Citation: Ballweg, Holger Matthias (2019) Ways of Guided Listening: Embodied approaches to the design of interactive sonifications. Doctoral thesis, Northumbria University.

Ways of Guided Listening: Embodied approaches to the design of interactive sonifications

Holger Matthias Ballweg

Supervisor: Dr. P. Vickers
Prof. A. Bouridane
Dr. A. Bronowska

Department of Computer and Information Sciences
Northumbria University

This dissertation is submitted for the degree of
Doctor of Philosophy

June 2019

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and acknowledgements.

Any ethical clearance for the research presented in this thesis has been approved. Approval has been sought and granted by Research and Business Services at Northumbria University under record number RE-EE-14-141029-5450bac0eb9f3 on 02/09/2016.

I declare that the word count of this thesis is 29,454 words.

Holger Matthias Ballweg
June 2019

Acknowledgements

I thank my supervisory team for their support and guidance.

I thank Kirsty Lindsay and Nick Caplan for their fruitful collaboration.

I thank the university for giving me the opportunity to do this research.

I thank my family and friends for their encouragement and advice.

I thank my partner, Shelly Knotts, for all her patience and support.

Abstract

This thesis presents three use cases for interactive feedback. In each case users interact with a system and receive feedback: the primary source of feedback is visual, while a second source of feedback is offered as sonification.

The first use case comprised an interactive sonification system for use by pathologists in the triage stage of cancer diagnostics. Image features derived from computational homology are mapped to a soundscape with integrated auditory glance indicating potential regions of interests. The resulting prototype did not meet the requirements of a domain expert.

In the second case this thesis presents an interactive sonification plug-in developed for a software package for interactive visualisation of macromolecular complexes. A framework for building different sonification methods in Python and an OSC-controlled sound producing software was established along with sonification methods and a general sonification plugin. It received generally positive feedback, but the mapping was deemed not very transparent.

From these cases and ideas in sonification design literature, the Subject-Position-Based Sonification Design Framework (SPBDF) was developed. It explores an alternative conception of design: that working from a frame of reference encompassing a non-expert audience will lead towards sonifications that are more easily understood. A method for the analysis of sonifications according to its criteria is outlined and put into practice to evaluate a range of sonifications.

This framework was evaluated in the third use case, a system for sonified feedback for an exercise device designed for back pain rehabilitation. Two different sonifications, one using SPBDF as basis of their design, were evaluated, indicating that interactive sonification can provide valuable feedback and improve task performance (decrease the mean speed) when the soundscape employed invokes an appropriate emotional response in the user.

Preface

Outlined here are the contributions by others, the author, and contributions made in collaboration with others.

The image sonification outlined in Chapter 3 and the data analysis was designed and implemented by the author. The data and some guidance on possible data analysis methods was supplied by Prof Bouridane.

ChiSon, the protein sonification system described in Chapter 4 was designed and implemented by myself, deciding on which protein and which data domain to sonify in collaboration with Dr. Bronowska. The experiment was designed in collaboration with Dr. Bronowska.

The SoniFRED experiment described in Chapter 6 was designed in collaboration with Kirsty Lindsay and her supervisor Prof Nick Caplan. They designed the exercise routine and the way feedback was recorded. Lindsay, Caplan, Vickers and the author decided collaboratively on the different feedback conditions. The author designed question 1.3, 1.4, 1.5 on Questionnaire 1 and part 2 of the second questionnaire. Preliminary analysis of the recorded quantitative data was done by Kirsty Lindsay and replicated by myself. The author conducted all qualitative data analysis and the ANOVA published in this thesis.

The beach sonification was designed and implemented by Dr. Vickers. The author adapted it to take live input from the FRED exercise device and integrated it into the system used in the experiment.

The author designed and implemented the train sonification and the system used to run the experiment.

The design framework described in Chapter 5 was developed in response to ideas set out by Vickers and in dialogue with him. The author chose the analysis-based approach to develop the framework, conducted the analyses, and designed the guidelines, as well as using them in their prototypical form in his SoniFRED sonification.

Table of contents

Chapter 1

Introduction

1.1 Background

Visual feedback is ubiquitous in the world today — indicator lights, symbols, and text flood our visual field from the time we open our eyes in the morning to shortly before we close them at night. Auditory signals are also employed to deliver notifications about messages received on a mobile phone or doors not closed in an automobile. Whereas we can sometimes discern which type of message is being received by the different sounds used, more information is only available to the eyes - the icon showing us which door is open or who messaged us. Maybe it would be better if we could gain more information from those sounds than just the arrival of new information.

1.2 Research motivation

This thesis explores multiple applications for interactive feedback, where users interact with a system and receive feedback according to their interaction. Similar to those indicators described above, the primary source of feedback in these cases is visual, while a second source of feedback is offered in another modality. In the case of message notification tones above, another form of feedback could be speech synthesis of the sender's name or even just assigning arbitrary sounds to each contact. This would enable the user to triage the handling of the message accordingly — but privacy concerns would be a major problem, as probably only a minority of people want other people around them to know exactly or approximately who is messaging them. In this case, the user just wants to know what kind of message was received or even just that a message was received.

On the other hand there is no reason to not indicate which car door is open - most cars contain enough speakers to indicate a direction as well as having something light up on the dashboard. Speech synthesis in this context could be perceived as uncanny depending on the execution, or unwieldy as localisation would become an issue, whereas just a directed signal would provide both the heads-up to have a look at the dashboard and information about the state of the car. Without any legend or training a regular user of the car could then use the additional spatial auditory feedback to save themselves a look at the dashboard. We could argue that this would equally well be served by speech synthesis as localisation can be dealt with in most cases, but maybe a spatialised speech synthesis ("Rear left passenger door open" reproduced by the rear left speaker) would be the best of both worlds.

This thesis explores questions around designing interactive sonifications that supplement existing visual systems and focused on developing the ecological theories of sonification first put forward by Vickers [117, 120].

Three different fields of application were explored: Pathology, structural biology, and physiotherapy. In each of the fields, a collaboration with researchers in the field led to the development of a sonification system.

1.3 Research aims and objectives

Questions addressed by this research are:

- How can interactive sonification be used to enhance visual feedback in the fields of pathology, structural biology, and physiotherapy?
- How can an ecological approach to perception be used to design interactive sonifications?

1.3.1 Objectives

The objectives in support of the overall aim are as follows:

1. Study existing tools and their interfaces alone and with domain experts in pathology, structural biology, and physiotherapy
2. Investigate where sonification could be beneficial in the given context
3. Study sonifications in the given context (classification of tissue samples, protein dynamics, exercise devices) and similar contexts or scenarios

4. Discuss and outline potential solutions
5. Perform experiments to test the proposed solutions
6. Formulate design guidelines based on the solutions developed for different contexts above
7. Use design guidelines to come up with solutions
8. Perform and analyse experiments to test the proposed solutions and therefore design guidelines

1.4 Thesis contributions

The thesis introduces a novel interactive system for the sonification of colorectal cancer tissue samples, which derives an auditory glance from scripted interaction with the main interactive sonification.

It describes a sound object-based interactive sonification framework for spatial data. With this framework an interactive sonification for protein dynamics was implemented.

This sonification is implemented as a plug-in for an open-source molecular visualisation software, released as open source software and also consisting of a sonification plug-in enabling the mapping of arbitrary data coupled to the protein's components to pitch.

It introduces a design framework for sonification called *SPBDF* extending research into design theory by Vickers.

Lastly, it introduces a real-time feedback system for a physiotherapeutic exercise device designed using SPBDF.

1.5 Thesis overview

This thesis investigates the design and use of interactive sonification to enhance visual feedback. It describes an interactive sonification of tissue samples, followed by a protein dynamics sonification. Then it describes a design framework and its development, followed by its practical application to the development of SoniFRED (a sonification tool for the auditory display of exercise machine performance — see Chapter 6). A discussion of all topics concludes the thesis. The chapters are organised as follows.

Chapter 2 provides an overview of the field of auditory display and its techniques. It then focusses on applications of sonification to interactive feedback tasks.

In Chapter 3, an interactive sonification system for use by pathologists in the triage stage of colorectal cancer diagnostics is described. The evaluation of different image feature extraction algorithms is outlined and the design of sonification strategies explained. The mapping from data to sound is explained and the use of an *auditory glance* of the image data discussed. The chapter concludes with a description of proposed and executed evaluations and their results.

Chapter 4 presents an interactive sonification plug-in developed for a software package for interactive visualisation and analysis of macromolecular complexes. It outlines the task requirements, as well as the design and implementation of the plug-in. The resulting framework for building different sonification methods in Python and an OSC-controlled sound producing software and sonification methods for the task are described. To conclude, an evaluation of one sonification provided by the plug-in with domain experts and its results are discussed.

Chapter 5 shows the development of a new framework for sonification design based around the idea of the subject position in film studies as first put forward by Vickers [117, 120]. This framework is directly derived from the projects outlined in Chapters 3 and 4. A majority of current sonification design theories are focussed on psychoacoustically optimised mapping or elaborate transformations from data into sound. This chapter explores an alternative conception of design: that working from a frame of reference encompassing a non-expert audience will lead towards sonifications that are more easily understood. A method for the analysis of sonifications according to the criteria put forward in the framework is outlined and then put into practice to evaluate a range of sonifications relevant to the tasks outlined in this thesis.

This framework was put to the test in a collaborative research project presented in Chapter 6. A system for sonified feedback for an exercise device designed for back pain rehabilitation is described. Two different sonifications, one using the framework as basis of their design, one developed by Vickers, are outlined. A large scale study evaluating the sonifications and its results are then discussed.

The following Chapter 7 draws together the results of the evaluations described in Chapters 3–6 and relates them back to the research questions outlined in this introduction.

Finally, Chapter 8 provides an overview of the material presented in the previous chapters. It includes recommendations for future research.

Chapter 2

Sonification

2.1 Introduction

The Sonification Report defines sonification as "[...] *the use of nonspeech audio to convey information.* More specifically, *sonification is the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation*"[69, p. 4]. Hermann proposes a more detailed taxonomy of sonification and auditory display, including definitions for other terminology related to it [54]. Auditory display is the overarching category, including sonification, audification, musification and others.

Basic audio alerts and notifications, e.g., the bell sound of a microwave notifying us of the time being up, form the simplest auditory displays [123]. Musification, on the other hand, is the musical representation of data [27], where sonification is extended by using musical structures to engage the listener. Frysinger [45] provides a brief history of auditory data representation up to the 1980s.

The other most basic form of auditory displays is audification, which Walker and Kramer define as "the direct translation of a data waveform into sound" [123]. Dombois and Eckel extend that definition to include a reference to "making sense of data"[34], which differentiates audification from mere playback of recorded sounds, which is arguably audification according to the first definition. Audification has applications among others in seismology [52] and astrology [33]. As the simplest form of auditory display, audification is often used as initial approach to a dataset or problem, often then superseded by more complex auditory displays.

With the first International Conference on Auditory Display (ICAD) in 1992 [68], a regular meeting of the very diverse community of researchers in auditory display was created. Another important factor in its wider use was also the advancement of computing technology and sound synthesis. This enabled much more complex ways of data processing and mapping, e.g., as employed in Model-based Sonification [57]. Frysinger argues that “[while] visual displays have been used for centuries without a good psychophysical framework, auditory displays are somewhat less intuitive, and the field will therefore benefit greatly from rigorous and accessible research”[45, p. 412].

Minghim and Forrest [77] state that a well designed sonification complementing a visual task will provide sound that is “non-intrusive, and should not affect, but instead improve, visual interpretation”. They also argue that sound can speed up the interpretation and interaction by relieving the user of visual work load. In working with data, they argue that getting the gist of portions of data could reduce the number of interactions the user has to make to get an understanding of the data (e.g., viewpoint changes).

There are different sonification techniques, some will be described below or where relevant. The *Sonification Handbook* provides an extensive overview of techniques in part III [60, p. 299], and more techniques are popping up regularly, e.g., in 2018 Hermann [55] introduced a new sonification technique called Wave Space Sonification (WSS), where a scalar field is scanned along a path driven by the data, which is supposed to enable controlling the definition of the auditory representation for the area of interest, as well as complex sound variations in reaction to changes in the overall pattern of the data.

More in-depth background for the projects can be found in their respective chapters. Here a general overview of common themes is provided.

2.2 Interactive Sonification

Hermann and Hunt defined interactive sonification as “[...] the discipline of data exploration by interactively manipulating the data’s transformation into sound” [56, p. 274]. Whereas some sonifications don’t demand interaction, musical instruments in general do. They can provide inspiration about possible modes of interaction. Whereas auditory display is a wide field with many subfields, interaction exists on a different axis, as overlap between the fields of auditory display and human computer interaction (HCI).

Choi [24] presents an “extensible sonification model” for interactive sonification. She introduces *Heuristic Listening* as an “informal model of a listener’s clinical attention to data sonification through multisensory interaction in a context of structured inquiry”. Outcomes from case study implementations demonstrate the framework’s performance.

Plazak *et al* [88] used distance sonification to provide audiovisual feedback of a surgical probe in image-guided neurosurgery. The results of a user study gave credence to the hypothesis that the proposed system may result in greater accuracy and increase the precision of individual surgical tasks.

2.2.1 Sonifications providing an overview

Franklin and Roberts proposed a novel path-based model [43] for sonification of 2D or 3D data. They came up with terminology to efficiently describe tour-based methods, where the sequencing of the data happens along a path through the data space. They talk about path, span, span angle and span envelope, where path is a series of path points (which can be in completely random locations), the span is the neighbourhood associated with the path point (which can be irregularly shaped), the span angle is the angle of interception between the span and the path, and the span envelope is the size of the span dependent on the time and the position of the path.

Ó Maidín and Fernström developed the *Sonic Browser*, a browser to navigate music or sound databases using sonic spatialisation [71]. The contents of the database can be represented as abstract shapes in a 2-dimensional viewer, their location and other attributes determined by properties of the melody, e.g., melodic distance. To ease the browsing and help the user to explore the space visually and aurally, they introduced a new kind of cursor function, which creates an aura around the cursor (a circle with user-determined diameter at the cursor), playing all melodies within the aura concurrently using spatialised sound. Though no formal tests were conducted, informal tests showed a high degree of satisfaction.

Hermann introduced the concept of Auditory Information Buckets (AIB) in his thesis [53]. A bucket is “filled” with data until its threshold is reached and it “empties” itself, causing an AIB sound event, in which the aggregated data is sonified. Depending on the threshold and the chosen sonification, the resulting sound event is more or less frequent or complex. AIBs provide a different kind of auditory zoom, giving the user a choice over the resolution in which information is summarised. In his example, AIBs are used to summarise the characteristics of the modes of a Markov chain Monte Carlo model-based sonification to give the user a clue about the shape of the attraction basin of a mode.

In Zhao’s dissertation, she uses the word “gist” instead of “overview”, “to differentiate from ‘seeing’ by eyes”. Due to the capacity of the human short-term memory, she advises to keep the duration of a gist under 10 seconds. She also suggests synchronising the auditory display with other modalities to enable low-vision users to combine information from multiple channels and ease collaboration with sighted users [133].

2.3 Sonification as augmentation of visual displays

Wei and Ibrahim [124] display complex data about dengue fever by using parameter mapping. They call their sonification an ‘ambient sound framework’. In an evaluation they found an improvement in the level of data interpretation when using their framework.

In their TaDa method, Barrass *et al* [11] outlined a design process that includes task analysis, sound examples, rules for design, and interactive sound design tools. Multiple scenarios were implemented and provided evidence that sounds can provide information hard to obtain visually, therefore improving a display’s directness and usefulness.

2.4 Sonification and listening

The field of psychology of hearing, or psychoacoustics, deals with how humans perceive various sounds, particularly our psychological and physiological responses. When we hear, we don’t just hear the thing we focus on, as in contrast to our visual perceptual apparatus we can only physically focus our hearing in a very limited way. Instead, our brain manages to extract relevant acoustic streams out of a mixture of background and relevant foreground [18]. This is also known as the “cocktail-party problem”. More on the topic of psychoacoustics can be found in [20].

Ferguson and Brewster [42] evaluated sonification mappings that are based on psychoacoustic sensations in two experiments where subjects were required to detect the level of focus of an astronomical image through sonification. They found that noise and combined roughness and noise came close to the visual condition.

