for developing new musical interfaces that use multi-touch interfaces by Hochenbaum et al. [25] among others.

Given pervasive personal *multi-touch* devices (e.g. iPhone, iPad, Android tablet computing) why focus on the *table,* not merely an application for surfaces? Primarily, the motivation lies in the collaborative potential of table-scale surfaces, especially for meetings, discussions and interaction. Further, the screen real estate afforded by a high resolution table surface for comparative and spatial manipulations of the data representation facilitates much richer interaction and more detailed analytical tasks. However, we acknowledge that the new possibilities afforded by rapidly improving sensors and systems, such as the *Kinect*, and Oblong Industries *G-speak*,[1] threaten to again redefine the methods

[1] http://oblong.com.

for interaction with computers. Portability matters, however, and tablet based technology looms large over tabletop interaction. It may be though that a tabletop computing perspective, transferred to a tablet technology retains a very different viewpoint to the more common situation, where desktop applications are refactored to use tablet user interfaces but retain many of the formats and conventions common to the desktop computer paradigm. A cursory inventory of the described interaction and display methods in the current prototype does not appear to conflict with the input and display methods that exist in tablet devices, and indeed even implementation similarities mean that possible conversion for tablet computers is likely to be relatively trivial.

1.3 Interactive data representation

Both recent advances in software tools and new data sources made freely available have had a significant effect on information visualisation practice. With the ability to create compelling data representations on commonly available computer hardware, an increasing number of artists and designers have developed interactive information visualisations as a communication medium in its own right [17]. Due to this proliferation of information representations, media, educational, governmental and advocacy organisations have embraced information visualisations for their ability to inform and engage its audience. Prime examples have been featured in various prestigious venues, ranging from the cover of scientific journal Nature [18], to renowned exhibitions in the New York Museum of Modern Art (MoMa) [3], to popular projects online (such as On's *They Rule*,[2] or Harris' *We Feel Fine*[3]). By articulating the common interests and goals of representation across academic fields, shared research problems can be addressed in a multi-perspective space and a broader cultural context [8, 27, 40].

Although these works have successfully conveyed meaningful insights to hundreds of thousands of people, there still exists a gap between 'scientific' and 'aesthetic' representations in their ability to recognise and learn from each other's capabilities. Warren Sack [34] recognises aesthetics of governance: the interpretation and articulation of meaning (instead of information) as a creative response to visual forms of contemporary art. Kosara [28] proposes criticism as a tool for identifying the differences between information representation and more artistic forms of visual communication. Pousman et al. [33] coined the term 'casual visualisation' as conveying an 'increased focus on activities that are less task driven, data sets that are personally meaningful, and built for a wider set of audiences'. Viegas and Wattenberg [42] use the term 'artistic data visualisation' for data depictions that embody a forceful point of view, thereby recognising the power of representation as a potential mass communication medium.

The current inadequacy of existing information visualisation techniques in serving the general public and policy makers is evident in the lack of engagement with open data at present. Examples of available yet under-explored data include that released by the Australian Bureau of Statistics, the U.S. established `data.gov` website, or the world bank's data catalog.[4] A designed solution is required because the information-consumers are essentially non-experts: policy-makers and the general public. Open data also holds significant potential to engage, motivate and persuade public opinion and permeate social networks and cooperation via interactive debate/discussion and feedback.

The information aesthetics [35, 41], interactivity, clarity and comprehensibility of a representation each contribute to its usefulness. The importance of understanding the role of information aesthetics in data representation, and in particular 'how insights and aesthetics interact [to] sustain insightful and visually appealing information visualisation' has been listed as one of the 'Top 10 Unsolved Information Visualisation Problems' [11] and was one of the 'Top Research Goals for 2010' [2]. This interest is also demonstrated by some of the initiatives at the world's most prestigious academic conferences, such as the 'Information Aesthetics Showcase' at the ACM SIGGRAPH'09, the theme of balancing '...art and science, design and research' at CHI'08, and the annual 'Infovis Art Exhibit' at the IEEE Infovis'07–09.

Democratising data enables ordinary people to engage with and take responsibility for the information that they create and communicate, through social dialogue, interaction and as a basis for decisions grounded in genuine knowledge. Once data is in the public domain, it can be repackaged and improved in unpredictable ways [19]; 'Citizens become more active and creative producers of information and recombined data' [31].

1.4 Interactive sonification as multi-modal interaction

The convergence of interactive sonification with information visualisation and touch modalities allows us to investigate the inter-operative, reinforcing and independent role of sound in interactive data representation. Studies show that manipulating the way a graph is drawn influences viewers' ability to extract information from it [45, 46]. Multimodal design practice, and consequently sound, has the potential to complement the visualisation domain, reinforcing the effectiveness and the immediacy of the information.

[2] http://theyrule.org/.

[3] http://www.wefeelfine.org/.

[4] http://data.worldbank.org/.

Most existing methodologies for sonification focus neither on the aspects of aesthetics nor on the musicality of the sounding outcome, rendering representations that are difficult for non-experts to understand. Auditory representations can be used to perform trend analysis, point estimation, pattern detection, and point comparison [43], however auditory representations may not always be the optimal method suited to these generic tasks. The strengths of auditory display, however—in highlighting subtle changes in values, illuminating gradual changes, presenting several data-streams concurrently, emphasising anomalies and outliers—have the potential to complement visualisation methods to achieve intuitive, rapidly understood depictions. Past work by the authors [7] focused on aesthetic sonification and user-centred interface customisation: giving the interacting user interactive control over the musical setting of various elements present in the sonification (e.g. kind of scale, key, pitch selections, timbre, spatial audio streams, register and range, rhythmic accentuations, and tempo). While customisation is useful, a theoretical framework is still required that is broad enough to unify various representational formats, technologies and software platforms, and which is capable of incorporating combined auditory and visual modalities into a hierarchical understanding of a data representation.

2 Data interaction framework

The interaction purposes that may guide a user are varied, but they are generally based around *story-finding* or the ability to make meaning from a dataset. Many authors discuss data analysis as a process of finding the *story* within the dataset.

Stories may consist of various types of meaning, such as:

- comparisons between the consequences of taking different courses of action;
- predictions for the future;
- understand complexity and develop new understandings;
- finding special cases which are distinct from the ordinary;
- understand a recorded version of the past;
- seeing what aspects of the world are related or irrelevant to each other;
- determining which principles are useful and which are untrustworthy.

