Citation for published version


DOI

Link to record in KAR

http://kar.kent.ac.uk/14576/

Document Version

UNSPECIFIED

Copyright & reuse
Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (e.g., Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research
The version in the Kent Academic Repository may differ from the final published version. Users are advised to check http://kar.kent.ac.uk for the status of the paper. Users should always cite the published version of record.

Enquiries
For any further enquiries regarding the licence status of this document, please contact: researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at http://kar.kent.ac.uk/contact.html
SONIFICATION OF SPATIAL DATA

Tooba Nasir
Computing Laboratory
University of Kent
Canterbury, UK.
tn37@kent.ac.uk

Jonathan C. Roberts
Computing Laboratory
University of Kent
Canterbury, UK.
j.c.roberts@kent.ac.uk

ABSTRACT
Sonification is the use of sound and speech to represent information. There are many sonification examples in the literature from simple realizations such as a Geiger counter to representations of complex geological features. The data that is being represented can be either spatial or non-spatial. Specifically, spatial data contains positional information; the position either refers to an exact location in the physical world or in an abstract virtual world. Likewise, sound itself is spatial: the source of the sound can always be located. There is obviously a synergy between spatial data and sonification. Hence, this paper reviews the sonification of spatial data and investigates this synergy. We look at strategies for presentation, exploration and what spatial interfaces and devices developers have used to interact with the sonifications. Furthermore we discuss the duality between spatial data and various sonification methodologies.

[Keywords: Sonification, spatial data sonification, information representation]

1. INTRODUCTION
Sonification is the representation of data into the sound domain using non-speech audio [16]. Through this mapping the user is able to make nominal, qualitative or quantitative judgments on the information being heard. That is, sonification can communicate a category or name, or the relative size of the data value (whether something is larger or smaller than something else) or the exact value of that data, respectively.

There is a growing interest into sonification, not only is it useful for accessibility to (say) represent information to users who are blind or partially sighted, but it enables more variables to be presented in one display and some information (such as rapidly changing information) is better suited to the sound domain.

In addition, geographical visualizations and other spatial data visualizations are important. For instance, we often utilize route maps to navigate and refer to world maps to locate a holiday destination or read off the x,y coordinates of a point on a scatterplot. Considering specifically geographical visualizations. (1) Spatial visualizations permit the user to analyze large datasets. (2) The required information is readily available. (3) Lots of different types of information can be co-located and hence compared, and (4) items that are in close proximity relate to each other and thus can be easily manipulated [8]. However, in comparison with geographical visualization the mapping of spatial data into sound is difficult.

First, the mapping of spatial data in Geographical Information Systems (GIS) is well developed and documented whereas sonification research is still in its infancy. In fact, map creation has been around for thousands of years and developers know how to allocate the graphical components because they follow well-formulated design guidelines. Hence the GIS community does not focus their research effort on the mapping process, rather the challenge in the community is towards the analysis, processing and management of multidimensional spatial datasets. Conversely, sonification of spatial data is not well developed. There are few guidelines and developers still focus on the mappings.

Second, in traditional visualizations the mapping is implicit and accurate: spatial data is positioned on an x,y grid and may be accurately located and hence comprehended.

While in sonification, although sound may be spatial, it is not inherent how to map the information, and the perception of the information is less precise. For instance, although sound may be mapped to a position in the azimuth plane, users are unable to accurately locate the position of the sound source as accurately as they could locate the information in an equivalent graphical visualization [20].

Third, spatial data in geographical visualizations are mapped to two-dimensional spaces while this need not be the case for sonification. For instance, someone explaining to their colleague the route from the workplace to their home is spatial information, but the communication medium (speech) is not spatial.