Volmar [122] argues in an historic overview of auditory knowledge construction that in the case of auscultation, the proponents did not prefer the auditory channel, but only the ear was capable of receiving signals from within the human body. The process of knowledge development around the “mapping” of sounds to lesions could also only developed over thousands of autopsies. He concludes in the last section of the online paper that “sonification researchers should pay more attention to how a sonic form of epistemic practices, facts, and arguments may contribute to specific research processes” instead of blaming ocular centrism for lack of interest in sonification, therefore focussing not on mere mapping but producing “sonic facts for sound arguments”, i.e., providing a mapping that is reproducible and related to the practical needs of the specific field.

Supper [111] argues that historical examples of sonification are often misinterpreted by sonification researchers, seeing them as examples of using sound to represent data instead of displays that happen to use sound as in the case of the Geiger counter. The debate between data and information is still ongoing, as in many cases it is not clear if sonification is sonifying data or information. Worrall's thesis [129] tries to give an overview of that.

Supper goes on to say that the sonification community currently is too focussed on building tools, partly because historically Kramer was a composer interested in making data music and tools derived from computer music tools are not very usable for domain scientists, therefore domain scientists soon lose interest or never are interested in using sonification, especially not the tools developed by sonification scientists.

Regarding big data and sonification she postulates that information does not just "pop out" of data if sonified and therefore problematic promises and popular misunderstandings about what sonification can do emerge.

2.4.1 Sonification and embodiment

Human auditory perception is a complex interaction between the perceiver, the environment and the sound attributes. Perception in an ecological sense is to be understood as a relationship between information available in the environment and the capacities, sensitivities and interests of a perceiver [25, p. 91].

There are several extant sonification design frameworks, typically approaching sonification from the point of view of the data types involved (e.g., [30]) or information requirements (e.g., [11]) but there remains a lack of heuristics based upon the more fundamental relationships listeners have with sound.

Sanderson and Watson [99, 100] extended ecological interface design (EID) to include the design of auditory displays. EID is a systematic approach aiming for understanding of cognitive work and exploiting perception to make cognitive work easier where possible. Their analysis points out that the main design problem in process monitoring sonifications is how to get the user's attention only when the data requires it. They propose to extend Cognitive Work Analysis (CWA) with an "attentional mapping" stage to accommodate that requirement.

Davies and Burns [28] also pointed out the usefulness and application of CWA and EID to the design of auditory displays. CWA consists of different modules to determine user needs and display requirements: Work domain analysis (WDA), control task analysis, strategies analysis, social and organisational analysis, and worker competency analysis. An explanation and examples of each analysis method applied to auditory displays is given in the paper.

They criticise auditory display designers who, instead of developing a product that meets the users' needs, look for ways to better train them on their, apparently not very easy to understand, system. They emphasise that, in the case of auditory display, the environment in which it will be used must be thoroughly analysed and considered. Furthermore, due to the fleetingness of the display, a functionality to revisit previously presented information is critical.

Roddy and Furlong [95] describe sonification as very positivistic research culture, derived from computer music and neuroscience viewing the mind as a computer. Aesthetics comes second after the translation from data to sound is established. They say that parameter mapping sonification is especially prone to making positivistic mappings, disregarding that acoustical parameters aren't perceived linearly. Auditory icons are therefore better because they actually represent something, leading them to the conclusion that parametrised auditory icons are the way forward.

Roddy and Furlong present a twin-pan schema [94], where they represent two values (x and $-x$) by distal cues. When one value goes up, the other goes down, like in weighing scales. Preliminary usability testing pointed towards the display working in that it was communicating embodied meaning about the relationship between the two values [93]. The study also tested a sonification of rain probability described in [93]. A SOURCE-PATH-GOAL model was used to represent the days of the week, fading in and out days. A CENTRE-PERIPHERY schema was used to represent the probability of rain: auditory icons of rainfall moved away or towards the listener according to probability. They tried to delineate days with a "beep", but study participants didn't understand that very well. The user study showed an "embodied" understanding, i.e., no linguistic or numerical knowledge of rain probability, but an "intuitive", or "felt" understanding. This shows its usefulness only for measurements of qualitative data types, because it enables intuitive understanding, but not when abstract numerical values are concerned.

In his doctoral thesis, Roddy [92, p. 60] argues for complex dimensions of mapping, as pitch, amplitude, timbre are too simple dimensions to represent embodied sounds. For example precipitation data can be represented by falling rain, where heavier rain is not easily expressed in one dimension like pitch or amplitude, as it is a complex interplay between dimensions.

2.5 Summary

This chapter provided an introduction into sonification and auditory display. It has also included an introduction to interactive sonification, the main technique used in the projects used in the following chapters. Specific literature related to the research areas of the projects will be reviewed in the corresponding chapters.

Chapter 3

Image Sonification

3.1 Introduction

The work on image sonification described in this chapter happened in conjunction with another research project at Northumbria around the classification of colorectal cancer biopsy tissue samples conducted by Bouridane and Peyret in collaboration with a hospital in Qatar (see [87]). While their work looked at automatic classification, the system described here supports pathologists in their work.

Pathologists are confronted with high workloads and high pressure in the analysis of biopsy tissue samples. Automatic classification systems for various types of samples are in development or in some areas already in use, but in many cases (e.g., in colorectal cancer diagnosis), do not work with the accuracy demanded by a potentially severe disease.

By sonifying extracted image features the system described in this chapter enables pathologists to analyse samples in an audiovisual manner. We propose this approach will help to improve the accuracy of analyses, while at the same time improving the working environment by introducing more channels for feedback, giving the pathologists an enriched view of the samples in question.

This chapter starts with an outline of existing methods for medical image sonification, followed by a description of the system developed in the course of this thesis and its evaluation. The source code for this chapter can be found online [1].

---

[1] See `https://github.com/mortuosplango/bettison`.

3.2 Image sonification

Image sonification is the translation of images into sound. Applications include prosthetic equipment for blind people, artistic purposes and biomedical image analysis. Sarkar *et al.* [101] provide a good overview of the field.

In one of the earliest reported image sonifications, Meijer [75] proposes an experimental system for auditory image representation. A greyscale image of up to 64 × 64 pixels is sonified by translating grey values into amplitudes. The image is scanned from left to right in about one second and every pixel in a column translated into the amplitude of a sine wave. The vertical position of the pixel determines the frequency — the lowest line is represented by the sine wave with the lowest frequency. The frequencies are spread linearly or exponentially between 500 Hz and 5 kHz. The image-to-sound mapping is thus isomorphic and according to the author ensures the preservation of visual information.

Yeo and Berger [130] take a similarly reversible approach in their work on raster scanning. Raster scanning means recording elements of a display image by sweeping the screen in a line-by-line manner. This could be seen as a special form of scanned synthesis as described by Boulanger *et al* [17] as a sound synthesis technique which “scans” a closed path in a data space periodically to create a sound. Yeo and Berger conclude that raster sonification creates a sound which “contains the ‘feeling’ of the original image” [130, p. 8]. In future work they wanted to explore its use for the sonification of a painting to provide auditory clues for its visual patterns.

Yoshida *et al* [132] present *EdgeSonic*, an interactive sonification to aid visually impaired users to recognise objects in an image by sonifying local edge gradients and the distance to the edge of the object from the user’s finger on a touch screen. In preliminary experiments they showed that this combination is effective for understanding basic line drawings, improved further by providing user training.

Banf and Blanz [9] propose a modular computer vision sonification model for the visually impaired. They aimed to provide visually impaired people with means to explore images interactively. They chose a touch screen instead of a mouse for the interaction device as it helps with orientation by providing a fixed x- and y-axis and a direct analogue to touching a relief. In their model they also account for colour by mapping complementary colours to *Complementary Instruments*, i.e., instruments complementary in timbre (choir (red), organ (green), strings (blue), bagpipe (yellow), and flute (white, black, grey)). Grey is treated as separate instrument. Rough textures are represented as vibrato. “Orientation maps” represent dominant structures within the image and are represented by hum-like sounds (horizontal or vertical structure) or percussion sounds (diagonal structure) an octave lower than the pixels. All sonification is realised with MIDI instruments.

In a follow-up paper, Banf and Blanz [10] also use higher-level machine learning-derived features in addition to the low-level display described above. Sound synthesis is used instead of MIDI for cross-platform reasons and to obtain more possibilities to directly sonify colour properties. Furthermore, speech synthesis is used to identify found objects so the user doesn't miss them during exploration. Auditory icons for objects recognised by the software are used ("meow" for cats etc.). They are looped while the user is in the region, but pass away in 3D sound during that time. Roughness of natural regions are sonified using brown noise as an acoustical representation. A user study with one congenital blind person and another with two congenital blind people was conducted. All participants reported the system being very intuitive, as well as easy to understand and quick to learn. The authors note though that, as it is a very general tool, special-purpose tools are likely to be more adequate for many every-day tasks, e.g., navigation.

In Martins' and Rangayyan's work [72, 73], an image is translated into the Radon domain, a process comparable to inverse tomographic image reconstruction. Different sonifications are used depending if the texture is regular or not [72].

3.3 Sonification of geometric data

Axen and Choi [6] traversed the paths of simplicial complexes and sonified the result using additive synthesis. They found that attributes of multidimensional objects (e.g., holes) could be more easily detected in the sonification than in the visual display. In further research they developed "auditory zooming" [7], where the sonification depends on the user's visual position in relation to the complex: the further the user is from the complex, the more data values will be mapped to one sound. When zooming in, the sonification gets more complex.

The Satin system represents geometrical information of virtual objects [103].[2] It is meant to be a multimodal and multisensory system, which enables the evaluation of virtual shapes of objects based on a haptic interface, stereoscopic visualization and sound. A shape is not directly sonified, but instead the curvature of the shape is translated into a frequency. When the user touches the shape on the haptic interface, a sound is produced. Discontinuities in the curvature are sonified by frequency modulation between the frequency before and after the discontinuity according to the distance of the user's finger to the discontinuity.

Zhao *et al* [134] developed a system to make choropleth maps accessible to visually-impaired users. The resulting application, *iSonic*, works according to her adaptation of Shneiderman's Information Seeking Mantra [104] to the aural realm. Zhao [133] proposes the following Audio Information Seeking Actions (AISA's):

---

[2] See also `http://www.satin-project.eu/`.

- Gist
- Navigate, Situate and Search
- Filter
- Details-on-demand and Adjust Information Level
- Select
- Brush

In the case of iSonic, first a gist is provided by playing all visible data points on the map in sequence according to a predetermined sweeping path, then the user can navigate the map with a keyboard or touch screen to find the information they want. They can filter the displayed states by different criteria, demand details of the state their cursor is situated in, which will be rendered using speech synthesis, play only selected states and switch between different views (which is referred to as brushing in visualisations) showing the same information (e.g., between a map and a table view with the same states selected).

In user studies participants had to determine a pattern in the displayed geo-referenced data by using iSonic. High accuracy of map discrimination and high levels of satisfaction and understanding of the tools and tasks are reported [32].

3.4 Auditory display of medical data

Hermann *et al* [58] experimented with the sonification of multichannel image data obtained from cells using multi-parameter fluorescence microscopy. They sonified the markers of the cells using short musical phrases to make it easier to memorise and distinguish cells. No studies were reported. They remark that the interpretation of sonifications has to be learned, therefore establishing standard sonification procedures (which still allow for extensions later) would benefit biomedical expert users to improve their interpretation proficiency.

Building on this research, a multi-parameter image analysis system for biomedical research was proposed using both visualisation and sonification [82]. The auditory display uses artificial words produced by concatenating diphone sequences to take into account musically untrained listeners and high-developed perceptual skills in most listeners. Each diphone represents a biomedical protein marker, resulting in protein patterns represented by words, where identical patterns are easily identifiable. They note the big unrealised potential in unused lingual variations. In a study, 60 non-expert participants had to identify if cells were identical to a reference cell or not based on the visualisation, the interactive sonification or both. They found no significant difference between the variances of the groups, meaning that the combination approach didn't improve performance, and the group using only sonification was the slowest (with evaluation times varying by just 15%). Nattkemper [81] remarks that the contrast to the successes reported in other works about multimodal data representation isn't surprising considering the possibility space of auditory display design in combination with a visual display and the small number of documented applications.

Jovanov *et al* [64] worked on the perceptualisation of biomedical data in the form of EEG. They use natural sounds (e.g., of a creek or bees) to provide a pleasant basis for sonification. The data stream is then used to modulate the pitch, volume or balance of the sound pattern. They argue that supporting user-specific templates is highly advisable, as the "perception of audio-visual patterns is very personal". They conclude that "[sonification] improved the ability to assess genuine dynamics of brain electrical activity and to perceive inherent spatio-temporal patterns of brain electrical activity".

Hermann *et al* [59] worked on vocal sonification of pathologic EEG features. They used indirect continuous parameter mapping sonification as the mapping used data-driven features to control synthesis parameters, instead of directly mapping the data to control parameters. No formal evaluation was conducted.

Cassidy and Berger built a sonification based on a model of the human vocal tract to analyse hyperspectral colon tissue images [21, 22]. By mapping the data to the radii of the tube sections in the model, the output sound would vary. They argue that the use of human vowel sounds creates an intuitive and easily-learned representation, but no formal evaluation was conducted. Yeo *et al* [131] note on the same project that they have "demonstrated the ability to distinguish between benign and malignant cell structures with auditory display", but don't provide any evidence.

In the first report on sonification in the analysis of three-dimensional medical image data, Rogińska *et al* [96] use immersive sonification to augment brain scan data. The user is inside an immersive 3D environment with 3D sound where they can navigate through the data. No formal studies have yet been reported.

In the field of cervical cancer diagnosis, Edwards, Hines and Hunt [38] worked on the segmentation of biological cell images for sonification. The microscopic examination of cervical cancer is the starting point to improve diagnostic accuracy by translating image features into a complementary auditory form. They decided against using a simple alarm sound if cancer was detected — the system is only meant to present information and leave the interpretation to the screener. The planned sonification was supposed to show how well a given cell matches typical normal or abnormal cells, but no auditory display existed at this point in their research.

These two dimensions — confidence and abnormality — were abandoned in the course of the project as reported in Edwards *et al* [39], maybe for reasons of simplicity. The researchers decided on a “badness” scale of ten steps, which formed a soundscape comprised of samples of cows mooing to indicate “normality” to a woman screaming on the other end of the scale. Participants in the study found the mapping to be baffling and not universal, as the ordering of the different sounds representing the “badness” wasn’t intuitive.

Following this study, screening cytologists of Leeds NHS Trust were asked to fill in a questionnaire about their listening habits and were asked to give feedback on several of the sounds and which cell images they would relate them to. The results were that a majority wanted to hear an alarm-like sound when an abnormal cell was found, but would also appreciate a constant quiet sound to “show the system is working”. The participants were also opposed to the use of real-world sounds and sceptical about the use of “musical” sounds. To help with finding suitable sounds, the researchers decided to conduct an online survey into sounds and their association with normalcy and badness. No results have been published yet. They concluded that the reliability of the rating and the “badness” are important parameters to sonify, as well as stressing the importance of selecting the right sounds.

Kagawa *et al.* [65, 66] tried to improve CT and MRT image diagnosis with sonification. They argue that imaging supporting methods, such as emphasised images, strain radiologists visually. Therefore they want to lighten the burden of the visual modality by using sounds. They used beeps, background music and earcons. In a small study with 10 engineering student participants they found that sound effects can contribute to exact diagnosis, but didn’t find a shortening in time for the operation of the user interface with sonification enabled.

Assistive systems can help narrow down the area to examine by finding regions of interests (ROIs) in the tissue sample. Peyret *et al.* [87] are looking into ways to use multispectral imagery to detect ROIs and classify them. That still leaves the pathologist with the final diagnosis and the problem of fatigue and heavy workload. How can sonification support the pathologist in this task?

The design of the software was guided by the working habits of pathologists and previous research into the auditory display of medical data as described in the literature review. By providing a multimodal data representation, bias due to one representation dominating as described in Ben-Tal *et al* [15] was hoped to be mitigated.

3.5 Overview

In this section an overview of the data available to the software is given, along with a description of the design process and rationales behind design decisions taken. The resulting plan and its components are then outlined.

3.5.1 The data

Every image has 4,194,304 (512 × 512 × 16) 8-bit data points. Every pixel in the image is a 16-dimensional vector. The data is spatial, therefore there is no time axis that could be used to order the data points in time.

3.5.2 Rationale behind the design

Sonification is inherently time-based. The ear can only discern a limited number of parallel events. Directly mapping the image data to sound would lead to sonic averaging, i.e., details would get lost due to the mass of sound events.

Reducing the resolution of the image before sonification would just shift that loss of information to the image processing step.