To extract any meaning from a dataset, it needs to be decoded and interpreted. This research investigates simple concepts at the basis of tabletop or multi-touch interaction and navigation with data representations, such as how it is possible to scan through data and explore data, and what can be changed about a representation and what must remain consistent.

Many different approaches to scanning through data exist in various fields—but they are best described by Shneidermann's information visualisation mantra, 'Overview first, zoom and filter, then details on demand' [36]. The problem is to maintain different levels of detail in one's mind concurrently, so that comparisons may be made at various levels. If a dataset is to be understood in these levels of detail then scanning through the different levels needs to be facilitated by the computer, rather than hindered by it. Developing a taxonomy of different interactions and their purposes is a precursor to choosing the best method for a given purpose.

As an example of an interaction, pointing at, or pressing on a data point, is an action that may be expected to make the data point 'give up' its information. But another possible response is that a 'playhead' could be attracted to the datapoint that was pointed at, or a selection could be made—there are actually many responses that may 'make sense' for one action. Similarly, pressing two points on a data representation may mean a selection of a data range has been made, or that a comparison is necessary, or that a zoom selection has been requested. Thus we can see that, despite the somewhat standardised and intuitive multi-touch gestures we may be currently familiar with (Pinch, Swipe, Scale etc.), there remains considerable ambiguity and orthogonality in the touch gestures available and the resultant, consequential affect on the interface. Below, we will categorise each of the interaction activity possibilities under four categories of interaction—arrangement, exploration, selection, and scale or zooming.

2.1 Arrangement

Sonification of data requires a mental model of the relationships between the data sonified, and other data that may not be directly within view, or being presented simultaneously. Using systematic visual arrangement of multiple data dimensions allows the sonification user to form and manipulate a strong mental model about how the data is structured. A sonification performed of a particular data dimension can then be related to another sonification of a dimension rapidly.

As the arrangement of representations is so important, some of the other common ideas about tabletop computing probably need to be de-prioritised. The first is allowing the resizing of graphics—if this is permitted non-systematically, it breaks the possibility of developing groupings and drawing comparisons in a fair manner. While there may be necessary situations where resizing is necessary and helpful, these need to be controlled to avoid a situation where different graphs give similar information in different ways. Orientation is another issue—many tabletop computing systems use orientation flexibility to maximise the multi-user collaborative possibilities of the systems, but this again breaks the design outlined above, although not to the same degree.

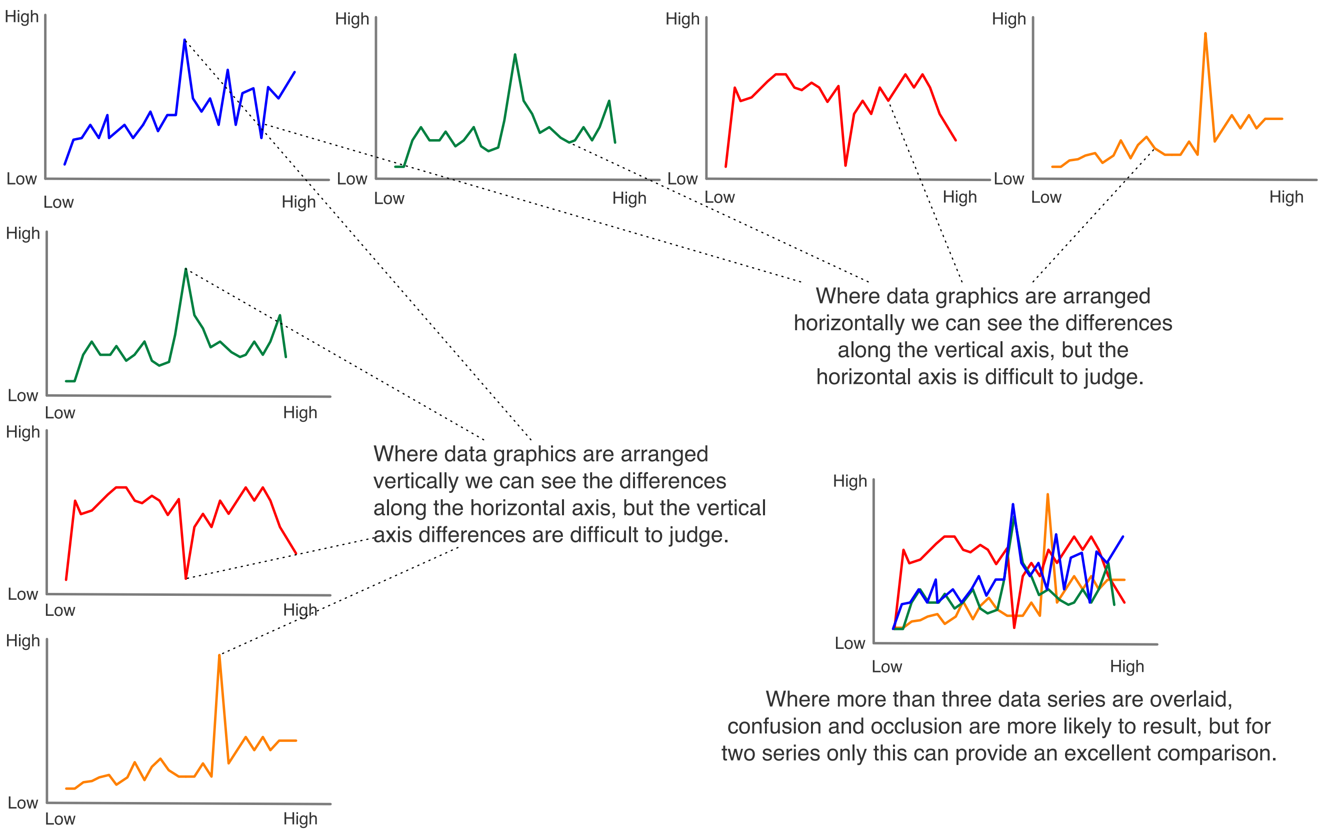

Fig. 1 Where spatial rearrangement of multiple graphs is possible, the different arrangements allow different information to be assessed

While this does restrict a user from customising the representation in this manner, it also frees the user to perform several other types of operations that would be difficult if this customisation were performed. For instance, an arrangement that is direct is to overlay two representations for comparison on the same axis. When the orientation and size of the graphics are constrained to a particular relationship, this is very easy to achieve. In turn, this overlaying can provide further possibilities for accessing and triggering comparative sonifications.