This paper focuses on sonification of spatial data, in particular geo-spatial data. We develop a categorization that divides the research into four overarching categories (section 2). This taxonomy is then used to categorize the main research papers in the area (sections 3,4 and 5). Then we
describe the different interfaces and devices that developers have used that allow the users to spatially interact with the information (section 6). Finally we discuss the duality between spatial data and various sonification methodologies (section 7). The scope of this paper is on the use of sound to represent data. There has been much research in the area of spatial sonification in virtual environments (VE) [13] where sound is used to enhance the sense of presence in the VE, or surround sound setups have been used to realise complex scenes or high fidelity realistic worlds. But, these latter examples are not included because they demonstrate acoustic renderings, rather than representing value information to the user.

2. CATEGORIES & BACKGROUND

This section details the categories and provides some background information. Some of the background information may be readily known by a sonification developer, but is included here for completeness and to develop the taxonomy structure. There are two main parts to this section: first, spatial and non-spatial data and sound, and second, the components of sonification.

2.1. Spatial & Non-Spatial data and sound

The overarching categorization groups the research into four parts, see table 1. Because this paper is particularly focused on the notion of spatiality and location the category on ‘Non Spatial Data with Non Spatial Data’ is not included and is out of the scope of this paper.

<table>
<thead>
<tr>
<th>Spatial Data</th>
<th>Spatial Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>with</td>
<td>with</td>
</tr>
<tr>
<td>Spatial Sound</td>
<td>Non Spatial Sound</td>
</tr>
<tr>
<td>(Section 3)</td>
<td>(Section 4)</td>
</tr>
<tr>
<td>Non Spatial Data</td>
<td>Non Spatial Data</td>
</tr>
<tr>
<td>with</td>
<td>with</td>
</tr>
<tr>
<td>Spatial Sound</td>
<td>Non Spatial Sound</td>
</tr>
<tr>
<td>(Section 5)</td>
<td>(Section 5)</td>
</tr>
</tbody>
</table>

Table 1: Spatial & Non-spatial Mappings

2.2. Non-spatial and Spatial Components of sonification

The well known semiologist Jacques Bertin [4] states that a visualization developer should perform a component analysis; to analyze both the components of the data and those of the visual domain, and work out an effective mapping from one to the other. He named the components of the visual system *retinal variables* which he used alongside x,y spatial components. We use this same categorization, but extended to sonification. Location information can be used to enhance the sonification or can be used to represent qualitative information. Sounds can be localized through four methods, shown in Figure 1. Hence there are two groups of non-spatial components and four spatial. (1) non-spatial audible variables, (2) non-spatial motifs, (3) Interaural Time Difference (ITD), (4) Interaural Intensity Difference (IID), (5) Doppler effects and (6) Environment effects.

**Non-spatial audible variables** are the building blocks for sonification. They include pitch, loudness, attack and decay rates, timbre, tempo and brightness. For instance, a developer may wish to communicate that a company’s stock is increasing over time, this is similar to a graphical line graph, thus they could map the value of the stock to pitch and hours of the day to time. As well as mapping the best variable to the data dimension, the developer needs to decide on the scale and polarity of that mapping [29, 28].

**Non-spatial motifs** are higher order components. They utilize the variables to communicate the information at a higher-level; they have a specific structure and may need to be learned. For instance, Earcons [5] utilize the audible variables to communicate different objects through sound motifs, and the similarity in the data is represented by similar motifs. Another example is by Franklin and Roberts [9] who demonstrated in their ‘pie chart sonification’ how a

---

1In this categorization we use the term motifs as a general term to describe any higher-order sonification mapping
human listener, oriented to face along x-axis

Interaural time difference - left to right perception

Interaural time difference - radial perception

Human hears information radially

Using ITD and IID a user can locate sounds above and below

Pitch changes as sounds move

Doppler Perception

Further away represents lower value

Closer sound source represents higher value

Interaural Intensity Difference

ITD and IID for vertical location

Sound gets dampened through the environment.