As the pathologist will also see the image, sonifying every data point is unnecessary. The project is not concerned with replacing the visual display, but supporting it.

Only sonifying the dimensions not being used in the visual display, e.g., if 3 of the 16 bands are displayed, leads to a slight reduction of data points to 3,407,872 (512 × 512 × (16-3)). This of course is problematic at best as most bands in the image will be combined in those three channels.

A common strategy is to introduce a time axis into our dataset. A simple way to do this is to designate one of the axes of the image as time, most often the x-axis, “sweeping” the image horizontally and using the data of the other axes as basis for the sonification. This would leave us with 8,192 (512 * 16) data points per time step, or 6,656 if we disregard three of the dimensions.

In order to avoid information loss, the time parameter can be stretched and display the data points spread over a longer time frame, e.g., by introducing a vertical sub-sweep inside the horizontal sweep. In this case, more information can be displayed, but the user experience can suffer as the time to display each image increases without regard to how ambiguous the image actually is and how long the pathologist would normally take to analyse it.

3.5.3 Getting rid of the time axis

Interactive sonification is a method to give users the possibility to interact with the sonification via an interface. Instead of just hitting “play” on a dataset, users can query sections of it and only those sections will be sonified (see Section 2.2 for more information on interactive sonification).

Considering the working habits of pathologists —- focusing their attention on specific sections of the image in turn, ignoring some sections of the image completely, with a clear objective in mind —- a sweep of the dataset will necessarily display a lot of unnecessary, and potentially distracting information.

With the interactive stage, the problems of introducing a time axis into spatial data are avoided by giving the users the possibility to create their own order of events suited to their needs.

3.5.4 How to not miss the obvious?

The downside of using interactive sonification is that some aspects of the dataset may remain undiscovered, as users only concentrate on sections that seem interesting to them judged by their visual appearance. In the use case described here, the system should support pathologists make more accurate decisions faster by assisting them in their task of image classification.

A trade-off between user-friendly non-intrusive interaction and an assistive auditory display has to be made to maintain a consistent mapping strategy between the two stages.

Therefore, the system can not just use one axis as time axis, as described above, as that would prevent the utilisation of the same sound for both stages. Therefore, a path is overlaid on the image, which triggers a pseudo-interaction and therefore sonifies a subset of data points spread evenly over the image in the course of the duration of the traversal of the path. Zhao [133] advises keeping the duration under 10 seconds due to the capacity of the human short-term memory.

The path is displayed to the user while the overview is played to enable them to map the sound to the location in the image more easily.

3.5.5 Two-stage sonification

As outlined above, both an overview and an interactive stage is called for to avoid too much of a focus on visual features and wasting time on playing parts of the image that are irrelevant. The auditory display is therefore split into these two stages, a short overview and an interactive stage:

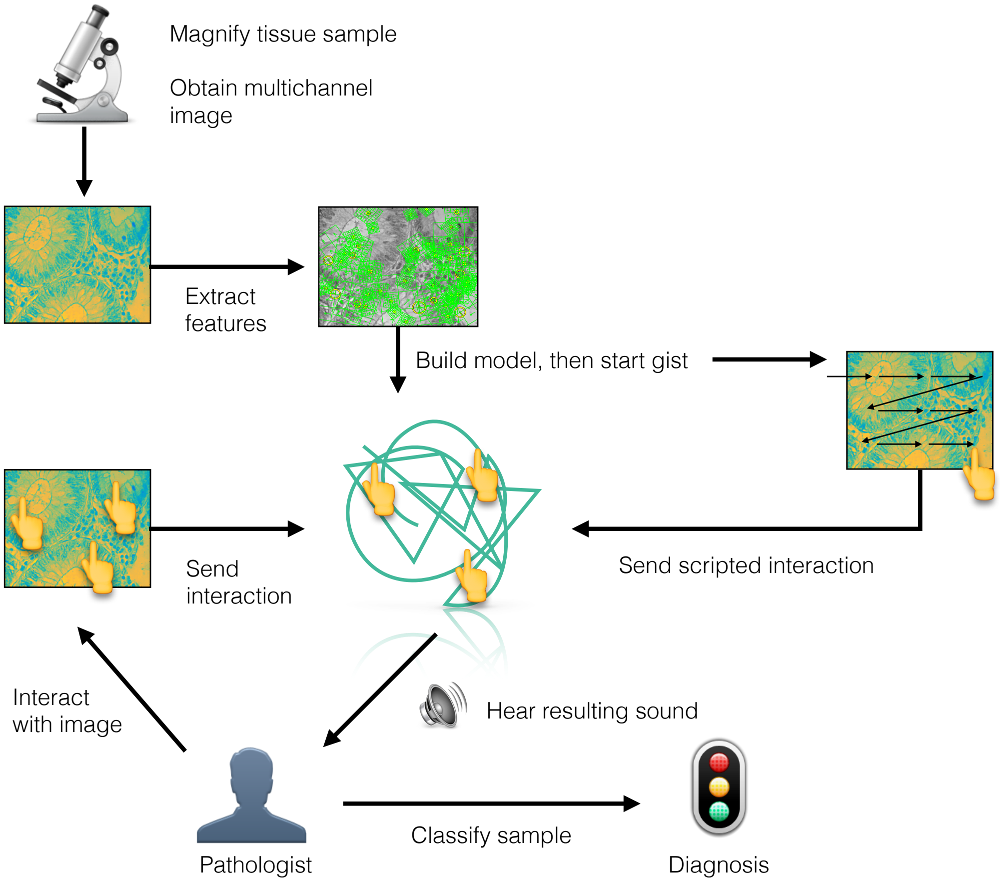


Fig. 3.1 An overview of the software: A multichannel image (an image with more than one colour channel — 16 in this case, but an RGB image can be seen as a multichannel image with 3 channels) is produced by digitising a tissue sample, and relevant features for cancer detection extracted. These features are mapped to sound parameters in two-dimensional space (the model), which the user can interact with via a touch screen interface. After building the model, a scripted interaction on the image is executed, resulting in a short overview sonification (gist) of the image being played. After this gist, the user can interact with the image, which leads to sounds being produced according to the model, leading them to classify the sample.

1. In the overview, the data is displayed in under 10 seconds by providing a scripted sequence of information requests.
2. In the interactive stage, the users can request information about arbitrary regions of the image, resulting in a sonification of that section of the data.

With the interactive stage, the problems of introducing a time axis into spatial data are avoided by giving the users the possibility to create their own order of events suited to their needs, whereas the overview is still there to showcase the information contained in the sonic domain.

The auditory display is consistent between overview and interactive stage, enabling the users to enquire about details they heard in the overview.

The overview is triggered every time a new image is displayed. The users can also trigger it again later if they wish to hear it again.

3.5.6 How to extract less abstract data points out of the dataset

To make the displayed data more helpful, the image is analysed and features extracted. In this way, the given data set is further compressed into a more helpful format for humans. Especially in the case of pixel data, a direct mapping to sound is not very useful most of the time.

Image feature extraction and classification is a very active research field. Multivariate image mining, i.e. feature extraction from multivariate images, of biomedical samples is reviewed by Herold *et al* [61].

Multispectral images are a subtype of multivariate images, where a sample is imaged for multiple wavelengths (i.e., an RGB colour image can be seen as a multispectral image stack with three bands). Other subtypes include multimodal images, where images obtained from multiple modes (e.g., images obtained from via optical and electron microscopy) are combined. Multivariate images (MVI) lead to increased information being represented, but they also lead to problems with how to efficiently analyse that data [61].

Different sonification models to represent the data and different feature subsets will be examined. Finding a meaningful mapping between the image data, the extracted features, conclusions gained from these features and sound is non-trivial.

Local Binary Patterns

Local binary patterns (LBP) describe the role of a pixel in its neighbourhood by taking the grey-level of the pixel and comparing it to its neighbours in a defined radius around it (minimum 3x3) — if the neighbouring pixel is lighter, that pixel gets assigned a 1, if it is darker, it gets assigned a 0 in the resulting binary string. Consolidating rotational patterns, one is able to categorise a pixel as section of an edge, a flat bit, or a corner (see [84] for more details).

Scale-Invariant Feature Transform

Scale-Invariant Feature Transform (SIFT), as described by Lowe [70], enables the extraction of feature vectors called SIFT keys from an image. Those features are scale-invariant, i.e., they can be used to detect an object described by them in a different image when the object is scaled, rotated, 3D projected, blurred or even partially occluded. It works by translating the image into scale space and then extracting keys that describe gradients and their direction. Those features “share a number of properties in common with the responses of neurons in inferior temporal (IT) cortex in primate vision” [70, p. 1].

Segmentation of cells

Roula *et al* [98] describe an evolutionary snake algorithm for the segmentation of nuclei in histopathological images. Segmentation of the nucleus is useful as the nuclear chromatin patterns provide important information on the state of the cell. Other segmentation methods rely on evolutionary approaches which are computationally complex. By replacing standard mutation with an “oriented stochastic mutation process”, they achieve fast computation times suitable for real-time quantitative pathology applications.

Chaddad *et al* [23] improved colon cancer cells detection based on Haralick’s features [50] on segmented histopathological images. They used a snake algorithm to facilitate rapid segmentation of cancer cells. By using just three Haralick’s coefficients (Contrast, Entropy, Correlation) extracted from multispectral images, they were able to find variations that led to successful classification in a probabilistic neural network.

Rathore *et al* [90] developed a segmentation method for colon biopsy images using local binary and ternary patterns, and Haralick’s features, which are then reduced using genetic algorithms and F-Score. These features are input into to random forest, rotation forest, and rotation boost classifiers to segregate the image into regions of normal, malignant, and connecting tissue.

Cohen *et al* present [26] a method to detect and extract the shape of the cell and determine malignancy from it, as more malignant cells have a less round shape.

Homology-based region of interest (ROI) detection

Nakane *et al* [80] describe a method to detect regions of interest (ROIs) in images using methods from computational homology, in this case the Betti-number of a region. Their research is in its early stages, but the use of homology, a concept employed in branches both of algebra and topology, provides a way to determine ROIs based on the degree of contact in a region, as cancer cells lack contact inhibition. They report that the system still reports many false positives, but successfully detects ROIs and has potential to be further developed.

Summary

The prototype described in section 3.6 makes use of local binary patterns. Scale-Invariant Feature Transforms were not used in this thesis, but present an important direction of image feature extraction. The segmentation of cells according to the methods described here was attempted but incomplete or missing descriptions of the algorithms used prevented their implementation. Homology-based detection is used in the second prototype described below.

3.5.7 Integral components of the software

Feature extraction When loading an image, features are extracted according to the needs of the sonification employed. This happens before the image is displayed if the features chosen can be extracted with minimal delay. If not, the image should be displayed first and the extraction happen in the background. Whenever possible analyses of possible images to be displayed could happen in the background while other images are on screen.

Interface Tissue samples are displayed on a big screen. Users are able to interact with them by using a secondary touch screen displaying the same image. Regions of interest (ROIs) extracted by analysing image features areas marked and the user is able to zoom according to the zoom levels provided by the microscope.

They are presented with an auditory overview of the currently displayed image on change of viewpoint.

They are able to explore the displayed image, probing arbitrary subregions by stroking or tapping the displayed image.

3.6 Prototype: Data Sonogram with local binary pattern features

A first prototype was implemented using Python for feature extraction as it has great support for scientific computing and many extraction methods are readily available as libraries. It used Spectral Python (SPy) [16] to load the multispectral images and perform principal component analysis on them. Then the images are reduced to the first five principal components and Scipy [63] is used to perform local binary pattern (LBP) analysis on each channel.

The extracted features are then sent to SuperCollider for the sound synthesis and SPy is used to reduce the original image to RGB. This image is then sent to Max/MSP for display.

SuperCollider [126] is used to sonify the data. It provides good performance and this prototype is based on the SuperCollider implementation of a data sonogram by Till Bovermann[3], where a travelling wave is mapped from the coordinates received and a granular sound-event triggered for every datapoint it hits. The amplitude is mapped to the distance of the travelling wave. Depending on the value of the LBP at the point and if it denotes an edge, a flat bit or a corner, the sound parameters of the sound event are varied, resp. flat bits are ignored.

Max/MSP[4] and Mira[5] is used to display the image on the computer and the touch screen (see Figure 3.2). Max with Mira enables rapid development of touch-screen interfaces as Mira can just "mirror" a Max interface to an iPad, removing the need to develop dedicated apps and turning any iPad into a touch screen interface. When the user interacts with the image, the coordinates are sent to SuperCollider to be sonified. When a new image is selected or zooming into a region is requested, the feature extraction backend is notified.

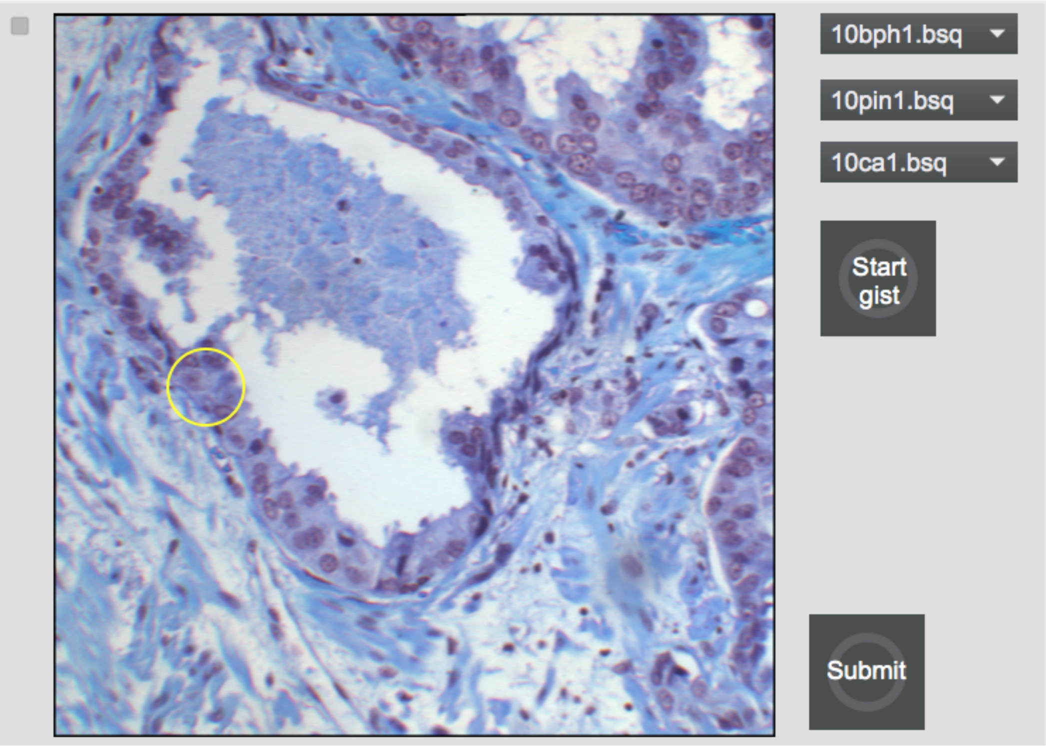


Fig. 3.2 The interface.

---

[3] http://www.techfak.uni-bielefeld.de/~tboverma/sc/tgz/MBS_Sonogram.tgz.

[4] `https://cycling74.com/products/max/`.

[5] `https://cycling74.com/products/mira/`.

This prototype essentially only displayed features already visible in the image to a non-visually impaired person. It was therefore decided to look for more higher level features that are region-based, in contrast to many of the extraction techniques surveyed above who give a score for the whole image instead.

3.7 Prototype: Interactive parameter mapping sonification of regions of potential interest

Another prototype was implemented, using the same interface and architecture as in Section 3.6.

A promising technique to extract the shape of cells [26] lacked essential information to be reimplemented and also would have taken the same direction as before, just replicating the visual cortex instead of augmenting the information available to the pathologist. Therefore extracting Betti numbers from segments of the image to determine a homological score of potential interest [80] was chosen, as it provided both feature extraction on an image-region level as well as relying on both more abstract and more easily implementable features.

Spectral Python (SPy) [16] is again used to load the multispectral images. Then each channel of the image is turned into a binary image using a threshold calculated with Otsu's method [2]. The feature extraction method described [80] uses $14x14$ fields, but as their images have roughly twice the resolution, $7x7$ fields are used in this case. For each field the Betti number is calculated using ChomP [6], a software package described in Mischaikow *et al* [78].

The extracted features are then sent to SuperCollider for the sound synthesis and SPy is used to reduce the original image to RGB. This image is then sent to Max/MSP for display.