Furthermore, spatial relationships between multiple small graphs can quickly reveal patterns within the information depending on how the graphs are arranged. Lining up graphs is a familiar approach to looking for patterns—if graphs are lined up vertically, then the arrangement can show the difference that is occurring on the horizontal scale, and conversely horizontal arrangement can highlight the characteristics of the vertical scale data (see Fig. 1). In the described user interface prototype, each data dimension is presented as a separate visual representation on the data 'canvas', where the interaction with the data can take place. These graphics are fairly familiar in visual appearance, but it is their interactive nature that is the point of departure.

Immediacy of interaction is necessary for rapid understanding of data by novices—achieved by representing relationships between parameters both visually and auditorally, rather than asking users to consider statistical concepts and develop numerical analyses of the relationships described. A first step to providing this immediacy of interaction is to provide a layout for the various data dimensions.

Controlling sonifications is further enhanced by using multiple visual data graphics arranged horizontally or vertically, as the system can assume that they are related to each other, and therefore comparisons between them are expected by the user. Sonification controls can be linked and automated across the set of graphs. Furthermore, by setting up two or more arranged groups of graphs pre-set views can be elegantly passed to a data representation. The act of lining up a data graph with a set that have already been customised to a particular range can be used to propogate the settings of the axes and other parameters from the other graphs to the new graph. As an example, two sets of graphs may be set to look at the stock market for the years 1925–1935 and 1985–1995, to investigate the size and effect of the financial crises that occurred during these two periods on perhaps four countries. Aligning a further graph with either of these groups could be used to investigate this particular period more closely.

2.2 Exploration and playback

Given the interactive possibilities now available, it is worth reconsidering the idea of a sonification 'playback', and attempting an abstraction of the concept. Playback, in the current context, is the systematic presentation of a particular data dimension, along which it is often expected there may be a trend, or at least some interesting information. The current focus is on time-series data (which is played along the time dimension usually), but any other dimension may be used as a 'playback' dimension along which to order data. The process of playback allows the user to listen for trends that occur over time, instead of requiring the user to find them in a representation. Most visual representations do not include the concept of 'playback', but for sonifications this idea is usually essential as most (but not all) sonifications are presented as individual sounds ordered by time. Purely visual representations, on the other hand, do not usually require this mapping to time, as the user can visually scan through the graphic examining each data point in whatever manner they wish. In order to build a multi-modal representation that exploits the benefits of sonification, a concept of playback must be extended to exploit interaction to allow reorganisation for time-based presentation.

Alternative approaches to representing data by using parameter mapping sonification over time often involve an interactive interface to control the movement through the data dimension. With a visual graph, many questions can be answered from a single graph due to the ability for the eye to scan through the data and make cross-references as needed. The use of sonification requires the same scanning ability, but the ears cannot easily scan through the memory of a sound, especially if it contains a large amount of data, and therefore interactive sonification holds much promise in a multi-modal context.

In the user interface described here, we have approached the time-axis as a moveable (manually or automatic) playback head that can be controlled allowing for repeated listening, by 'scrubbing' or scrolling in a forward and backwards direction. This method, however, has some drawbacks when experienced and compared with typical sonifications. The conscious effort needed to steadily move a playhead through a set of data makes listening difficult, and the alternative—having the playhead follow behind the finger at a steady playback speed—results in a 'stickiness' that negates selecting different parts of the time-series rapidly. Therefore, at least two methods of playback are needed in an interface of this nature—playback of a selection of datapoints at a regular tone presentation rate, and one for simply selecting a single datapoint and having it sonified.

Furthermore, when the idea of sonification 'playback' is extended to comparisons between data dimensions, it becomes apparent that rapidity of playback is important. To get an idea of five data dimensions one may wish to choose one of them, and make a comparison between the remaining four one by one. If the sonifications generated by each of these comparisons were to last 15 seconds each, the practicality of this interface would be heavily limited by the short-term memory of the user. Rapid sonification speeds, or methods that create stationary sounds with representational qualities, are thought of as a solution to increasing the efficacy of this design.

2.3 Selection

Selection is the process of marking the boundaries of a subset of the data, from a larger dataset. For each selection, in data that only has one dimension (such as time) there are two boundary points, in 2-dimensional data there are four, and so on. In any case however, there are a set of boundaries to be selected, for a number of different possible reasons. Selections may designate an area of interest or concern, within a larger context, or they may be used to discard useless or irrelevant information.

Selections also exist *between* data dimensions. The way in which data dimensions are arranged and accessed in a statistical workflow determines much about the ease of making a comparison or looking for a pattern. Selection allows processes (whether statistical or representational) to be applied to appropriate and interesting data, and between different data dimensions.

2.4 Zoom and scale

Scales are extremely important in any data representation as they provide the context by which the data is understood, in terms of the parameter being described, and the method or units used to describe it. The level of detail, and conversely the amount of data viewable, are determined by the scales. Scale size and transformation (e.g. logarithmic/linear) can have a transformative effect on the ability of the representation to show information (often through highlighting regularity along a particular scale).

Comparisons between representations require scales to be comparable in at least one of a number of possible ways. Either the scales should be identical, meaning that the quantities represented will be the same, or they should be similar in another way, so that although the quantities represented are not the same, the meaning is similar. For example, when comparing time-series graphs, the time intervals on the x-axis (or axis allocated to time) should scale at the same rate, so as to retain the relative relationship between graphs being compared. An example to clarify: comparing median income from various countries may be done by using current US dollars as a common factor, with scales set between \$0 and \$150000—but this may make it difficult to perceive any meaning about income levels from countries where the values may be 10 or more times smaller. Various strategies exist for setting appropriate scales, but the two main options are normalising the data scales to the data range, in order to see comparative *change* or *shape* between data dimensions, or standardising the scales to a shared range in order to compare the absolute *magnitude* of different data dimensions.

Authors such as Nees and Walker have discussed the possibility of including axis information as auditory context within a sonification [32], but this becomes somewhat unnecessary in the current context, as the user may employ the visual axes to assist them. Theoretically, this should provide less complexity in the sonification.