Figure 1: A and B show Interaural Time Difference (ITD) to provide left right and radial information, C demonstrates the Interaural Intensity Difference (IID), D shows that through ITD and IID vertical positions may be detected, while E shows that location may be found through Doppler effects and F through the environment.

Structure similar to Morse-code could be used to present quantitative information.

The principle of Interaural Time Difference (ITD) is that there is a phase difference between the sound arriving at the left ear compared with the right. Whereas the principle of the Interaural Intensity Difference (IID) is that objects which are closer sound louder. Multiple speakers allow the user to perceive the sound from different locations. In fact, ITD on its own permits the user to locate sound in the azimuth plane whereas ITD along with IID allows the user to perceive sounds azimuthally and in elevation [15].

Doppler and time-based effects. Factors, such as the Doppler or frequency changes, give a listener perception of source distance and movement from the listener’s position perspective. The siren of an ambulance on the road grows closer and louder as it approaches a listener and then starts to fade as it rushes away. Echo is another distance location method which could be used to sonify distances.

Finally, the environment which the sound is displayed will effect how the sound is perceived. This effect can be used to locate objects. The environment encompasses factors such as reverberation, reflection and sound occlusion. For instance, in a furnished carpeted room the sounds generated by people or machinery is soft while in an unfurnished room or tiled room the sound echos and reverberates. Hence, if a user knows the position of some objects then the user will be able to locate the sound as it moves behind different objects in turn. Furthermore, if the user knows the exact path by which a sensor moves (such as a maze, or a zig-zag path on a 2d image) then the user can understand different values of the data as the sensor moves through the world.

3. Spatial Data with Spatial Sound

In this section we discuss related research of spatial data mapped to spatial sound.

3.1. Interaural Time Difference (ITD) - Left-Right Perception

Smith et al.[27] presented multidimensional datasets in sound using stereo effects to provide the location information. They presented the data both through sonification and visualization, and the user could zoom into the display and change the orientation of their avatar for the sonification. Each data component was represented by a glyph in the visualization and a sound motif for the sonification. The user could notice trends in the data through the texture of the graphic display and they could hear the sounds from the motifs coming from different locations to gain an understanding of clusters and groupings in the data.

Minghim and Forrest [21] presented scientific visualizations of scalar volume data sonified using stereo balance for
direction and orientation information, and timbres to represent surface structures.

3.2. Interaural Time Difference (ITD) - Radial Perception

Franklin and Roberts [9] presented five different mappings for graph data to sound. In four of the presentations the user was placed in the center of the pie-chart facing towards the zero percentile in the azimuth plane. The edges of each pie segment was represented by a sound located on the circumference; the different designs changed whether the start and end point was sounded, and whether the start point always was normalized to the forward position. The location information was generated through ITD’s with the user wearing either headphones or using surround sound speakers. The fifth design utilized non-spatial audio by representing the pie values by Morse-code. Their work evaluated the five designs and showed that the non-spatial Morse-code version was the most accurate, with the non-normalized version being the next accurate. They also discussed issues to do with the Minimum Audible Angle [20], saying that because the accuracy of spatial sound perception depends on the radial location of the sound source, the act of locating pie segments to the immediate left or right of the user is least accurate.

3.3. Interaural intensity difference (IID) - Loudness vs Distance

Gardner [12] suggested that a logical mapping of sound attributes to a spatial dataset is loudness mapped to distance. However, to our knowledge no researcher has solely used IID to represent spatial data.

3.4. ITD & IID - High & Low Perception

The Interaural Time Difference (ITD) along with Interaural intensity difference (IID) gives a listener the perception of sounds azimuthally and in elevation, see Figure 1D. The use of speakers provide a listener the spatial perception of an azimuth plane and the elevation of the sound source.