The Betti number for each layer of each segment is multiplied with 110 Hz if greater than a threshold of 6 as only very high Betti numbers (> 7) are considered interesting [79], otherwise it defaults to 110 Hz, the base frequency of this tone. These 16 frequencies are used in an additive synthesis process using sine waves with a slight ($\pm 1\%$) frequency drift to liven up the resulting tone.

Therefore each spectral layer contributes its potential of being a region of interest to a tone, where a tone with less lower partials can be interpreted as potentially more interesting region.

The resulting tones are then low-pass filtered with a second-order filter with a cut-off frequency of 400 Hz. This is to prevent listener fatigue as all regions are played continuously.

---

[6] Available at `http://chomp.rutgers.edu/Projects/Computational_Homology/OriginalCHomP/software/`

When the user interacts with the image as described above or the glance is triggered, the filter cut-off frequency is lifted to 18000 Hz for a short amount of time revealing the full spectrum at this region, whereas otherwise the regions blend into a background sound scape.

This background soundscape evokes bell-like associations when Betti numbers are lower or less spread-out, sounding more metallic the more unique higher partials are in the mix.[7]

3.8 Evaluation

Due to budgetary constraints no user study could be conducted as part of the other project. A pathologist at a local hospital agreed to discuss the project though and provide feedback. She provided a tour of the lab and demonstrated her work environment and the way she assesses colorectal cancer tissue samples.

After demonstrating the current prototype, she provided very interesting feedback and it became clear that our prototype and the needs of pathologists only partly overlapped, e.g., whereas the prototype worked with a static image, she used the microscope to hone in on different parts of a larger tissue sample, defeating the use case of the auditory glance as that would mean having to scan the whole image beforehand. All in all she worked much faster than the system would enable her — especially with the additional delay of listening for clues. If the preliminary survey had happened or a closer collaboration with a pathologist in earlier stages of the project would have been feasible, this could have been prevented.

3.9 Possible other sonification strategies

Taking inspiration from sonocytology, a method to record the movement of cell walls using atomic force microscope, which enables some conclusions about the state of the cell [97], one approach could be to extract the shapes of the cells and use these to build a physical modelling-based interactive sonification, where cells can be touched to listen to them, in a way artificially creating an sonocytological audification.

Another angle could be focused on the translation of higher level features into a model-based sonification context. For example, a network of relationships between cells could provide a model of resonators that the user could interact with.

A dimensional reduction from the 16 dimensions of the multispectral image to a more handleable 5 could be done without considerable loss of information, then it would be possible to map those 5 dimensions onto physical modelling synthesis system.

---

[7] See `https://youtu.be/5BaNbXZt78M` for a video demonstrating this sonification.

3.10 Conclusion

This chapter outlined the work done towards interactive sonification of image features. Two prototypes were developed and a pathologist gave valuable feedback for the second prototype, which pointed towards a wrong approach in terms of the HCI element of the project. The application cognitive work analysis (CWA) as discussed by Davies and Burns [28] to this project could have lead to a different outcome if access to domain experts is possible.

The outlined prototype using Betti numbers used filtering to put parts of the sonification in the background and highlight parts that the user interacts with. Along with using auditory glances that are generated by scripted interaction this leads to a sonification that is at the same time static and consistent, therefore corresponding to the state of the data it is based on as well as signalling that the sonification still works (as was remarked as desirable by participants in the study conducted by Edwards *et al* [39]), but at the same time being explorable by the user or auto-explored by the glance.

These ideas around a static data set being translated into a nearly-static soundscape with interaction highlighting the parts the user is interested in while at the same time providing an overview of the whole dataset in the background are further developed in the project described in Chapter 4.

Because of budgetary constraints and with a genuine user study in Qatar unfeasible, other use cases for interactive sonification were explored, which are elaborated on in the following chapters.

Chapter 4

Protein Dynamics Sonification

4.1 Introduction

This chapter outlines a system for the sonification of protein dynamics. The chapter is based on and extends a paper previously published by Ballweg *et al* [8].

Molecular graphics and visualisation has a long tradition in analysing and interpreting computational chemistry and structural biology data. High demand for powerful software, rendering the structural and dynamic attributes of macromolecular datasets in accurate yet visually elegant and intuitive ways, has resulted in the development of several molecular graphics packages and platforms, open-source (VMD, UCSF Chimera, PyMol) as well as commercial (MOE, Schrodinger).

UCSF Chimera is one of the leading software packages for interactive visualisation and analysis of macromolecular complexes (protein-protein, protein-ligand, DNA), their structure and dynamics [86]. It is open-source licensed for academic use, has a long history, a considerable user base, and is constantly in active development. As it provides a Python API it is easily extendable and has a strong user base developing plug-ins.

Recent advances of visualisation platforms suitable for macromolecular settings exposed the limitations of visualisation as a technique; namely, its difficulty in dealing with intrinsic dynamics of molecular targets, and limitations in the number of molecular attributes visualised simultaneously. Most molecules are not static in time, and different parts of the structures might be more or less flexible. To account for flexibility is very important in certain aspects of structural biology and rational drug design, and representing the flexibility in an accessible and intuitive way is of a crucial importance for medicinal chemists, in particular for users without extensive background in computational chemistry. Another limitation is the number of attributes that can be visualised simultaneously. Only a certain number of attributes can be shown at any time by varying colours or shapes, but being able to assign several molecular attributes to a molecular fragment (such as group of atoms, residue, protein domain) and access these in a straightforward way could be pivotal for the drug design community. Both functionalities would also be highly useful in research outreach contexts and for crossing boundaries between disciplines (e.g. science-inspired art projects).

In these respects, enhancing molecular graphics software packages with sonification plugins could make dramatic differences in the accessibility of macromolecular data.

To assess its applicability and feasibility, the macromolecular system was chosen, in which intrinsic dynamics plays a pivotal role in its biological function — the human nF-kappaB inducible kinase (NIK). It is a central component of so-called non-canonical nF-kB pathway, which is upregulated in many inflammatory conditions and cancers, such as T-cells lymphoma (TCL). This is what makes NIK an attractive target for cancer research — finding an inhibitor could open new possibilities for treatment of TCL, which has a very poor prognosis in general. The area around the Adenosine triphosphate (ATP) binding site, which can be druggable by inhibitors, is surrounded by highly flexible loops, including the activation loop, which directly controls the biological activity of the enzyme (Figure 4.1). The dynamics of the so-called hinge region is also involved in regulation of the ligand/drug binding to the protein. Despite their biological importance, these features may be challenging to spot in conventional visualisation strategies, i.e. when a single structure of the protein is visualised.

4.2 Related work

Barrett [12] uses Ambisonic to render data into sound in her Cheddar application implemented in Max/MSP. She draws from spectromorphological concepts established by Smalley [107] and combines it with psychoacoustic insights to create a tool for interactive parameter-mapped sonification. Data is mapped to up to six parameters, three of them spatial, two user-defined and one representing a time step or stamp. Scaling and data filtering is controllable in real-time. She argues that as mappings can be tweaked in real-time, it is possible to design an application that is suitable for a broad range of De Campo's [29] design map, rather than taking his approach of designing non-real-time software for a specific dataset. Distance cue processing is deemed ineffective and CPU-intensive and dropped from Cheddar in favour of more effective features. She chose higher-order ambisonics (11th order 2-D and 5th order 3-D as provided by IRCAM's spat objects [1], the order chosen from practical work, not from formal listening tests; Harpex is used for binaural rendering.) as it is not fixed to a speaker layout, as in the use of Vector-based amplitude panning (VBAP) or wave-field synthesis

---

[1] See `http://forumnet.ircam.fr/fr/produit/spat/`.

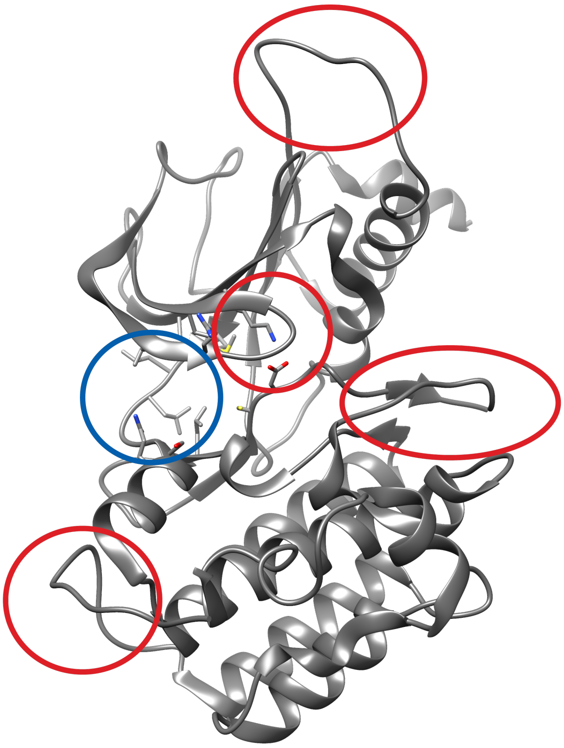


Fig. 4.1 *Ribbon diagram of the kinase domain of human NIK (PDB code 4IDT). The secondary structure is represented as grey elements (helices, sheets, and loops). The flexible loops are marked in red circles. The hinge regions is marked in blue circle. The residues involved in the ligand binding are displayed and coloured by atom.*

(WFS). Cheddar uses granulated sound samples for sonification, synthesised sounds are only used for testing. This is due to their richer spectral possibilities and potential relation to the source data set.

Sonification of proteins and DNA has a long history, dating back to Hayashi and Munakata's mapping of DNA sequences for analysis [51]. Following that, there are many more examples of artistic and scientific auditory display mapping the building blocks of proteins, especially of DNA, to pitches or other attributes of sound. Dunn and Clark provide a very musical example of this, in an artistic collaborative project to sonify protein chains [37]. Garcia-Ruiz and Gutierrez-Pulido provide an extensive overview of auditory display of molecular structures [46].

The CoRSAIRe project [41, 114] developed a multimodal environment for protein docking analysis. In the project, a researcher works in a virtual environment interacting with proteins. Sonification is used to augment the display and the interaction of proteins in 3d space.

Multiple sonification plug-ins have been developed for molecular visualisation software: Grond and Dall'Antonio's [48] SUMO framework is a sonification plugin for the molecular visualisation application PyMOL. In their framework, they implemented two example sonifications: an amino acid sonification and a B-factor sonification.

Rau *et al.* [91] use sonification to augment events extracted from molecular dynamics simulations in MegaMol, a visualisation middleware for visualising point-based molecular datasets. They used OpenAL to provide spatial audio and auditory icons to highlight events happening in the simulation. This aimed to prevent the user from missing them due to occlusion, e.g., H-Bonds forming and breaking in the simulation. Presenting the plugin to collaborators in the field of structural biology, they received positive feedback, though they did not see any immediate advantage for their day-to-day work.

As a side note, the FoldSynth environment for protein folding synthesis [115] provides an (as of yet) undocumented sonification plugin.

The existing systems described here either are not publicly available (anymore) (SUMO, Rau's system), depend on non-standard technology (CoRSAIRe), or are undocumented and not self-explanatory (FoldSynth). The system presented here seeks to alleviate these problems of access by publishing the source code of the plug-in, which is itself running on open source software, and providing documentation. The ideal outcome would be a sonification plug-in directly integrated into the main codebase of a major protein visualisation software, broadening the user base for sonification by making it trivial to access it.

Furthermore, this chapter presents a collaboration with domain experts followed by an evaluation of the system with potential users.

4.3 Design

This section explains the components of the design of the application and how they were chosen.

4.3.1 Spatialisation

As other immersive approaches were successfully using spatialisation as part of the sonification technique, this approach was also tested in these experiments [91, 114]. As the sound in this system is bound to the visualisation it also seemed the most intuitive choice to bind sound sources corresponding to parts of the protein to virtual spatialised sound sources.

The plugin is intended to easily be integrated into the current workflow and technical setup of the intended users. Therefore headphones were used as preferred delivery method.

This also enables the use of head-related transfer functions (HRTFs) to render a 3D sound field on standard stereo headphones.

4.3.2 Interaction

As macromolecules such as proteins, nucleic acids, and their complexes can be very large and flexible and information overload of the users was a concern, interactive sonification was employed to enhance the user experience. The user can click on elements of the molecules to get sonic feedback and a temporary colour change of the component which is sonified. This visual feedback provides a point of reference to the user, especially as spatial distribution in the sound is not easy to discern when the zoom level is small and the distribution is concentrated in the centre of the field.

In most tasks where molecular visualisation is used, not only the current element but also its surroundings are important, e.g., as the type of flex an element shows is inherently dependent on how flexible the surrounding elements are. Therefore a travelling wave paradigm was implemented, i.e., an interaction based on the idea of a wave circularly spreading outwards in all directions from a point in space. In this case, a wave front spreads outward from the point the user interacted with, similar to the Data Sonogram method of Model-based Sonification introduced by Hermann and Ritter [57]. For simplicity reasons the wave does not actually take into account the physical centre of the elements it spreads to, just spreading equally spaced in time to adjacent elements. The wave "loses energy" as it travels, rendering later elements of the sonification less pronounced. The visual feedback is coupled with the wave front, i.e., the moment an element is "hit", the temporary colour change occurs with the colour intensity relative to the energy of the wave (see Figure 4.2).

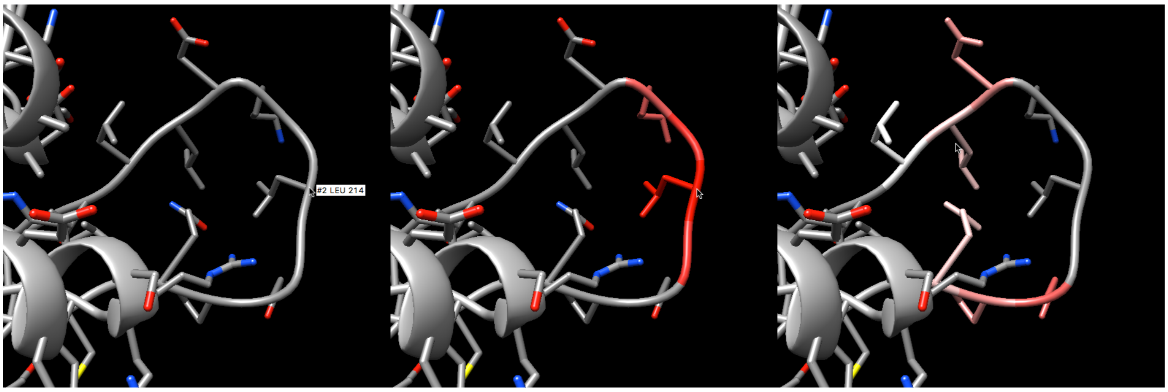


Fig. 4.2 *Travelling wave interaction visualisation. From left to right: before clicking on an element, after two steps of the wave, and the end of the wave (fading from red to white as it loses energy).*

Three different wave propagation methods would be possible in the described scenario (currently only (1) is implemented in the plug-in because of time constraints):

1. propagation to the directly connected neighbours of the origin element;
2. propagation as in the first case, but also to elements connected with H-bonds;
3. propagation in three-dimensional space according to the radius of the wave's reach.

More propagation methods could provide the user the choice to concentrate on the immediate vicinity of the element (1), the logical vicinity (2) or the spatial vicinity (3). Initial testing found that in some applications it is beneficial to stagger the wave front in mode (1) to prevent simultaneous events. This means that instead of playing the element the user clicked on and then the adjacent elements, the adjacent elements are staggered, so if the user clicks on element 3 in a chain of five elements, the elements are played in the order 3, 2, 4, 1, 5 instead of 3, 2+4, 1+5.

Controls are provided to change the propagation rate and radius of the wave, as well as enable or disable the staggering of the wave.

4.3.3 Backgrounding of sound objects

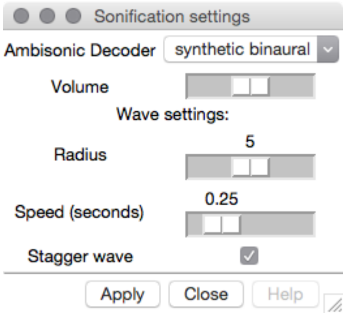
When mapping parameters to pitch, high-pitched permanently sounding objects in the soundscape can overburden the user, especially in the stereo version. The described sonification design demanded some elements of the molecules deemed important to be sonified permanently though. Therefore an interactive property was implemented to keep sounds in the mix, in a volume corresponding to their location in space, but without dominating the soundscape.