3 User interface design

In this section we will detail the user interface design. Time-series data is presented in the form of small interactive graphical representations. We are focused on a refactoring of the way in which graphs are presented when interaction is made possible, and when sonification can be used. When a representation is static, many elements that are included, such as gridlines, scales, captions, legends and labels have to be carefully designed, and sometimes compromises are made so they are not too distracting. With the benefit of interactivity, a representation may hide or show various elements depending on a user's preference, purpose or current activity. This also allows the prioritisation of the data over the other elements.

A reasonable assumption is that the data representation should be accessed and altered through a multitude of different multi-touch gestures in one mode. However, even a small number, when implemented, proved to be difficult for interaction purposes. As there appear to be two main categories of interactions which are very different (data playback/comparison and data arrangement/scaling), it seems that frustration is increased when the system misinterprets a gesture (for instance when it moves a representation rather than sonifying data). Furthermore, some of the simplest interactions (dragging, scaling) were easier to perform, and using them for only one purpose meant that other actions became difficult. Therefore, we used a modal design, and grouped various functions together, so that they did not annoy the user when an interaction was misinterpreted (see Fig. 2, with discussion below). The two modes used (selected using a toggle button in the upper left corner of the canvas) were:

1. Arrangement and movement of each representation, as well as zoom/scale alteration.
2. Exploration, sonification and comparison between different data representations.

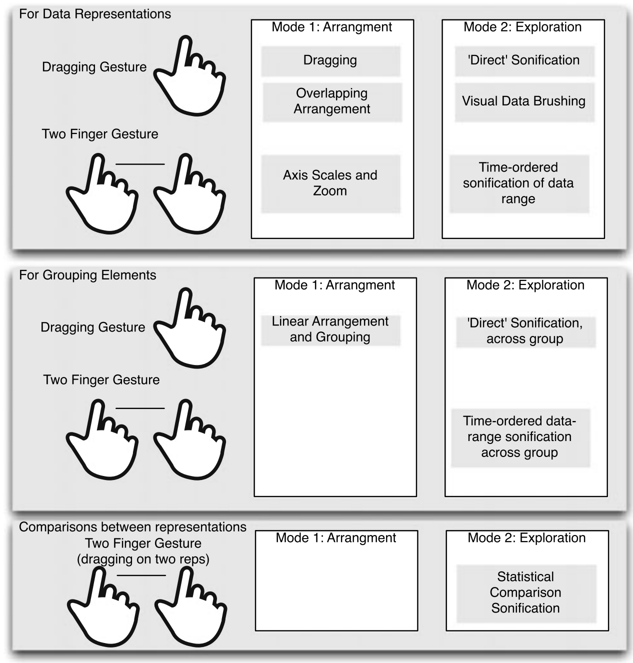

Fig. 2 A figurative classification of the modes and gestures in the described user interface

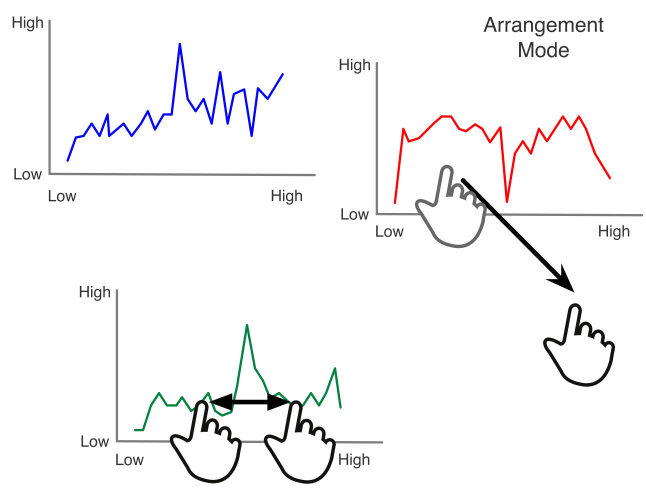

Fig. 3 Arrangement is the mode by which the various representations can be moved around the data canvas, and aligned into groups

3.1 Mode 1: Arrangement

In the Arrangement mode, for representations:

– Dragging moves the representations around the space;
– Two fingers scales the axes.

For grouping elements:

– Aligning the representation with the group with propogates the x (time) and y-scales.

Arrangement (Fig. 3) is the mode in which interactions control the positioning of the representations, the scales, and

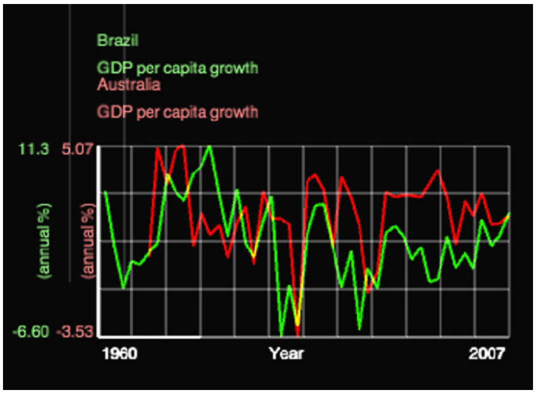

Fig. 4 When representations are overlapped, axis elements reorganise to aid in visual representation

zoom of each of the representations, and the arrangement of the representations into groups ordered vertically or horizontally. This is also the mode where the data source for the representation can be changed.

A single finger is used to move the representations around the canvas. This is quick and natural, and is less a metaphor than a direct manipulation. Multiple representations can be moved with multiple fingers or hands, or by multiple users simultaneously. Two fingers are used for timeline scaling, allowing a viewing range to be selected through using the zoom metaphor. Continuing this metaphor, the timeline range can also be panned within the timeline.

For arrangement, a natural approach to make a comparison is to overlay two representations (Fig. 4). As the system tracks the location of other representations in this mode, the overlay of the two graphs triggers the contextual labels and scales to be rearranged to facilitate this interaction. Also, as each representation has a different hue associated with it, and the y-axis labels as well as the central line-graph are drawn in this particular hue, so colour correspondence can be used to associate the lines with the appropriate scale labels.

The other interaction discussed is the arrangement of multiple graphs into a sequence. This is possible manually, graph by graph, but as the purpose of the grouping is for interaction, it is useful to have a visual grouping element so that interaction with the group can be directed through it. This grouping element is used to align the representations, but it can also be used to interact with the sonification of the data for each in the group of representations. This means that the interaction methods for single representations can remain consistent and unchanged when they are grouped together (see Fig. 5).