Work in spatial sonification of spatial datasets includes sonification of atmospheric and weather data for storm activity recorded for a geographic region spanning over 1000 km [23]. The data used for sonification was taken from modeled storm activity at different elevation levels and six out of nine recorded variables were used for sonification, i.e. atmospheric pressure, water vapor, relative humidity, dew point, temperature, and total wind speed. Each variable was mapped to the pitch of a sound sample of a distinct timbre. The sonification was based on a customized 16 speaker arrangement where the speakers were mapped to geographical location points on the mapped data in north-south and east-west directions. The final sonified storms were presented as compositions of the sonified variables. Ringing bells sounds were used to mark the time and elevation of each composition.

Some work has been done to help blind people understand geographical maps, allowing active exploration and navigation and auditory feedback for ‘details on demand’. Zhao et al. [31] presented spatial sonification geographical distribution pattern of statistical data. The statistical values on a map of the USA data were mapped to pitch while the 2D location of the geographic region was mapped to the sound location. Five patterns for active exploration were created such that the map could be traced on a vertical strip, a horizontal strip, a diagonal strip, in a cluster or in a non-pattern. A keyboard and a tactile tablet were used as user interfaces; the numeric pad on keyboard was used for navigation through the map with the arrow keys moving the user in their respective directions. The tactile tablet allowed a user to activate a region of interest at a finger position. When the user moved over a geographic region they could hear the sonified non-speech value associated to the region; the region name and the statistical value were presented to the user as speech feedback. The user could hear any combination of the sonified output options.

3.5. Doppler & Time Effects

A good example of using Doppler or time effects for sonification is presented by Hermann and Ritter [14]. They say “in our world, normally passive objects are silent ... sound occurs when the system becomes excited”. In their work they present a virtual physics that models the vibrational process. They provide an example of virtual sonograms for exploring the trajectory of particles on a two-dimensional plane.

Saue [25] proposed a temporal mapping of data to sound, where the sound changes depending on the users’ position. Two and three dimensional spatial datasets were sonified and the sound data mappings were time dependent. The datasets chosen for sonification were subsets of seismic [26] and medical imagery, ultrasound images and a micro listener’s movement inside the human body. The data was represented as streams and the sequence order of these streams was represented as implicit time. Each data sample was assigned a data to sound mapping and the samples were run through the predefined mapper at a specific speed. For two and three dimensional datasets two alternative sound mappers were used. The first technique involved creation and sonification of trajectories in the dataset.

The trajectories defined an implicit time equal to the 1D case, and could be played automatically at a constant speed or through a pointing device interaction. The second method was based on spatializing sound maps and assigning each with an individual time. A parameterized sound
located in 3D space was associated to an object when selected. Exploration of the dataset was based mainly on orientation. The data mappers were defined over points and regions and computed local maxima of the object, to extract position and value information from it. These computed values were then mapped to sound through the sound mappers. All localized sounds related to listener’s position. The sound parameters were scaled relative to the maximum and minimum values in the objects. This scaling resulted in the zooming in effect such that the listener was moving towards the sound source. Saue argued that choosing a temporal sound mapping for spatial data strengthens and supplements data comprehension [25].

3.6. Environment effects

There are no obvious examples of people using environmental effects to describe spatial data. For example, if the user knows the environment and notice how the sound changes through that environment then they will understand where that source is located. The closest work is that of path based sonifications, which are included in the next section.

4. SPATIAL DATA - NON-SPATIAL SOUND

Non-spatial sounds are usually used to represent quantitative information. Non-spatial sound components include motifs, auditory icons and Earcons [5]. Speech feedback is also used as a non-spatial sound representation technique for textural data. These non-spatial representations of sound enhance a sighted user’s perception of the graphical and sonified representations of the dataset and provide the information to blind users that they are unable to see. For example, the Talking Tactile Tablet [17] consists of tactile sheets embossed with raised lines and textures describing images, maps and diagrams. A symbol, icon or region on the map can be pressed to get the non-spatial audio information about it. The Tablet reads out the name of the selected object and outputs a sound associated with that object.