The permanently sounding objects are therefore low-pass filtered to 300 Hz (subject to further evaluation) after their creation. This is similar to the process used in the image sonification (see Chapter 3). This was to place them in the perceptual background of the scene by removing their dominance in the soundscape. The filter is lifted for a short amount of time only when a "wave" hits these elements or the user clicks on them, restoring the backgrounded sound objects to their former spectral glory.

4.4 Implementation

4.4.1 Framework

UCSF Chimera was chosen as the target platform as it is widely used, open source, free for academic use, and relatively straightforward to extend. It provides support for all data


Sonification settings

Ambisonic Decoder: synthetic binaural

Volume

Wave settings:

Radius: 5

Speed (seconds): 0.25

Stagger wave ☑

Fig. 4.3 *Sonification settings GUI. Allows the user to change the Ambisonic decoder used and the volume of the sonification. The "Wave settings" part relates to the Data Sonogram-like travelling wave interaction. The radius determines how many elements the wave reaches and the speed determines the wait time between elements. "Staggering" refers to not simultaneously playing parts of the wave that would be played simultaneous otherwise.*

formats (PDB and mol2) and applications (visualisation of molecules and docking tasks) targeted.

The plugin is written in the Python programming language. Chimera takes care of translating the source files in PDB or other formats into hierarchical data structures representing models, made up of residues, atoms, and bonds, etc. It provides triggers for changes in these data structures which is used to detect loading and deleting of molecules, updating the sonification accordingly.

All other data processing is done in Python with the help of the open-source scientific computing library Numpy [116].

SuperCollider is used as the synthesis engine, as it has good support for all major operating systems, supports 3D sound with Ambisonics, and has good sound synthesis plugins [74].

The plugin communicates with SuperCollider over UDP using the Open Sound Control (OSC) protocol. As all data processing is done in Python, other sound rendering options are possible as well, e.g., using Pure Data or Cycling74's Max, as the OSC commands used are relatively language-agnostic (see Table 4.1). This is similar to the SUMO molecular visualisation plugin, which linked audio and video processing via OSC, but used Max instead of SuperCollider [48].

In Python, the individual parts of the sonification are represented by "sound objects", data structures that correspond to synthesis processes in the sound rendering software. In the current implementation they correspond to *Synths* in SuperCollider, but they could also correspond to data structures in other (music) programming languages. Each "sound object" has a unique id shared between sound rendering software and plugin, and can be modified by

Table 4.1 *Simple sound object OSC protocol. Sound objects correspond to synthesis processes on the synthesis engine and data structures on the client.*

OSC command	Explanation
`/obj/new id sound_type [attr_name attr_value]*`	Add new object with id and sound type
`/obj/modify id [attr_name attr_value]*`	Modify existing object by id
`/obj/delete id`	Delete object by id
`/reset`	Reset everything (delete all sound objects and samples)
`/sample/new id path`	Load sample at path to this id
`/decoder/set name`	Switch Ambisonic decoder
`/volume/set volume`	Set global volume between 0 and 1.0

setting arbitrary strings to numerical values, corresponding to arguments to those synthesis processes.

SuperCollider's *sclang* is used to implement handling of sound generation and synthesis processes. Sound is spatialised with the help of the Ambisonic Toolkit (ATK) [5]. Sounds are placed according to the position of the represented part of the molecule (atom, residue, bond) in relation to the camera position of the user's viewer. Chimera provides a trigger for changes in viewpoint, which is used to recalculate all sound objects' spatial positions and update the sound objects accordingly.

As the listener's coordinates are taken from the camera position, the user is not able to manipulate the listening position separately from the viewing position. Possible future work could include placing a separate "listener" in the scene, or positioning the listening position in front of the viewing position to prevent having a majority of sound objects bunched up in the centre of the sound stage and helping with discerning them.

The plugin's GUI enables switching between different Ambisonic decoders. Several decoders for headphone output based on head-related transfer functions (HRTFs), are provided, with a choice between KEMAR and synthetic HRTFs, enabling 3D sound experience over stereo headphones. A UHJ stereo decoder is also provided to enable output over Stereo speakers. The plugin is free software under the terms of the GNU General Public License and all source code for the above can be found online [2].

4.4.2 Sonification for molecular docking

The first task investigated is the docking of small molecular drug compounds to their cognate protein receptors, in order to design tight-binding (potent) and selective inhibitors (a process called iterative lead optimisation). The same task is routinely performed by medicinal chemists in order to select the best binder from a list of small molecular compounds, a process known as structure-based virtual screening. In this application, the chemists need

---

[2] See `https://github.com/mortuosplango/chison`.

to understand the structure and intrinsic dynamics of the binding site of the protein to draw conclusions about factors governing the binding potency, specificity, and selectivity.

This information allows for the rational design of drug-candidates with the optimal pharmacological profile in order to minimise the number of adverse effects. As some parts of the binding sites can be very tightly embedded in the structure they can be challenging to visualise in a way that is required for the task. Also, the dynamics of the binding site, which are notoriously difficult to inspect visually, may sometimes play a pivotal role in governing the ligand-protein associations (e.g., HIV-1 protease inhibitors, hERG potassium channel and cytochrome P450 binders).

An auditory display was designed to illustrate the electrostatic and van der Waals' interactions influencing the enthalpic contribution to the free energy of protein-ligand association, and the atomic positional fluctuations (APFs), which are the measure of the conformational flexibility of the protein target and ligand molecules (entropic contribution to the free binding energy), in order to support the user in the processes of virtual screening and iterative optimisation of the lead molecule.

Auditory icons and interactive parameter mapping sonification (PMSon) were used to give interactive feedback.

Auditory icons represent the H-bonds present between the ligand and the protein.

PMSon is used for the display of the ligand atoms. The constant sound of the ligand is represented using phase modulation synthesis. It is mapped linearly to the modulation frequency, whereas the grid van der Waals' score is mapped exponentially to the carrier frequency. Therefore each sound object in a ligand has the same carrier frequency, while the modulation frequency depends on the individual atoms' charge. This provides a unique overview sound for each ligand depending on the van der Waals' score while providing feedback on the charge of the object on click. The addition of vibrato livened up the sound by adding a bit of randomness to the pitch and wanted to improve source separation according to the principle of common fate [18], where streams that change in parallel (in this case, change in pitch) are perceived as one stream.

Inspired by science fiction film soundscapes a space ship docking metaphor was chosen, with the H-bonds closing represented with an auditory icon modelled after a magnetic seal docking on an air conduit. This sound was chosen to provide a metaphorical link to the task that is executed by the hydrogen bond — providing the strongest linking force between different not directly connected structures or parts of structures.

The protein's residues are sonified only on interaction, where the B-factors (also known as Debye–Waller factors, a measure of the intrinsic flexibility of parts of a protein) are mapped to the pitch of the sound objects. Interactive PMSon was used to show which components are more flexible than others. B-factor values were mapped exponentially to the pitch of a slightly distorted sine wave (see Figure 4.4) produced by $f(x) = \tanh(\sin(x) \times 2.8)$. This provides a relatively small spectral and CPU-footprint while still having some timbral qualities of a square wave, making it easier to localise than a pure sine tone.

By clicking on a residue or atom, the user can play the corresponding pitch of that region and of the neighbouring residues or atoms according to the wave settings.

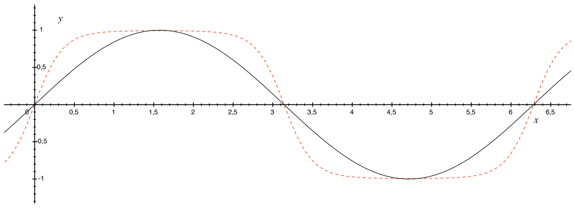


Fig. 4.4 *Sine wave (black) vs. distorted sine wave produced by* $f(x) = \tanh(\sin(x) \times 2.8)$ *(red, striped).*

4.4.3 B-factor sonification

The second task is the sonification of Debye–Waller factors, also known as B-factors, and/or atomic positional fluctuations (APFs) of atoms or amino acid residues, related to their intrinsic flexibility. Combined with a simple animation of the sonified parts this will help the researchers understand the dynamic behaviour of the protein of interest better.

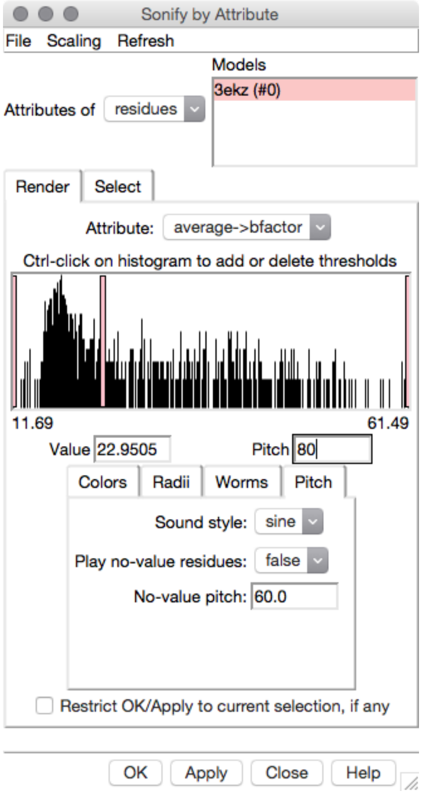
The same interactive PMSon for displaying the B-factors as in the previous paragraph was used. Additionally, the 20% of residues with the highest values were continuously displayed. They are displayed as sequence of sound events with a percussive envelope. The wait time between events is inversely proportional to the B-factor value. Vibrato was added as in the other task. Additionally, the vibrato gets more pronounced in higher pitch ranges (if the MIDI note number is bigger than 80) to create a threshold as used in [85].

By clicking on a residue, the user can play the corresponding pitch of that region and of the neighbouring residues according to the wave settings.

4.4.4 General-purpose sonification GUI

In addition to these tasks, the plugin comes with a GUI that enables the user to specify which data they want to sonify. A widely used Chimera tool that enables the mapping of attributes to rendering parameters (e.g., colour or radius) was extended to also give an option to render in pitch. In future versions of Chimera this general-purpose sonification option could be integrated into the current user experience.

The current version can be seen in Figure 4.5. After choosing a structural element level (atom, residue, or molecule) and an attribute, users can define markers in the histogram view and which MIDI pitch they represent. Values in between these markers are mapped with linear interpolation to the chosen pitches. They can specify a sound, as well as if a pitch and which pitch is played in case there is no value associated with the element. The user then can interact with the molecule by clicking certain parts of it and configure the resulting wave front in the sonification settings (see Figure 4.3).



Models: 3ekz (#0)

Attributes of: residues

Attribute: average->bfactor

Ctrl-click on histogram to add or delete thresholds

11.69

61.49

Value 22.9505

Pitch 80

Sound style: sine

Play no-value residues: false

No-value pitch: 60.0

☐ Restrict OK/Apply to current selection, if any

Fig. 4.5 *Sonify by attribute GUI. In this case, mapping the average B-factor of the residues to pitch. There are 3 markers defined in the histogram, each setting a pitch. The pitch values in between markers are linearly interpolated.*

4.5 Evaluation

In an evaluation, eight participants were asked to evaluate the prototype for displaying molecular dynamics. Seven out of eight participants are working in the field of structural biology or computational chemistry at least on a PhD candidate level. One participant is an electroacoustic composer. Each was shown a MD-Movie (molecular dynamics movie, a simulated movement pattern of the protein), a coloured representation of B-factors and the sonification displaying the same protein (see Figure 4.6). The participants were asked to identify the most flexible region of the protein in each representation. Monitoring the response and asking the participants to fill out questionnaires before and after the study aimed to determine: how the system could be integrated into their workflow; if the software system with sonification could have a positive influence on drug design tasks; and how the system could be improved. Minimal instruction was given on how to use the sonification, just explaining the mapping and the interaction possibilities.

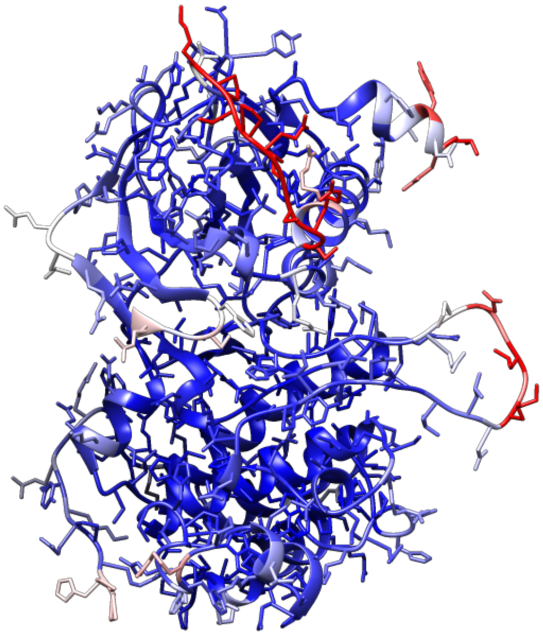


Fig. 4.6 *Coloured representation of B-factors in human NIK protein. Red parts represent high, white middle, blue low B-factors.*

Before using the software, the participants were asked to fill in a short questionnaire about their listening habits, their working habits and their work place. They were also asked what they imagined proteins would sound like (see Appendix C for the questionnaires and Appendix B for the consent form).

Seven of eight participants filled out the questionnaires. Four reported working in a quiet environment. Five use Chimera daily. Five participants had some level of musical training. None of the participants reported associating any sound with proteins. Besides the composer, nobody reported using sonification or sound in their work. When asked if they listened to music while working, which would imply they had some means of sound reproduction already available but also that the audiovisual interface could interfere with their usual habits, six participants reported to listen to music while working at least once a day.

Participants were audio- and video-recorded while using the system. After using the sonification, the participants were asked to fill in a questionnaire about their experience using the software. Based on the questionnaire results the audio and video recordings were not deemed helpful and not further analysed.

The questionnaires and informal chats with the participants showed a wide range of responses. One experienced researcher entirely dismissed the whole idea of sonification and the proposed setup. Comparing it to the visual modality, this researcher was quick to point out that it was much easier to see than to probe the protein for sonic information, and that this endeavour therefore is a waste of time. The less senior participants were more open to the ideas behind this work in general and its possibilities. Three participants remarked that although the visual modality was quicker to present the information, the different interaction type and the relatively lower speed of interaction lead to a more contemplative exploration.

In general, no participant had problems using the sonification to fulfil the task. Seven out of eight found the most flexible region using the sonification.

In order to gain a better insight into their attitude towards the sonification, we included seven possible descriptors into our questionnaire, for the subjects to rate their experience on a five point Lickert scale. Five out of seven participants agreed with or strongly agreed with the sonification being “beneficial” to the task at hand, four being “musical”, none being “annoying”, five being “helpful”, none being “fatiguing”, one being “transparent”.

On the other hand, three participants disagreed with or strongly disagreed with the mapping between data and sound being “transparent”.

Five participants reported they were interested in using this sonification in their work, four participants were interested in using it in research, four in teaching, and two in an art context. One participant remarked that it “would be good to use if teaching students or individuals who do not have a lot of experience in using chimera or other applications”. Four reported they would be interested to use sonification in general in their work.

In conclusion, there seems to be at least some general interest in sonification in the field. The general work environment noise level appears to be sufficiently quiet to use sonification. As most participants listen to music at work, the hardware for sound reproduction is already there. The interaction in the mapping turned out to be intuitively understandable even for participants which do not regularly use Chimera. The difference between participants perceiving the mapping as transparent and the sonification as transparent points to a problem in sound design of the overall sonification, showing an imbalance between the volume levels of permanent and interaction-triggered sounds.

4.6 Video Example

The described tasks and the general purpose sonification GUI is illustrated in a video example[3]. It shows first the docking task, then the B-factor sonification, and last the general purpose GUI.

4.7 Conclusion

This chapter presented a sonification plugin for the molecular visualisation application UCSF Chimera. Two sonifications for important tasks in computational chemistry were developed, alongside a general-purpose sonification GUI.

Presenting the sonification to potential users yielded valuable, mostly positive feedback. Using a travelling-wave paradigm related to Data Sonogram scheme for interaction with the molecular structures proved intuitive for novel users. The backgrounding via low-pass filtering for permanent sounds seemed to be valuable, but the filtering needs to be adjusted to account for the feedback.

Participants in the evaluation were interested in using sonification in their research and teaching, and the implemented approach to interaction was intuitively comprehensible.