The grouping element can also used to propogate scale information (either the time scale or the y-axis information) across multiple graphs. Therefore, as soon as a graph is aligned to the group it can take on the new data scales,

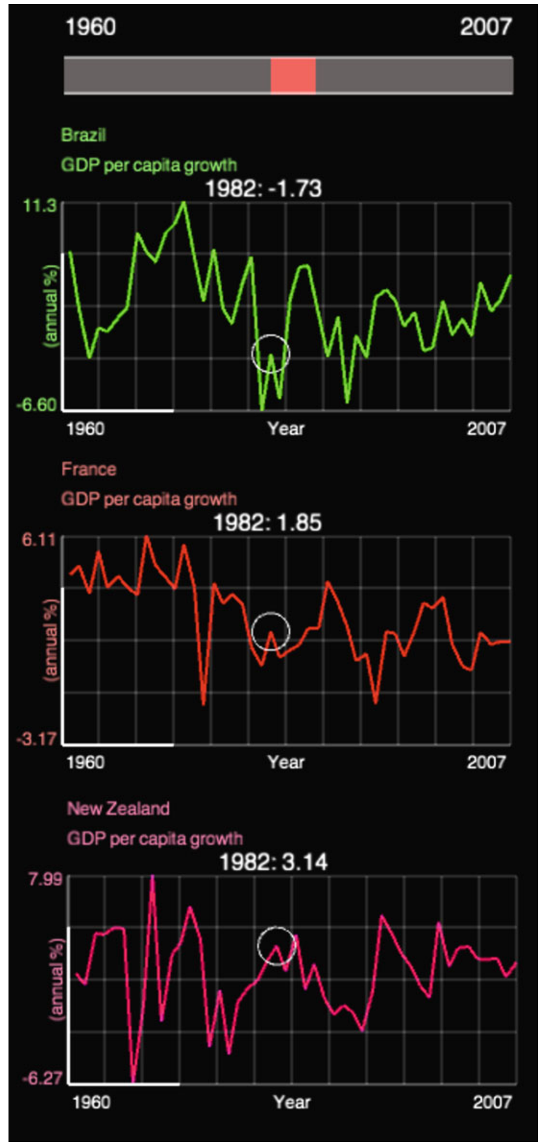

Fig. 5 When representations are arranged vertically, comparisons can be made between events related to the time axis. Additionally, sonifications of the entire group can be controlled using the grouping element (the *bar at the top*), which results in each representation being brushed

reverting to the previous scales when the representation is removed from the grouping element.

3.2 Mode 2: Exploration

In the Exploration mode, for representations:

- dragging sonifies the data and shows the information;
- a 2 finger gesture makes a selection and rapidly sonifies it.

For grouping elements:

- by dragging on the alignment vector all the aligned representations are played at the same time;

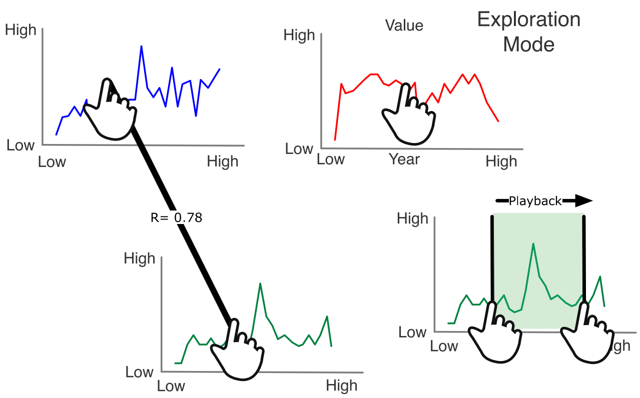

Fig. 6 Exploration is the mode by which the representations can be sonified and investigated

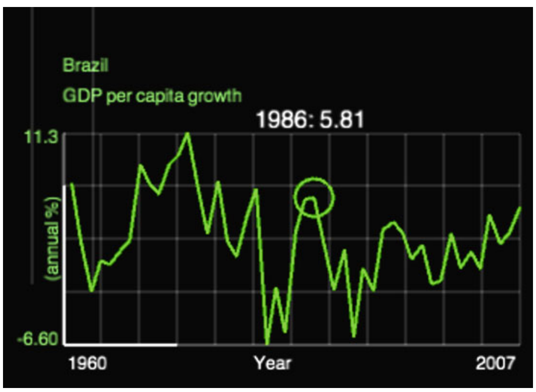

Fig. 7 Point estimation simultaneously performed in both the visual and auditory modes highlights the contributions of each mode

- two fingers on the alignment vector selects a time portion and plays that portion rapidly, but again for all the representations in the group.

To make comparisons:

- dragging on two representations makes a comparison between them;
- dragging on a representation and an alignment vector makes a comparison between each representation in the alignment vector and the original representation.

In the Exploration mode (Fig. 6), the data can be brushed (where particular portions of the graph line numerically displays datapoint values) and sonified with a single finger (Fig. 7). The data is shown numerically for each datapoint on the graphical representation, and simultaneously sonified as they are touched by the user. Where the representations have been organised into an aligned group, the single-finger brushing will still function as expected for single representations, while if the grouping element is brushed the group will all be sonified simultaneously. Furthermore, the use of two fingers to delineate a selection allows the playback of that portion of the time-series, which begins as soon as the two-finger selection gesture is finished.

Furthermore, after the data has been investigated and sonified graph by graph to gain orientation and a sense of the data, analysis may continue with comparisons made between different data representations. To do this, all that is necessary is to touch two different data representations simultaneously. Values calculated between the two are sonified, leaving the visual context undisturbed. Sonification methods for making these comparisons will be discussed below.

4 Sonification methods

We have mentioned already that other authors have discussed data sonification models for interactive systems. Tünnermann and Hermann have described a 'shockwave' method of data sonification playback [39], and Bovermann extended data exploration to physical space using *Tangible Data Scanning* [10]. These methods seem to focus on multi-dimensional data, such as the *iris* dataset [4], which often possess no obvious time dimension as does time-series data.

Multi-dimensional projections, like that envisaged by Bovermann and others, are one way of approaching these types of complexity, but another method is discussed by Wilkinson et al. [44] and Tufte [38], which is the use of multiple small graphics to represent the conditioning of a dataset. These authors show that through the use of multiple 'facets' or conditioned graphics, a data representation can avoid being a single, complex, representation. As discussed above, the use of such arrangements can also be a fertile manner in which to allow a mental model of data to influence decisions of how to listen to the data.