There are various examples of path-based sonifications; where a path is placed through the spatial data and sampled sequentially. The sampled points are then sonified. For example, the well known vOICe [19] application displays a 2D image on a designated path. Madhyastha and Reed [18] noticed that various people were sonifying two dimensional datasets and hence presented their toolkit named Porsonify. Franklin and Roberts [10] further explore this concept, detailing that the path has a direction, occluding front along and a path envelope.

Sonification of responsive well-logs [2] is another example of a path-based sonification. In this case, the datasets used were seismic surveys, well-logs and directional well-logs. The sonification was based on the metaphor of a virtual Geiger Counter and integrated with a three-dimensional visualization developed for well-logs [11] where data attributes were mapped to a bivariate color scheme and on a sliding lens. Various timbres (e.g. cello, trombone and bassoon) were used to represent different variables and multiple attributes could be simultaneously played. The data was sonified for different resolutions. A closer and clearer sonification of features such as peaks and boundaries was made possible through sweeps over an area of interest. Directional well-logs were sonified spatially. A virtual sound source was placed away from the user, pointing in the direction of the data to show that spatial sound conveys spatial correlation and spatial patterns.

Alty and Rigas [1] presented a non spatial sonification of geometric shapes. The shapes were presented as objects on a graph. These objects were represented by sound which conveyed the objects’ shape and position on the graph. The coordinate size was mapped to pitch in a chromatic scale, while the X and Y coordinates were distinguished by timbre. Short distinct earcons represented control actions i.e. shape selection, shape resize and dragging, and loading and saving files. The system was presented with a visual as well as audio interface. The graph area was scanned for sonification output in any one of the three possible scanning techniques: Top-down scan, center scan where the scan started in the center of the graph and grew outwards in a circle, and ascending scan which scans the objects in space in ascending order of their size. The system produced stereo sound output.

Finally, Bennett and Edwards [3] and in particularly Bennett in his PhD thesis described a method of sonifying diagrams. In their work, the x,y positions of objects on the display were sonified. Higher pitches were allocated to higher values of x.

5. NON SPATIAL DATA - SPATIAL SOUND

Ramloll et al.[24] presented a spatial sound mapping for an audio tactile line graphs. Users were positioned on the x-axis and could hear the graph, which was represented by pitch, as they followed the line with a haptic display, such that when the line is above the x-axis the listener hears the sound coming from their left ear, and below the x-axis from their right ear. This was used alongside a haptic force-feedback device (the Phantom) to allow the users to feel the graph at the same time as hearing it. This is shown in Figure 1A. Furthermore, they incorporated speech into the system to enhance the haptic display.

While most research work has been focused on the mapping of single data series, some researchers have also explored the possibility of sonifying multiple non spatial data series in order to make multiple data series graphs more accessible for visually impaired people [6]. Musical notes mapped on graph data, the y-values were mapped to the pitch of musical instruments. As the y-values on the graph
go higher the pitch of the musical notes increases. This technique was used to sonify two and three data series at the same time.

6. INTERFACES, EXPLORATION AND DEVICES

The present challenge in sonification of datasets, be it a spatial dataset or non spatial, is not the mapping alone but also an interface and user interaction with the data as well as its sonification. Logically a spatial interface would be required for interaction and exploration of a spatial dataset. We categorize these interfaces into four types that allow a user to interact with data and sonification spatially i.e. mouse, keyboard, tablet (graphic tablet or a tactile tablet) and haptics (force feedback).

Mouse interfaces. Smith et al. [27] presented an icon based auditory display. Users could move a mouse over the icons to activate an auditory texture. The formation of these textures was dependent on how the mouse moves i.e. slow or fast, linear or circular, and small or wide display area. The resulting sonification thus provided the user with spatial information based on the texture formation.

Saue [25] also used a mouse as an interaction device with the display to move an active listener around the dataset regions on the display. A user could mark places of interest in the dataset and go back to play them at a later stage. All localized sounds related to listener’s position.