The implemented spatialisation provides a more direct coupling of the sonification to the visualisation, mapping the visual field to the sound field with good accuracy. It provides a level of overview when parts of the structure are behind the viewpoint as they will still be part of the sound field, adding more realism to the perception of the protein. This could be seen as annoying or not helpful as the sonification is supposed to enhance the visualisation for the user, the standard way to look at proteins is having them in front of the person, not the person virtually sitting in the protein. If they were in the protein but seeing it in front of them the localisation would be even more confusing. An option to toggle between a ‘realistic’ and a ‘filtered’ sound field should be provided.

The current implementation of the travelling wave is slightly arbitrary in the way it travels evenly spaced out in time, disregarding the spatial distribution of elements. It provides an extension of the interaction in Chapter 3 though and a kind of “glance” at the vicinity of the chosen object when set to travel faster.

A structure that looks static but actually moves, as in this case, could be imagined as soundscape of a creaking staircase or a large structure that looks static but produces sound. The approach to sonification design in this chapter was lacking a clear structure and metaphor

---

[3] Available at `https://www.youtube.com/watch?v=IjQrL1T3h0s`.

for the whole task and process. In the next chapter we investigate how the process of designing a sonification can be improved based on the lessons learned from this project.

Chapter 5

Subject-Position Based Design Framework

> At ICAD 2011, Emery Schubert, a musicologist, held a deliberately provocative talk, beginning as follows: "I am here to teach you, ICAD people, what sonification is about. My claim is that music has the potential to sonify our emotions". If this claim goes fundamentally against the vision of the great majority of researchers and sonification pioneers, it illustrates the ambivalent relationship between the two domains: there is often an artistic dimension in sonification design, and music can sometimes be composed on the basis of scientific data, or in the perspective of communicating objective information. [36, p. 10]

5.1 Introduction

This chapter describes a new framework for sonification design. It is called "Subject-Position Based Design Framework" (SPBDF).

This framework was developed as a response to the ideas set out by Vickers [117, 120] and which he expanded with Hogg and Worrall in [121]. These ideas did not consist of a step-by-step guide, more a general insight into how sonification design could be approached in a different way. A driving factor for the development of SPBDF was the frustration with the slow process of adaptation of sonification, coupled with disappointment over very data-focussed sonification designs, which seemed to forget about the user. Another factor was a reflection on the design process undertaken and described in the previous two chapters, its perceived lack of direction and role in the shortcomings of the resulting sonifications.

The development of this framework took place over the course of the different projects presented in this thesis. The sonifications described in the previous two chapters were developed without adhering to a specific design framework or a formalised design theory. Each embodied the author's thoughts about sound design in their way, which discussions with Vickers and the literature referenced here catalysed into the framework described in this chapter.

Once a rough set of principles was extracted from the ideas and previous experience, the framework was further developed through analysing existing sonifications created by others based on these principles, representing a very vague understanding of how this theory could work. These analyses were then compared to studies or other evaluations conducted on these sonifications, and the framework refined and concretised accordingly.

The following sections outline the background, the motivations for developing this framework, discuss the practical implications of the framework, document the analyses of sonifications, and show its application to the design of the SoniFRED sonifications (see Chapter 6 for more details).

5.2 Background

Visualisation is not (usually) illustration. Our viewing experience is mostly based on our surroundings, paintings, illustrations. We first have to learn how to "read" a visualisation before we can fully make use of it. Nothing is intuitive about a visualisation. It does not exist in the natural world.

Sonification, likewise, is not (usually) music (though see [118, 121]). Our listening experience is mostly based on listening to speech, environmental sounds and music. Training people to listen analytically is one of the big challenges to establishing sonification as an alternative or complementary mode of data representation [102, 121]. Visualisation expert and commentator, Robert Kosara, wrote: "Sonification is not visualization for the ears, it follows completely different rules" [67]. And it is these different rules that make it so interesting and so challenging to do well.

Sonification design is in its first generation. The vast majority of research effort has been directed at designing sonifications for very narrow use cases with results and design heuristics that are not generalisable. Sonification has suffered from a lack of agreed, well-formulated evidence-based design guidelines/heuristics or even rigorous theoretical underpinnings. Much of the work is ad-hoc, dependent on the preferences and prejudices (and even ignorance of aesthetic and perceptual issues) of the sonification designer. There are several extant design frameworks, typically approaching sonification from the point of view of the data types involved (e.g., [30]) or information requirements (e.g., [11]) but there remains a lack of heuristics based upon the more fundamental relationships listeners have with sound.

Wolfe [128] connects sonification with the mysticism of negation: “Both practitioners of sonification and apophatic mystics believe that certain types of information are incomprehensible through traditional analytic means and can only be understood through experience. In this way, sonification can be thought of as a source of mystical information.” [128, p. 304] She explains that “Kataphatic methods suggest experiences are ultimately describable, while Apophatic methods claim that the information is ineffable and can only be experienced.” [128, p. 304] She cites Steiner [108] as saying that the listener “[...] must get beyond the point of caring whether, for him, the sound is pleasant or unpleasant, agreeable or disagreeable, and his soul must be filled with whatever is occurring in the being from which the sound proceeds.”[128, p. 308] She argues that the spiritual and mystical traits of sonification are not fully appreciated and sonified information provides a structure through which artists “can investigate facets of the world and our place within them”[128, p. 308].

Though in an artistic context these reflections can be seen as positive, they show one of the drawbacks of a lot of sonification design — obscuring information seems to but should not be the result of an auditory display designated for a specific use case.

5.2.1 Subject Position

Maybe the current approaches to sonification design can only lead to obscuratism: Vickers, Hogg and Worrall [121] argued that listeners are culturally primed to hear pitched sounds (or unpitched sounds, as in Pierre Schaeffer’s, Cage’s, and many other’s works sampling or curating sound) in a musical sense rather than as data, so the underlying data is overshadowed by the listener’s knowledge and experience. They further argue that the sense of hearing is associated to the other senses through our experience of comprehending our environment, i.e., should not be divided from the source of the sound. Sonification therefore has to take the aesthetics of the sounds into account and guide the listener by developing a ‘subject position’ — the part of the perspective on an object that is non-individual, and therefore inherent in the object of perception.

They base this idea of the ‘subject position’ in sound on the work of Clarke [25], who defines it as follows:

> The subject-position of everyday life is overwhelming one of transparently active engagement. But many aesthetic objects and circumstances change this seamless state of affairs by radically limiting the perceiver’s capacity to intervene in, or act upon, the immediate environment in a free-flowing manner. Under these

> circumstances, perceivers may become much more aware of their perspective on the objects of perception.Part of that perspective is utterly individual . . . [b]ut an important component is also built into the material properties of the object of perception, and is therefore a shaping force (at least potentially) on every perceiver. [25, pp. 124–125]

So every person perceives things partly differently, but their interpretation is also partly restricted by the object of perception. Therefore sonifications themselves impose limits on the individual's room for interpretation. They conclude that information content and aesthetics are not two different axes of design, as music can also be both structure and emotion [121, p. 15]. Therefore "aesthetic enters at the point of constructing the subject position such that a hearing-in [sic] becomes more probable, that is, something in the aesthetic of the sound has to match the phenomenon being revealed"[121, p. 18].

5.2.2 Directness

Another factor in mapping is directness. The subject position and its construction can help designers think about embedded and restricted possible interpretations inherent in their design. This restriction of interpretation is related to how direct an auditory display points to its external reference. Vickers and Hogg [119] used the term indexicality to describe something that points to some other thing that is external, e.g., an idea or an entity. This something could be, e.g., a noise or a sign. How indexical a mapping is can be seen as a measure of arbitrariness — as in semiotics an indexical signifier is non-arbitrary, therefore having a direct physical or causal connection to the signified.

In terms of sonification design, data drives the parameters of an acoustic signal in parameter mapping sonification whereas in model-based sonification user interactions with the data produce the sound [57]. Maximum indexicality would be achieved by a system that directly derives the sound from the data itself. Audification would therefore be at the top end of the indexicality spectrum. In contrast, low indexicality would be a more symbolic or interpretative mapping.

Caddick-Bourne has been cited as suggesting with regards to the representational aspects of auditory displays

> . . . [that when] a work represents some data then it seems that the data must, in some more substantial way be part of how the representation is properly experienced, so we somehow experience the thing in terms of what it represents ([19] cited in [121, p. 96]).

For audification, because the coupling between sound and data is very close, the representation is direct and arguably the experience is also directly in terms of the data. On the other hand sonification can be more or less direct. Does directness have an influence on the usefulness of a sonification?

There are examples of sonifications with low indexicality, which were not deemed helpful when evaluated: Edwards *et al.*'s sonification of biological cells (also outlined in Section 3.4), where image features of the cell samples were mapped to a 1000-point 'badness' scale indicating cancerousness. The scale was mapped to ten different environmental sounds, each representing 100 points of the scale, where a cow mooing represented the lowest level of the scales, and a woman screaming the highest level. It turned out that participants found the mapping difficult to use, and described it as too arbitrary and not a clear linear mapping of cell properties.

On the other hand, sometimes low indexicality can hardly be avoided: Taylor [113] sees Morse code as sonification. For the English codebook, patterns were chosen based on the frequency of the letters. Therefore the mapping is quasi-arbitrary, which in this case is not a value judgement as the system the code operates in (pulses of arbitrary lengths) does not support non-arbitrary values. He further argues that Morse code and spoken language share this symbolic nature of their mappings. He quotes from Fitch's *The Evolution of Language*: "...arbitrariness is almost automatic if you start with a vocal system, for the realm of the iconic is rather limited in vocalizations. Onomatopoeia can buy you some animal names, and some emotional expressions, via imitation, but not much more. But the flip side of the coin—too often overlooked—is that arbitrariness is a crucial step to a fully open field for semantic reference, and this is something that we gain almost automatically with the capacity to link meanings to vocal signals..."[113, p. 1].

5.3 Motivation

Subject position based design theory could help reach people unaccustomed with sonification. Auditory display can be seen as being based in a spectrum of different types of sounds: music, functional sounds, and environmental sounds (see Fig. 5.1).

Functional sounds in this context are "designed" by humans to fulfil a certain function outside of their production and enjoyment, whereas music denotes organised sounds produced as aesthetic objects, which is a function as well, but is being ignored for simplicity.

The environmental soundscape encompasses all sounds around us at any point in time, theoretically also encompassing both functional sounds and music. In this context it is meant to represent sounds that are not "designed" by humans for a specific function (e.g., bird song

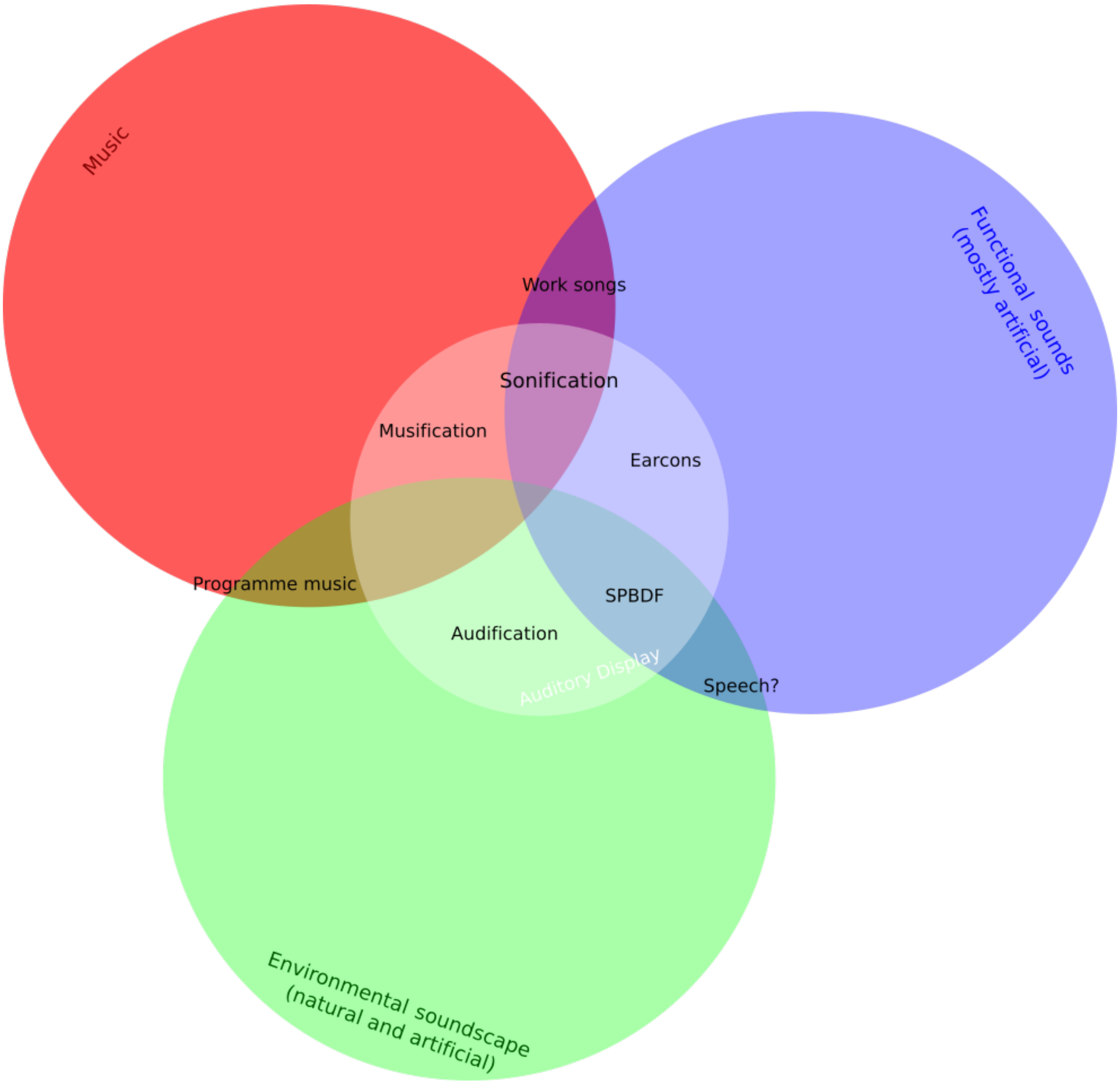


Fig. 5.1 Sonification design based around a spectrum of different types of sounds.

might be entertaining, but it was not designed by humans; engine sounds of cars might give indications about the state of the engine, but were not designed to provide this information).

Work songs are positioned as overlap of music and functional sounds as they provide a function outside of entertainment in the synchronisation of workers' tasks.

Speech could be seen as both functional sound and environmental soundscape, as it evolved over time but could also seen as being "designed" by humans as writing fixed words and many words are created deliberately, even if their origin might be obscure.

Programme music (e.g., Vivaldi's *Four Seasons*) on the other hand takes inspiration from our environment, but transforms it into a composed form for entertainment.

Earcons are designed to communicate specific information to the user, positioning them squarely in the functional sounds category.

In contrast, audification "plays" the raw data. It could be seen as the simplest type of sonification, where each data point is mapped straight to a speaker cone's movement. This playback process might be "designed", e.g., by speeding up the playback to change the frequency range of the earth's seismic activity to be perceptible by humans, but the resulting sound is seen here as environmental in nature, as the mapping from data to sound is not designed in the sense that no complex mapping occurs (one could obviously argue that mapping DOES occur as the data probably will have to be scaled to the range necessary for the DAC process to occur).

Musification is seen as auditory display designed primarily for entertainment, not providing another function.

Sonification in this schema sits at the overlap of functional sounds and music, as it is both designed for a certain function, but in many cases uses musical tropes like melody or rhythm to render information.

Without additional long term training, only a small number of people can use and understand traditional sonifications. This design framework is placed outside of the realm of music or functional sounds, instead residing in the realm of the soundscape we are embedded in. This is the only realm of sound that a majority of people is in any way trained to listen to analytically. Music is perceived more as a background activity for the majority of people, and alarm sounds will be perceived analytically in some cases, but in the majority the sounds are too interchangeable and low in information to really afford any challenge to us. Consequently the only level of analytical listening this framework should expect from its users is their — frankly most probably stunted — ability to discern broad categories of sounds: thunder from birdsong, a bicycle from a car, a car engine from a lorry engine, shaking a half-full from shaking a full water bottle.

Therefore we propose to approach the problem from the other side and design sonifications as extensions of common soundscapes. This framework builds on this narrow basis to design sonifications that can make sense to a majority of people without falling back on minor chords meaning bad and major chords meaning good, or high pitched sine waves meaning bigger than low pitched sine waves.