In many circumstances sonification research focusses on complex or novel methods of sonification. Instead of focusing directly on the transformations from data to sound, this research focuses on various sonification presentation formats and navigation solutions in relation to the current interactive multi-modal context. Therefore, a simple parameter mapping approach will suffice initially. We have chosen parameter mapping to the pitch of frequency modulated tones, with small variations in the onset time and tone colour used to provide a measure of distinction between the different data representations. The pitch mapping is a mapping from the visible on-screen y-axis range to a two-octave pitch range quantised to midi note numbers and arranged along a log-scale. Setting the scales on the representation adjusts the mapping input range, but the output pitch range is generally fixed (at least currently).

In terms of the mapping to time, however, we have at least three main approaches: a direct mapping of interaction to sonification (yielding a flexible time mapping), a sonification of a portion of the data (a stricter time-mapping), and a sound cluster as a representation of the data midpoint and spread (time as an axis is ignored).

'Direct' sonification of the values touched or 'brushed' is a primary presentation method discussed, as this is the method that may seem the most obvious when considering a combined visual and auditory representation. The typical response when presented with a graphic on a multi-touch display is to touch it. It is actually, however, an uncommon sonification format in other current sonification software frameworks, as often they don't incorporate visual graphics or interaction in this manner. This is analogous to point estimation in either a visual or auditory graph, but is performed with one's finger (Fig. 7).

Sonification of a portion of values is another essential method. While sonification brushing is highly interactive, and allows rapid response, it destroys an important advantage of representation in the auditory domain; the ear's ability to detect regularity in time. 'Direct' sonification uses an arbitrary time-mapping based on the interaction events. To create a time-ordered sonification we need to make a selection of range between two time points. Then a sonification using a regular time-mapping for that selection of data points can restore the use of the ear's ability to detect regularity.

The third sonification method, as an alternative to an ordered sonification, is to rapidly sonify randomly chosen values from the data dimension, yielding an averaging, density cloud type of sonification. If the sonification is carried out at a fast enough pace, a type of averaging occurs that produces an impression of the midpoint and spread of the data. These 'clouds' can be quickly compared with each other, ameliorating the need to use short-term memory for comparisons. Of course, depending on the parameter being compared this technique may make assumptions about the equivalency of values from different time periods (it is better suited to situations where the values are independent of each other). A quantity that fluctuates around a mean, like a currency value for instance, is a good candidate for this type of sonification, so as to obtain an impression of its general range.

4.1 Comparison sonifications

Many statistical comparisons may be performed between two time-series, and our purpose in this research is to provide a template interaction method, to which various statistical methods may be applied. As an example, we will discuss a comparison based on Pearson correlation between two time-series. In a typical scenario, however, many more parameters may need to be controlled than will be discussed in this example, as statistical comparisons often entail significant numerical specificity (although using a one-time dialogue box may be all that is necessary to adjust a comparison's settings to adequate values).

Correlation between two time-series describes the degree that the variance in one time-series may correspond to the other. This does not, on its own, indicate a causal relationship between the two, but is usually reason for further investigation of a possible relationship. Correlation relationships are often expressed as values between -1 and 1, where numbers close to 1 are a high correlation, numbers close to 0 indicate little relationship, and numbers close to -1 indicate an opposite relationship.

Given that a comparison of this nature can be described, at least in a cursory manner, by a single value, a design for a sonification of this value is relatively straightforward. The numerical relationship $(-1, 0, 1)$ is obviously quite idiosyncratic to measures of this type. A sonification design for values in this range is as simple as playing a mid-range note, and then a note up to one octave above (for a 1) or below the note (for a -1). If the value to be sonified was 0.99, (a high correlation value), then the sonification would consist of a mid range note, and then a note that is almost an octave higher. The playback of two notes is almost instantaneous and pitch-mapping is easily perceptible, allowing rapid correlation comparisons to be carried out without reading the visual display for any significant period of time.

Many alternative sonifications of correlation may be proposed—for instance, one incorporating a representation of the significance (p-value) of the correlation value, or another that attempts to represent correlations in terms of the raw data rather than a numerical reduction. We have specifically focused not on the sonification designs themselves, but on the possibility for sonification comparisons of this type to be executed rapidly. This helps maintain a persistent mental model within the wider framework of data investigation and sonification.

5 Discussion and evaluation

This article has presented an interface for multi-touch tabletop interaction and sonification of data. This design aims to simplify the presentation and use of time-series data, by allowing rapid selection, arrangement, exploration and comparison of different time-series data. We especially utilise multi-modality to foreground different aspects of the data and multi-touch interaction to perform comparative and analytical tasks while listening to the data through sonification.

By heuristically evaluating the implications and possibilities of the interaction methods we seek to validate these navigation principles. The integration of data exploration methods in a single interface is the purpose of this approach—as opposed to a purpose-built representation for each dataset. The ability of a representation method to engage a user is important, as an unengaged user will find it difficult to comprehend the data the representation is attempting to convey.

5.1 Sonification in multi-touch environments

Sonification provides an important opportunity to develop new methods of prioritising the 'narrative' over static graphical representations. Especially for time-series data, the time-based method for representing the data reflects the narrative that exist within time-series data. The use of selection and 'playback' is a departure from the way in which data is usually explored by visual representations (through scanning and cross-referencing), and can provide a more systematic and narrative discussion of time-series data. The viewer is able to examine the sonification in their own time, enabling a more intimate, detailed and focused interaction than occurs in non-interactive sonification playback.

Sonification's difficulties, such as point estimation, memory, playback format etc. are greatly ameliorated by the use of an interactive multi-touch presentation method. Point estimation is evidence of the way that multi-modal interaction can enhance data understanding—sonification adds a narrative element, and a binding of the time axis to time in a way that strengthens the flexible, but arbitrary mapping that a visual representation creates. When a point on a line graphic is touched, the surrounding context is sonified as the user settles on a particular value—the user understands the value in relation to its data context from the sonification, but the visual graphical elements (both the axis elements and the data brushing) provide another type of context, by relating the data to the numeric decoding of the x- and y-axis.

The arrangement possibilities for the representations in the visual modality are not just used for visual investigation. They allow a mental model of the data (its groupings, contrasts and sizes) to be developed in an iterative manner before particular sonifications are played and investigated. Furthermore, when the data dimensions are systematically arranged they can be exploited for comparisons in flexible and user-centred manners. Given that much of the process of data analysis is finding and testing hypotheses in the data, allowing simultaneous access to many different perspectives on the data is an added efficiency.