Poll’s storm data sonification [23] provided an easy exploration and selection of the dataset with a mouse. A user selected an elevation level for the sonified storm activity and the speaker location on the map and pressed the sound icon on the display to play the storm. The graphical and sonic interface presenting geographic and elevation information for the storms.

Zhao and Shneiderman [31] used a keyboard along with a tactile tablet for an interface on their sonified geographical map. This combination allowed the users to navigate the map easily and north-south and east-west direction and select an area of interest using the tablet to retrieve more information. Another interface with Tactile feedback was combined with sound to teach spatial information in a digital map exploration and get audio feedback on locations of interest [22].

Barras and Zehner in their ‘responsive sonification of well-logs’ [2] chose to use a 3d haptic interface in a Responsive Work-bench. A probe and a dial control panel were used for the interaction. The information was sonified using a virtual Gieger counter. The sonification probe, with 3D spatial tracking and 6 Degrees of Freedom, was used to explore the visualization. Users could move it it vertically on the display to interact with virtual objects in the 3D graphical interface.

Weinberg et al. [30] allowed the users to interact with the system with tactile controllers and generate their own spikes in the environment for sonification and for interaction with other users. A video display supported comprehension of the sonified dataset. They also used a GUI to depict frequency bands which allowed the user to interact with the dataset choose the audio output format such as live or recorded. The sonification was presented on speakers and represented the sound projection in brainwaves. Users interacted with the system with tactile controllers and could generate their own spikes.

7. DISCUSSION & CONCLUSION

The research of this paper demonstrates that researchers have not fully utilized the maximum potential of spatial sound. For instance, there are only two examples of researchers using Doppler and time effects to represent distance and no obvious examples of researchers utilizing environmental effects to visualize data. The work by Hermann and Ritter [14] is an excellent example of how motion can realize two dimensional effects, but there is unquestionably more research to be done here. For example, echo location or other factors such as reverberation and spatial occlusion could be used to visualize spatial information.

There are many non-spatial variables that can be used to sonically realize the information, as detailed in section 2.2. One important area is speech output. It is often hard to understand quantities from sonifications, but speech provides the user with exact quantifiable information. For example, Zhao et al. [31] used speech feedback to verbalize statistical values of a geographical map. A natural extension of this is to use a two-dimensional tablet alongside the speech interface (especially utilizing tactile overlays). The Talking Tactile Tablet [17] and other such devices provide a natural two dimensional two-way interaction, allowing the user to tactually and spatially interact with the data and listen to appropriate information. Furthermore, haptics devices provide spatial and possibly additional information [24, 31, 30], but the inclusion of haptic devices with sonification especially spatial sonification is also infancy.

Auditory displays have been used to express aspects of information that are difficult to visualize graphically, this is certainly true of the multivariate information that was presented by Smith et al.[27] (see section 3). Auditory information also enhances visual information when used in conjunction with its visual equivalent or augmented with haptics or tactile. Sonification of a non spatial dataset has the potential to convey important information that might either be hidden from the human eye or is negligible in a visualization overview [7]. So this is obviously another area for further research.

There are definitely challenges with the perception of data through sound, and spatial sonification relies upon several models and assumptions. More accurate models such
as HRTF’s should be used to create accurate positional mappings, and error metrics such as the Minimum Audible Angle [20] should be referenced to create appropriate mappings and effective evaluations.

In conclusion, there is definitely a synergy between spatial data, spatial sonification techniques and spatial interfaces to provide the exploration. But, spatial sound is certainly not the only way to visualize spatial data. While the majority of researchers have used spatial sonification for spatial datasets and have used spatial interfaces for interaction and exploration of the sonification and the dataset itself, as can be seen from the categorization of related work in this paper, spatial sonification has to be used alongside non-spatial variables to maximise the perception of the information.

8. REFERENCES