5.4 Soundscapes and future proofing

Soundscapes are part of our environment and can change considerably over time. How do we make sure they are relevant for a broad (age) range of (future) users? Some soundscapes stay stable over a longer period, like in the mobile phone application Shoogle, which represents the messages received by a phone by equating the phone to a box of objects [125]. Shoogle was designed with the paradigm of a box of objects being shaken to determine its contents. Boxes aren't ephemeral metaphors. The ephemeral part of Shoogle is the actual implementation — developed for the now obsolete Symbian operating system and without released source code.

Geiger counters represent the amount of radiation by clicks. Fundamentally these clicks were produced by primitive means so essentially today we could produce more sophisticated sounds to represent this information, but as a rule these clicks remained the same.

Car sirens are using models of old sirens as starting point of their sound even though the mechanical versions of those sirens are hardly in use nowadays. The same goes for some car indicator sounds that are based on a mechanical clicking element or electric car sounds — how far do you go away from the original sound and make it easy for the user or the pedestrian to still realise the underlying meaning?

Many metaphors in desktop environments — including the desktop — must seem strange to younger users — how many people never really understand the standard folder icon until they work in an office with these types of folders? Some applications already replaced the floppy disk icon indicating saving with a different icon.

One could imagine a sonification based on obsolete noises of computer technology — hard drives speeding up and reading, dial up modem sounds, or floppy disk motor noises — that would actually be outdated today and hard to understand for younger people who didn't grow up with these sounds.

No soundscape, not even soundscapes in nature (especially in a time of mass species loss and climate change), will stay the same forever, but we should at least try to make sure we don't base our sonification on a soundscape that has an expiry date.

5.5 Framework

SPBDF is about finding metaphors users can relate to and that are related to the data being sonified. It provides guidance to help designers find these metaphors and translate them into auditory displays.

To summarise, it is based on the assumptions that

- training in analytical listening is severely lacking in the average user and can't be assumed;
- in light of this, it can only make sense to map the most important aspect of the data to one sound stream;

- the resulting sonification should provide pointers to this most important aspect of the mapping or even have them embedded;
- mapping of more aspects in other streams or the same stream may be desirable, but should not remove the focus of the mapping strategy from the average user's capabilities;
- meaning will never "pop out" (see [111]).

Therefore the coupling of data to sound has to be more strict, so the nature of the data gets reflected in the sonification. This is a paradigm that feeds into subject-position based sonification design.

The development of a sonification according to this framework comprises the following steps, which will be described in more detail in the following sections:

1. What is the task? What is the most important part of the data for fulfilling the task?
2. Find a soundscape related to data domain and task
3. Find the most salient feature of the soundscape and map as little data as possible to this one feature (or combination thereof)
4. (optional) Map more data to less salient aspects of the one feature (or combination thereof) or less salient features to enable more aspects being discovered by the user

What is the task? What is the most important part of the data for fulfilling the task?

The design of a sonification should not begin without determining the task that the sonification should help the user with. Clarity of description is paramount, as without a good task description it is impossible to determine which parts of the data are the most important ones to sonify.

Sometimes it can make sense to apply dimensional reduction to a data set if there are more than a couple of dimensions that are all relevant to the task.

Find the Soundscape

The next step is to find the scene to place the determined mapping in. Therefore a soundscape needs to be found that is related to the task the sonification should fulfil. This is particularly hard if the task is very abstract or its soundscape unknown. In these cases sometimes the data domain or its meaning can help to find a fitting soundscape. Creating a mind map around the context and free associations with it can help in that case. Sometimes asking the users what sounds or mental images or contexts they would associate with the topic or their activity as happened (in this case unsuccessfully) in the questionnaire used in the evaluation of the protein dynamics sonification described in Chapter 4, or in the case of the inspiration of the steam train sonification coming directly from one of the researchers working with the device as described in Chapter 6, can also be helpful or a useful starting point.

Inspiration for the composition of the soundscape itself, or alternatives to it can be found, e.g., by searching or browsing the online sound database on *Freesound* (see https://freesound.org/), where people from all over the world and different disciplines can share and catalogue sound clips [4], or in the sound archive of the British Library (available at https://sounds.bl.uk/). Further helpful resources could be video documentaries or — potentially even better — raw video material around the topic or task or potential soundscapes, which can help delineate the contributing elements of a soundscape by, e.g., seeing which part of a machine moves in time with the sound produced.

The selection should also take into account the target audience of the sonification, e.g., no steam trains in a country without train travel.

Find the most salient feature and map to it

With a soundscape selected, the lowest common denominator for understanding the soundscape has to be determined, i.e., the understanding of the data should not rely on a car mechanic's understanding of engine sounds or a medical doctor's understanding of heart murmurs. As most people are not very well trained at listening, only common cultural experiences can be relied on.

The most salient feature might also be a combination of several features. In that case a common vector to map these to the data has to be constructed. A rich mapping from data to sound is desirable, i.e., changing just the pitch but not a rhythm or timbre might not be enough.

(Optionally) map more data

Once a crystal-clear mapping of the most important domain for the task at hand has been determined, more data mappings can be added into the mix to provide long-term usability. A simple mapping might be good for first-time users but regular exposure should also lend itself to a broader experience of parts of the data.

5.6 Tutorial

In this section two examples of using the framework to design a sonification is explored. One case of an abstract task / data domain and one of a concrete domain.

5.6.1 Abstract domain: stock market monitoring sonification

Bob likes to stay up to date with his stock markets, especially in these volatile times. He wants to know if any of his shares significantly change their value, as well as how they relate to an index (e.g., FTSE100).

Task/Data domain: Inform user of stock value fluctuations and general stock market movement.

Soundscape: In this abstract case we would look at soundscapes that could reflect the energy of the trading as well as the events that are important to the user.

Possibilities we might think of might include a cafe with falling objects, spoons, glasses, plates, things being filled for upward motion, crowd noise levels for index. Or objects hitting the floor or bouncing off the ceiling representing the companies' products - which would be hard for services.

If we had a less monitoring focussed task, we could think of a request-based model of objects in a box changing size and therefore sounding differently if we shake this virtual box.

In the end we might end up with a trading floor soundscape. This might be a bad choice as trading floors might become obsolete soon and the soundscape might not be familiar to a broad range of people, but for this specific use case and user group it might be the best choice at the moment.- We would have the general crowd noise level to map to the index as well as voices reading the name of the stock with a questioning inflection for falling courses and an affirmative inflection for rising courses.

Find and map: The most important axis is essentially an event-based one, where big fluctuations are represented in events denoting the different stocks and their movement. In our case that would map to the voices of traders shouting the name of the stock title over the crowd noise of the trading floor. The inflection would indicate the movement of the stock - questioning for down (as it expresses "doubt" about the stock), affirmative as up ("confidence" in the stock). The voice pitch - maybe there could be 5 traders (soprano, alto, contralto/countertenor, tenor, bass register), where each stock starts in the middle at the start of the day and changes traders as it changes hands and its course diverges from the starting course of the day.

Optionally: map more : We can then map the crowd noise level to the index chosen. This could quieten as the index rises and louden as the index falls, leading to more frantic action in that case as people are presumably more likely to want to sell shares than hold them in that case.

5.6.2 Concrete domain: river ecosystem

Alice wants to analyse and compare different river ecosystems and their health over a some time period. She is interested in water levels and speed, number of species, and level of three pollutants in the water.

Task/Data domain: Represent river health over a time period. There are many dimensions to this data so we should use dimensional reduction to come up with a general health indicator. We also need to represent the passing of time as the sonification might happen in different time scales.

Soundscape: Obvious soundscape would be a river, with boats, birds, a water wheel.

Find and map: The general health indicator generated by combining the other data dimensions is our main axis. Our main sound of the soundscape is the gurgling of the water and the turning of an old water wheel. These two sounds are constantly part of the soundscape and are coupled - a faster-flowing river would also turn the wheel faster as well as creating more noises. So the general indicator will be mapped to both the amplitude of the gurgling noise as well as the frequency of the water wheel. In our case, a faster river indicates a worse off ecosystem, as it invokes more feelings of fear (flooding, drowning, being swept away). This decision on axis direction should be discussed with the user though, as they could have different associations with regard to river health.

Optionally: map more : We can then map other elements of the soundscape to the other axis, number of species could be represented by the density of birdsong. Three different boats could indicate the three pollutant levels. Time, as an external factor to this sonification could just be indicated by speech synthesis marking the year / season / month, depending on the speed of the sonification. We only indirectly mapped water levels and speed in this case. These could be indicated by other factors in the soundscape, but as we used the most salient feature for our overall indicator it would maybe worthwhile to leave them out as they could be inferred by the user.

5.6.3 Summary

This short tutorial showed how the framework could be applied to two different tasks. In each case the soundscape selected was directly related to the task at hand. In the river ecosystem example we left some axes out of the mapping to make the whole mapping less complex and used the most salient features of the mapping for a single axis representing a summary of the other axes. In the stock market example we used a more event-based mapping as the central point of the task does not need to be permanently represented. By just using speech we skip the whole abstract mapping of an abstract entity (e.g., an insurance company) to a sound and remove some of the learning curve for the user.

5.7 Analysis-based design

This section documents analyses of existing sonifications in a similar vein that Clarke analyses a Jimmy Hendrix performance [25, pp. 48–60], combined with the methods and steps outlined in this chapter. The goal was to look at these designs in the frame of SPBDF and determine what subject position these designs embody and how the nature of the data relates to the sound produced.

The long-term goal of such analyses would be to determine, as a sort of musicological task focussed on sonification and meta-study, if sonifications with clearer subject positions fare better in evaluations. This would support the application of the described design paradigms for the design of future sonifications.

Analysing sonifications is not a very common thing to do, but maybe should become a part of auditory display literature as much of the analysis currently happens informally in peer review, conference discussions, or workshops, instead of in a public setting for current and future sonification designers to reflect upon.

5.7.1 Audio graph of the pound’s decline since Brexit

The Press Association released a video with the title *Audio graph of the pound’s decline since Brexit* in Autumn 2016. It is a sonification of the course of the value of the pound sterling[1]. It was chosen for analysis as it represents a sonification aimed at a non-expert audience,

---

[1] `https://www.facebook.com/PressAssociation/videos/1237013123029982/`

making the the data domain immediately understandable, and the mapping (described below) was deemed fascinating by the author.

The sonification is described as showing the value of the pound compared to the US dollar over a course of time. It starts with a recording of a brass band. After the first two notes, the recording seems to be slowing down, in a way similar to vinyl record players or magnetic tape machines, i.e., it can be discerned from an unexpected ritardando as part of the band's performance by the pitch descending at a rate proportional to the slowdown.

After the third or fourth note it should become clear to most people familiar with the tune that the recording used here is of the national anthem of the United Kingdom ("God save the queen"). By comparison to the graph being displayed in the video, it becomes clear that the speed of the recording is associated with the course of the pound, where a higher speed is mapped to a higher course.

A similar slow-down effect can be perceived in portable magnetic tape players with failing batteries or when switching off the power to a vinyl record player. It is commonly associated with a loss of (electrical) power. Although the media mentioned might not be familiar to younger people, the effect is often used in popular music, TV, advertising, and radio. Therefore, with the knowledge that the tune played is the national anthem, a loss of power can be directly related to the United Kingdom, represented by the pound's course, losing power: the pound gets weaker, as the power gets weaker, a direct correlation.

The timing of the recording superpositions the end of the first verse with a change in leadership (David Cameron stepping down) indicated in the video and the change in the pound's course fluctuating on a slightly higher level than the previous crash. Theresa May's "reign" starts on the words "the queen".

Speed and pound are fluctuating further up and down, but always in a considerably lower speed than the original. The "flash crash" from 3rd October 2016 is marked by the last cymbal hit on "[God save the] queen", this "queen" being stretched downward and faded out at the end of the video

Comparison to the steps outlined above

Task/Data domain: Illustrate the changes in the value of the pound over a period of time. The important data domain is the value of the pound in comparison to the US dollar. Other aspects of the data could be the change over the course of a day or week, or historic comparisons.

Soundscape: The course of a currency is a quite abstract task to sonify. Money could be represented with the sound of coins or the rattling and waffling of a bank note counting machine.

In this case another abstract soundscape associated with the country which the currency is associated from was chosen — its national anthem. This can be easily associated with the country itself, but if it was not played by pitched coin noises the immediate association to currency is not apparent.

Choosing an instrumental version of the anthem enables more focussed listening as the processing of language and the additional comical effect of pitch changes in the human voice are avoided.

UK Rock band Pink Floyd's song "Money" could have been a direct reference, but maybe not apparently linked to the country.

Find and map: The variation of speed and pitch can definitely be argued to be the most salient domain of the soundscape.

Optionally: map more : In this case the timing of the song and corresponding lyrics corresponds to events in the timeline as well as matching the time span to be illustrated.

Conclusion

The sonification uses the normal speed of an — in the UK — universally known song to demonstrate the reference point.

It represents an indirect pitch mapping, but without the need to establish a reference point as the regular speed of the song is known (an advantage compared to a pure tone as more variation in the signal).

The mapping is timed so that major events fall on explanatory parts of the lyrics. Even though the recording is an instrumental rendition, the lyrics are easy enough that most UK citizens probably know where each word belongs. Using an instrumental version also has the advantage that the semantic content is less distracting. Additionally, a pitch shifted voice could have added an unintended comedic effect. Furthermore, choosing a song with lyrics also provides an anchoring point for talking about parts of the sonification and/or the data it represents ("the second verse" instead of "at five seconds in").

The mapping can be seen as very problematic though as it can easily be interpreted as political polemic: a lower pound course is not necessarily a bad occurrence or represents a country losing power, as it, e.g., enables to underbid other countries' export goods, and therefore should not be related to a mapping that can be associated with a loss of power. It implies one interpretation, not an objective representation of the data. Potentially the gross domestic product mapped to the national anthem would have been more appropriate, as this would have provided a better indicator of power.

5.7.2 Speed skating sonification

Stienstra *et al* [110] describe a movement sonification as feedback for speed-skaters (i.e., ice skating but faster and with longer blades). The aim of their *Augmented Speed-skate Experience* is to give real-time information about, e.g., the pressure on the foot, to long-track circuit speed-skaters.

This sonification was analysed as the paper provided ample evaluation and the problem domain was related to interactive feedback systems for physical exercise, as was the system described in Chapter 6. The amount of evaluation was judged by visual inspection without reading before conducting the analysis as to not bias the outcome.

Speed-skating is a relatively young sport with research into its mechanics and optimisations just at the beginning.

The researchers developed a device for each skate of the athlete, with a 100 Hz data acquisition rate, 3d acceleration sensor, and front and back skate pressure. Data gathering and processing resulted in a 150 ms feedback latency.

Sonification

Continuous parameter mapping sonification was used, where parameters were mapped to band-pass filtered pink noise (i.e., noise with a logarithmic distribution of energy over the spectrum). Adjustable parameters were centre frequency, intensity, and bandwidth of the filter. Each mapping was personalised to the subject's auditive perception.

Movement data interpretation layers used for mapping:

- direction and speed of athlete
- pressure changes
- balance per speed-skate
- athlete balance
- change of balance
- forward speed over pressure

- stroke frequency

Main mapping (referred to as **Pearl* in the paper and below): separate for each foot, pressure was mapped to loudness and balance to pitch (high pitch: front pressure dominant, low pitch: back pressure dominant).

Analysis of main mapping (*Pearl)

Using filtered noise as the basis of the sonification should be a good choice as it is related to the sound scape in an ice-skating environment. Pressure to loudness can also be a good representation of the data as increased pressure feels like more muscle movement/power/volume. This sound is also related to the skates scraping on the ice, or the wind in the ears of the athlete while skating at high speed.

Mapping the balance in pressure to pitch is not a straightforward choice. It is an arbitrary parameter mapping decision as it does not reflect the idea of balance itself, especially as there is no notion of being in balance as there seems to be no zero point provided in the sonification. On the other hand the mapping is embodied in an abstract way as humans are taller when standing on their toes when compared to standing on their heels, so forward balance to high pitch could be seen as more relatable than backward balance to high pitch.

As the device gives real-time feedback these mappings can be easily explored even though the balance mapping could be less arbitrary. Through filtering, the balance mapping gets integrated into the soundscape in a sensical way in this context.

Comparison to the steps outlined above

Task/Data domain: Improve speed skating training by providing the athlete with instant feedback on their use of the skates. The most important data domain seems to be the pressure on the back and front of the skate.