Sonification can also be used to provide extra analysis tools that mean that users do not need to switch visualisation modes to undertake analysis. This flexibility allows for the user to have a clear idea of the basic representation and relations between data dimensions, but to also listen to abstractions of that data (such as perhaps an auditory histogram). This is convenient, as nothing needs to be 'changed back' to finish the process, maintaining the representation context while providing analysis results quickly.

5.2 New sonifications for multi-modality

What should comparisons between two time-series sound like? It is possible that in the visual modality comparisons between two series in a line graph format may work well. In the auditory modality other strategies may be more useful; one strategy is to explicitly calculate a new series using a comparison between the original two series; Alternatively, simultaneous sonifications of the two series could be attempted if distinction between the series is made more efficient through the use of acoustic strategies such as spatial separation or timbre discrimination [37].

The point at which a sonification starts when a gesture is enacted is another factor that is extremely important with multi-modal designs such as this. There is no play button (or mouse button to click), and therefore a sound that starts as soon as a gesture is initiated can confuse unless the gesture is absolutely correct and intended. Alternatively, a sound that occurs as you make the gesture (dragging etc.), can become annoying, but is also highly immediate. A sound that happens when you release the gesture is another possibility, although that requires some carefully practised actions, to make sure that the gesture is exactly what was intended. Solutions to this design issue need to take several considerations into account:

1. What is the likelihood, and cost of, making an incorrect gesture?
2. Is the sound created information carrying?
3. Does the sonification require a playback process—what is the duration of the sound—and is it an immediate correspondence to the gesture?

Many of these issues can be solved simply by using faster sonifications.

The surprising thing about this research is the different manner in which sonification is viewed when the presentation format is changed. Sonifications can be inherently unsatisfying, in that during their playback, the lack of scales and markers makes it difficult to maintain a mental model of the data under investigation, let alone other data dimensions that may be related. Furthermore, restarting the sonification is often difficult. However, when sonification is partnered with the persistence of a visual display, as well as easy and immediate ways of replaying the sonification, it appears the listener is freed to attend to and enjoy the sonification. Whether this anecdotal result is actually the case among groups of users is an interesting question for future research.

5.3 Further research plans

Use of time-series data is appropriate, but the other major form of data that is important is multi-dimensional statistical data, such as that is created by scientific experiments or observations. Many more complexities exist with this type of data, such as how to incorporate categorical (non-numerical) data, and how to find common axes between which to organise data dimensions.

With the capability to perform statistical comparisons between two series, another question naturally arises as to how these should sound. This is due mainly to the multimodal context, as a sonification on its own does not invite comparison—most of the effort is spent understanding the information, but a sonification that can be accessed by touch is possibly quicker to be compared against similar series. Some work has been done on the use of sonification for statistical comparisons (for instance [15, 16, 20]), but there exists a gap somewhere between the high-level sonifications dealing with multivariate data, and the simple parameter mapping sonifications. There are few sonification analogs to common statistical procedures for comparing and contrasting data series and performing simple statistical testing. This research has highlighted a context that would benefit greatly from them. It also provides them with a significant amount more representation context, so that the necessary abstraction can be processed in concert with lower-level representations. This is difficult to provide in an auditory only context, and so a multi-modal approach holds much promise.

This research has sought to heuristically evaluate the navigation principles described. An evaluation using data analysis users would also be of value and is a next step.

5.4 System implementation

The system described is implemented on a rear diffused illumination multi-touch tabletop computing system. This system is based around the *Bricktable* design by Hochenbaum and Vallis [24] but is generally a simple multi-touch table with a surface area of approximately 1 m by 0.6 m and standing 0.8 m from the ground. This configuration allows for between 1 and 3 users to be within easy reach of the tabletop, while standing on one side of the table. A short-throw projector is used for projection of the image, and infra-red illumination of the hand positions is captured by an infrared camera (modified PS3 Eye). The projection surface used was drafting film, and the finger movements were captured using Community Core Vision[5] and a PS3 Eye camera. The use of sonification necessitates loudspeakers of some type, and to provide a stereophonic sound-stage for the users (at least) these must be positioned in a relatively controlled manner. Typical multi-user table top configurations, which allow users to move around the table and re-orient themselves as they wish, may result in the loss of control over the auditory image with respect to the listener.

The sonification synthesis implementation was programmed in the java programming language (along with the rest of the application), using FM synthesis that was implemented with Ollie Bown's *beads* java audio library.[6]

[5] http://ccv.nuigroup.com.

[6] http://www.beadsproject.net.

6 Conclusions and limitations

We have presented some investigations into user interfaces for interaction with statistical data using visualisation and interactive sonification. We have discussed the issues of arrangement, zoom and scale, selection, and playback of data. We have also discussed the way these concepts are implemented in a user interface with a set of two general modes for interaction with data representations—movement and arrangement, and exploration and comparison.

This research has focused on linking concepts regarding data analysis and representation, in the auditory, visual and touch modalities, into a common framework expressed in a user interface design. We were surprised by the different representation requirements for sonification that exist when its context is changed. The usefulness of a supporting visual context for sonification therefore goes well beyond redundant encoding only. The need for future sonifications of statistical tests and comparisons of various natures was discussed. The benefits of rapid overview sonifications, with durations in order of 1–2 seconds, rather than 10–15 seconds, was made clear when multiple time-series are presented and made 'touchable'.

Much further research is needed to build a truly flexible system for mainstream data analysis. Many practical issues still need to be addressed for interactive data analysis, such as data output formats and standardisation. Interactive data query building is probably also important when searching for relevant datasets with which to compare existing data, especially where multiple sources of data exist.

Methods for representing, investigating, sonifying and comparing representations are numerous—this paper has outlined approaches for organising the interaction between representations in terms of general navigation categories.

References

1. Adcock M, Barrass S (2004) Cultivating design patterns for auditory displays. In: Proceedings of the 10th international conference on auditory display, Sydney
2. Burkhard R et al (2007) Visualisation summit 2007: Ten research goals for 2010. Inform Visual 6(3):169–188
3. Aldersey-Williams H, Hall P, Sargent T, Antonelli P (2008) Design and the elastic mind. The Museum of Modern Art, New York
4. Anderson E (1935) The irises of the gaspe peninsula. Bull Am Iris Soc 59:2–5
5. Barrass S (2003) Sonification design patterns. In: Proceedings of the 9th international conference on auditory display, Boston
6. Barrass S, Schaffert N, Barrass T (2010) Probing preferences between six designs of interactive sonifications for recreational sports, health and fitness. In: Proceedings of ISon 2010, 3rd interactive sonification workshop, Stockholm, Sweden
7. Beilharz K, Ferguson S (2009) An interface design for aesthetic and interactive sonification. In: Proceedings of the 15th international conference on auditory display, Copenhagen, Denmark

8. Bertelsen O, Pold S (2004) Criticism as an approach to interface aesthetics. In: Nordic conference on human-computer interaction, Tampere, Finland
9. Bovermann T, Groten J, de Campo A, Eckel G (2007) Juggling sounds. In: International workshop on interactive sonification, York, UK
10. Bovermann T, Hermann T, Ritter H (2006) Tangible data scanning sonification model. In: Proceedings of the 12th international conference on auditory display, London, UK
11. Chen C (2005) Top 10 unsolved information visualisation problems. IEEE Comput Graph Appl 25(4):12–16
12. Collins A, Apted T, Kay J (2007) Tabletop file system access: associative and hierarchical approaches. In: IEEE international workshop on horizontal interactive human-computer systems (TABLETOP'07), pp 113–120
13. Collins A, Kay J (2008) Collaborative personal information management with shared, interactive tabletops. In: CHI 2008, Florence, Italy
14. Cowie A (2010) Collaborative concept mapping at an interactive tabletop. PhD thesis
15. Ferguson S (2009) Exploratory sound analysis: statistical sonifications for the investigation of sound. PhD thesis
16. Flowers JH, Buhman DC, Turnage KD (1997) Cross-modal equivalence of visual and auditory scatterplots for exploring bivariate data samples. Hum Factors 39(3):341–351
17. Fry B (2004) Computational information design. PhD thesis
18. Fry B (2008) Nature hapmap cover. Nature 437(7063):1 (cover)
19. Geiselhart K (2010) Towards global governance—interactive technologies and global accountability measures. Public Commun Rev 1(1):37–46
20. Hermann T (2002) Sonification for exploratory data analysis. PhD thesis
21. Hermann T (2008) Taxonomy and definitions for sonification and auditory display. In: Proceedings of the 2008 international conference on auditory display, Paris, France
22. Hermann T, Bovermann T, Riedenklau E, Ritter H (2007) Tangible computing for interactive sonification of multivariate data. In: International workshop on interactive sonification, York, UK
23. Hermann T, Niehus C, Ritter H (2003) Interactive visualisation and sonification for monitoring complex processes. In: Proceedings of the 9th international conference on auditory display, Boston, USA
24. Hochenbaum J, Vallis O (2009) Bricktable: a musical tangible multi-touch interface. In: Proceedings of Berlin open conference 09, Berlin, Germany
25. Hochenbaum J, Vallis O, Diakopoulos D, Murphy J, Kapur A (2010) Designing expressive musical interfaces for tabletop surfaces. In: Proceedings of the 2010 international conference on new interfaces for musical expression, Sydney, Australia
26. Hunt A, Hermann T (2004) The importance of interaction in sonification. In: Proceedings of the 10th international conference on auditory display, Sydney, Australia
27. Judelman G (2004) Aesthetics and inspiration for visualisation design: Bridging the gap between art and science. In: International conference on information visualisation (IV'04), London, England
28. Kosara R (2007) Visualisation criticism—the missing link between information visualisation and art. In: International conference on information visualisation (IV'07), Zurich, Switzerland
29. Kramer G (1993) Auditory display: sonification, audification, and auditory interfaces, Perseus, New York
30. Kramer G, Walker BN, Bonebright T, Cook PR, Flowers JH, Miner N, Neuhoff JG (1997) Sonification report: status of the field and research agenda. Tech rep, National Science Foundation
31. Mayo E, Steinberg T (2010) The power of information: an independent review
32. Nees M, Walker B (2006) Relative intensity of auditory context for auditory graph design. In: 12th international conference on auditory display, London, UK, pp 95–98
33. Pousman Z, Stasko J, Mateas M (2007) Casual information visualisation: depictions of data in everyday life. IEEE Trans Vis Comput Graph 13(6):1145–1152
34. Sack W (2006) Aesthetics of information visualisation. In: Paul C, Vesna V, Lovejoy M (eds) Context providers. MIT Press, Boston
35. Scha R, Bod R (1993) Computational esthetics. Inf Infbeleid 11(1):54–63
36. Shneiderman B (1996) The eyes have it: a task by data type taxonomy for information visualisations. In: IEEE symposium on visual languages, Boulder, CO, USA
37. Song H, Beilharz K (2007) Concurrent auditory stream discrimination in auditory graphing. J Comput 3(1):79–87
38. Tufte ER (1992) Envisioning information. Graphics Press, Cheshire
39. Tünnermann R, Hermann T (2009) Multi-touch interactions for model-based sonification. In: Proceedings of the 15th international conference on auditory display, Copenhagen, Denmark
40. Moere Vande A (2005) Form follows data: the symbiosis between design and information visualisation. In: International conference on computer-aided architectural design (CAADfutures'05). OKK Verlag, Vienna
41. Vesna V (2007) Seeing the world in a grain of sand: the database aesthetics of everything. In: Vesna V (ed) Database aesthetics—art in the age of information overflow, electronic mediations. University of Minnesota Press, Minneapolis
42. Viegas F, Wattenberg M (2007) Artistic data visualisation: beyond visual analytic. In: Online communities and social computing. Lecture notes in computer science, vol 4564(15). Springer, Berlin, pp 182–191
43. Walker BN, Nees MA (2005) An agenda for research and development of multimodal graphs. In: Proceedings of the 11th international conference on auditory display, Limerick, Ireland
44. Wilkinson L (2005) The grammar of graphics, 2nd edn. Springer, Berlin
45. Zacks J, Levy E, Tverksy B, Schiano D (1998) Reading bar graphs: effects of extraneous depth cues and graphical context. J Exp Psychol, Appl 4(2):119–138
46. Zacks J, Levy E, Tversky B, Schiano D (2001) Graphs in print. In: Olivier P, Anderson M, Meyer B (eds) Diagrammatic representation and reasoning. Springer, London