Soundscape: The soundscape of ice skating is dominated by the noises skates make on ice. For the skaters themselves the noises of the air rushing by their ears and clothing are added. Using filtered noise in this context can be seen as an abstraction of the existing soundscape. Depending on the nature of the headphones used, this synthetic version of the soundscape probably mixes well with the environment.

Using a soundscape more similar to the skate environment, e.g., using contact microphone recordings of the skates on the ice or the ice itself, could be more confusing than helpful to the skater.

Find and map: The chosen soundscape already comes with its mapping parameters in this case. The most salient parameter would be the centre frequency of the filter. This parameter was mapped to the balance of pressure, as outlined above the most important data domain.

A more salient mapping would also take into account the bandwidth and intensity of the filter, as they interact with how easy the frequency is perceived — high bandwidth and low intensity would lead to a centre frequency that is much harder to perceive.

Optionally: map more : The loudness is then mapped to the pressure itself, which influences the main mapping considerably, depending on the lower limit of the mapping. The amount of pressure applied definitely should be mapped to something and loudness has a physical correspondence in pressure — the more air pressure is applied to, e.g., a brass instrument, the louder the resulting tone.

A richer mapping from the two parameters balance and pressure would have been better in this case.

Analysis of alternative mappings tested less thoroughly

These were assessed through short rounds consisting of seven laps each, with questionnaires after each round. In the following each of these alternative mappings are briefly analysed according to SPBDF without referring to the paper's evaluation. After the analysis of all mappings, the results were compared to the evaluations described in the paper.

A: balance to pitch — low front, high back: This should be less relatable than the original mapping. High back might be good as a warning though, as high is more of a warning than low and standing too far to the back is not a good thing in skating.

Comparison with published result: correct

B: balance not separate for each foot but combined: This mapping is probably less confusing and more helpful, as there is one less audio stream to concentrate on.

Comparison with published result: incorrect, needs spatial information, as it is otherwise hard to learn and lacking information.

C: difference in balance over time instead of balance: Differences of balance is closer to what humans experience, so this mapping might be more useful. It depends on the details of the mapping, which are unfortunately not very well explained.

Comparison with published result: mapping not that good, not informative to the athlete

D: not balance, but pressure front loudness with fixed high pitch, pressure back loudness with fixed low pitch, left and right: Four streams of loudness. Embodiment-wise it is more reasonable, but not what you would expect.

Comparison with published result: Perceived as valuable, but shifted attention to forces instead of balance

E: sonification of sideward acceleration for each foot. higher pitch — higher acceleration. no sonification of balance: Data represented well in mapping, as higher scraping frequency would result in higher frequency. It seems not that useful in other regards though.

Comparison with published result: "no correlation between movement and sonification was perceived"

F: only sonification when both skates on ground: No feedback when skating on only one skate, which seems nonsensical for the task at hand. The muting of sound is very arbitrary.

Comparison with published result: incorrect, useful for specific purposes

G: third stream: stroke frequency to pitch: stroke frequency to pitch is a weird mapping, why not to rhythm or something like that?

Comparison with published result: other stream was perceived clearly, maybe because of mapping in middle of stereo field

Conclusion

After the evaluation of other mappings, the **Pearl* mapping was perceived as most information rich and un-coercive. The other mappings were assessed to be good for specific purposes, but not that good in general.

The analysis using the framework proposed came to similar conclusions as outlined above. The **Pearl* mapping analysed in SPBDF is deemed a successful sonification, which would be similarly designed in the steps of the framework.

Stienstra observed: "To listeners who never used the augmented speed-skate experience, the sound-scape recordings of the sonified movement sound like meaningless wind. To [the speed-skater testing the system], on the other hand, this wind caused by her movement provides a continuous rich information flow." [109, p. 34]

5.7.3 Slowification

Hammerschmidt and Hermann developed an in-vehicle sonification to help people drive at the right speed [49], pointedly called “Slowification”. It uses the inherent 4-channel sound system in most cars to pan the radio signal front-to-back in accordance with the driver’s adherence to the current speed limit. When the user drives too fast, the audio signal moves away from him, panning towards the back speakers. When they drive too slow, it moves towards the front. When they drive the optimum speed it stays in the middle.

This was chosen for analysis because its similarity in task with SoniFRED (see Chapter 6), as both sonifications aim to regulate the speed of the user.

Analysis

The relationship between data and sound comprises of the metaphor of the sound being a “bubble” on an elastic string that follows the car around, but wants to maintain the right speed. Problems with this approach are a very narrow focus on a driver that is not distracted by other passengers or other factors. Assuming the effect of the spatialisation is enough to bring focus back on the driving needs to be thoroughly evaluated.

Comparison to the steps outlined above

Task/Data domain: Represent the difference in speed of the car to the speed limit on the current road. The data vectors are the speed of the car, the speed limit, and the resulting difference.

Soundscape: Instead of designing a soundscape related to the task, this sonification borrows the soundscape already present in many cars in the form of the car stereo. The chosen soundscape is very subtle in that it blends into the current environment. Other possible soundscapes would have been the noises of the engine or an imaginary landscape being traversed by the car.

Find and map: The actual sonification takes place in the spatialisation of the soundscape. The more the speed of the car deviates from the speed limit, the further back the soundscape is panned in the car.

Optionally: map more : No more mapping was applied.

Conclusion

The idea in this sonification is very creative and suitable for the environment the task is situated in. The actual sonification is a spatialisation of an input signal, which sounds very elegant on paper but could be prone to practical issues in the actual implementation. An evaluation has to show if the effect of the spatialisation is enough to change the focus of the driver back to the speed of the vehicle.

5.7.4 Image sonification system (Chapter 3)

In the image sonification described above, images are displayed as well as sonified. The sonification takes its data from features extracted from the images.

Comparison to the steps outlined above

Task/Data domain: Provide another feedback channel for the analysis of tissue samples to determine cancerousness. The data comprises of features extracted from the images on a grid level. The features relate to potential cancerousness of the tissue region analysed.

Soundscape: Abstract sine-wave based cluster of frequencies, one tone per channel per field in the grid. Without interaction, each field is low-pass-filtered. On user interaction, the fields touched are played with higher volume and without filtering. The soundscape evokes bell-like associations when triggered. When no interaction takes place it manifests as drone with slightly detuned partials, resulting in a slow change in amplitudes of the partials, livening up the soundscape.

Find and map: The mapping is part of the soundscape: the frequencies are multiples of a base frequency of 110 Hz, the multiplier determined by the extracted Betti number, which has itself a relation to the potential cancerousness of the region of the image.

Optionally: map more : There is no more data to map in this case.

Conclusion

With image sonification, especially of abstract images, it is really difficult to think of a soundscape appropriate for the task. The resulting sonification here falls back on parameter mapping paradigms but creates a kind of emergent soundscape out of a small number of components — sine waves and low-pass filters. Coupled with the amount of layers — 16 layers per region as there is one Betti number extracted for each of the 16 channels of the image — an artificial soundscape is created specific to the image viewed. The mapping to a multiple of a base frequency creates a natural spectrum of evenly spaced frequencies, automatically creating an exponential mapping with higher Betti numbers not creating a huge spread in pitches, but a different spectrum.

In lieu of an evaluation, in the context of the current framework this sonification could be deemed not strictly adherent, but at least compatible with the framework, as the soundscape is artificial (in this artificial task domain understandable) but still adheres to the guidelines around mapping (unsurprising as one-dimensional data points in a two-dimensional space are mapped according to a pointer into the data).

5.7.5 ChiSon (Chapter 4)

In ChiSon, the intention was to sonify data about the flexibility of parts of the protein. This happens with a parameter-mapping sonification employing pitched sine waves.

Comparison to the steps outlined above

Task/Data domain: Structural protein data, where the flexibility of the protein is the focus of the task. The task is not entirely clear, as structure-based drug design is not a straightforward task. Another motivation is that the software does not support simultaneous visualisations of as many data dimensions as are available.

Soundscape: A mix of different metaphors and techniques is presented as parameter mapping sonification of a sine-like tone with a percussive envelope is employed to represent the flexibility values, where pitch is mapped to flexibility, whereas hydrogen bonds are represented with an earcon designed to sound like a science-fictional space ship docking to something. This bit could have been part of a bigger science-fiction space opera soundscape, but in isolation doesn't work that well. The soundscape is very abstract and not easily associated with the task.

As surfaced in the evaluation, no association of proteins and sound was in the participants' minds and the very abstract nature of the task does not lead itself easily to a soundscape.

Find and map: One aspect of the mapping is the spatialisation depending on the zoom level and position of the viewer. The user can therefore choose which part of the protein they focus their attention on. Coupled with the on-demand sonification by clicking on parts of the structure gives the protein a virtual physical form and presence in space. The mapping of pitch to flexibility is reasonable in that in e.g. string instruments, a thicker string can not represent high frequencies as easily as a thin string (compare violin to contrabass strings). As the task is not entirely well defined though, the resultant sonification has no clear structure either.

Optionally: map more : No more mapping was applied.

Conclusion

The mapping in this case again uses sine waves in lieu of a soundscape, as the data provides no obvious clues to what soundscape could be employed. Considering the nature of the data a tree metaphor with creaking branches or some other structure with inherent flexibility and sound could be considered, but it is hard to imagine how data mapped to that soundscape would be perceptible to the user. More research into potential metaphors is needed, but as the participants in the evaluation reported, the current mapping is not very transparent.

5.8 Conclusion

This chapter described a novel design framework for sonification. Additionally, its application to the analysis of sonifications was demonstrated.

In building on Clarke's ideas applied to sonification by Vickers, Hogg, and Worrall, a new design framework is created, which addresses some of the shortcomings in the design processes outlined in Chapters 3 and 4: a structure and process is proposed that can lead to sonifications which address the requirements of the task more closely. This is achieved by focussing on finding the most important data to map to, and mapping it to a soundscape that contains a subject position that is conducive to fulfilling the task. The focus on finding a soundscape instead of creating a mapping to sounds from scratch provides a different approach to a lot of sonification design currently conducted. The attention to clear mapping might be the obvious choice, but this focus is sometimes lost in the vast possibility space of sound synthesis and reproduction.

The proposed framework advises to only map more after the main mapping is clearly worked out. Mapping more data is encouraged though as it can lead to a more useful sonification for long-term users, providing opportunity to discover more aspects of the data and improve task performance over time.

The analysis of sonifications is proposed as another aspect of the framework, put into practice in the analysis of five sonifications (including the ones presented in the previous two chapters). It was useful in further refining the framework, and could inspire new ways of thinking about existing sonifications, reflecting on them in a way that can lead to more insight into where sonification as a broad field could be heading. The analysis of the speed-skating sonification and its comparison to the documented real-life evaluation showed a good overlap between the predicted and observed outcomes. The analysis of the image and protein sonifications provided reflections for future work in these areas.

Its application to a concrete project and the resulting evaluation are outlined in the following chapter.

Chapter 6

SoniFRED

6.1 Introduction

A study to explore a new theory to alleviate back pain was being conducted at Northumbria University, which postulates that atrophied small muscles in the back are responsible for it. These muscles atrophy because of our sedentary life style. Researchers found out about this phenomenon by studying astronauts coming back from long missions in space having back pain even though they were keeping an intense exercise schedule on their missions.

The study used a new exercise machine for physiotherapeutic use called FRED, which stands for Functional Readaptive Exercise Device [127]. The goals of its developers are to engage the key spinal and abdominal muscles *lumbar multifidus* (LM) and *transversus abdominis* (TrA), which are suspected to be causing lower back pain when atrophied. The original FRED research is described in [31]. The device itself is a modified cross-trainer offering minimal resistance (see Fig. 6.1), meaning the user is standing on an unstable base in comparison to the resistance normally experienced on a cross-trainer. Users are required to walk slowly on this device standing freely without using their arms and hands to support themselves. In order to walk forward as slow as scientifically recommended (ca. 0.4 Hz), they need to use the rear heel to counterbalance the pull of gravity and control the descent of the front foot. The aim is to operate FRED while sporting an upright posture and walking in a smooth, controlled manner with minimal variations in speed.

The only feedback provided in the study was a visual output of the current speed as a chart on a computer screen. The user had to try to get as close as possible to the centre line and aim to keep the plot line as flat as possible.

Upper and lower boundaries were also shown, which the users were supposed to stay inside of. A lower boundary of 0.2 Hz as well as an upper boundary of 0.6 Hz can indicate a loss of control: too much as well as too little pressure on the rear heel can cause cessation of movement or a too fast movement forward, indicating a loss of the proper balance required for the exercise.

These boundaries also enable the user to display some variability in movement speed, which will become less in the course of the exercise. To measure this progress, the physiotherapists determine a maximum deviation from the individual mean as a target rate based upon their current ability and any physical characteristics that might impact upon how well they are able to use FRED. For example, a novice might expect to reach a standard target deviation of 0.15 Hz while someone who is able to keep within the range of, say, 0.35 Hz $\leqslant f \leqslant$ 0.45 Hz would have a target deviation of 0.05 Hz. Once the therapist has determined a patient's target deviation it is interesting to know at what points they fail to achieve it.

Users report that walking in the prescribed way on the device is very difficult at the start as the movement involves putting some pressure on the back foot as well, and gets gradually easier with exercise.

This way of delivering feedback has certain drawbacks: both the user as well as the supervising therapist have to keep their eyes on the monitor constantly; lack of visual acuity or visual fatigue in the user can deteriorate exercise performance; and unfamiliarity with graphical plots can lead to confusion.

Auditory feedback could enable the therapist to monitor the performance of the user as well as making notes or attending to more than one user at the same time.

This presented a supreme use case for sonification, as this would give both users and supervisors a better range of movement and maybe additional engagement in the exercise. The researchers on the study were happy to collaborate on implementing an additional form of feedback.

Therefore, it was agreed to sonify the FRED data by providing feedback on the user's instantaneous walking speed, their variability, and each time they exceed 0.6 Hz or drop below 0.2 Hz.

An offline version working with recordings of the data from the study was implemented by Vickers. It involved a sonification using sounds related to the sea shore. Another sonification was developed later and a live version of both sonifications taking data directly from the device was implemented by the author.

Both sonification designs aim to find a mapping that corresponds to the feelings and attitudes to evoke in the user as outlined below and in Chapter 5.

In the following sections, relevant related projects will be discussed, followed by a discussion of the process of developing the SoniFRED system and its sonifications[1]. Then

---

[1] The system is available at `https://github.com/nuson/SoniFRED`.

Table 6.1 Sonification Metaphors

Indicator	**Beach (trad.)**	**Train (SPSon)**
Speed	Frequency of waves, smoothed and clipped at range boundaries	Chuffing frequency, directly coupled to wheel position; train wheels on tracks
Too fast	Foghorn	Train brake
Too slow	Ocean buoy bell	Train whistle
In the ‘zone’	Seagulls	Subtle tunnel reverb

the evaluation and its methodology is discussed, followed by its results. A discussion of the results and a summary conclude this chapter.

6.2 Background

Sonification as feedback for exercise has been implemented in a variety of ways. Sigrist *et al* [105] provide an extensive review of feedback for motor learning.

Stienstra *et al* [109, 110] describe a movement sonification as feedback for speed-skaters. The aim of their Augmented Speed-skate Experience is to give real-time information about relevant parameters (e.g., the pressure on the foot) to long-track circuit speed-skaters (a detailed analysis is provided in Chapter 5, Subsection 5.7.2).

Real-time feedback in general can help users with improving and/or correcting their technique while performing an exercise. E.g., gymnasts training on a pommel horse benefitted from research by Baudry *et al* [14], who demonstrated that real-time augmented auditory feedback can be used to correct complex movements and improve retention.

Eriksson *et al* [40] developed a visual and auditory feedback system for runners to improve the retention of a modified running technique on a treadmill. Both auditory and visual feedback was tested in an experiment and in almost all subjects both feedback conditions resulted in an adjustment of their technique.

Frid *et al* [44] examined the use of haptic and auditory feedback in a task where subjects had to virtually throw an object via a point-based haptic device. They found a lower error rate in the audio-visuo-haptic condition when using sonfication based on a physical model synthesis of friction sounds compared to the visual-only condition. They also report an increase in ‘perceived intuitiveness’ in most conditions that included haptic and/or auditory feedback.

Fig. 6.1 Functional Readaptive Exercise Device (FRED): a modified cross trainer which offers minimal resistance requiring smooth control to recruit muscles in the lower back and abdomen.

Hurkmans *et al* [62] provided evidence that partial weight-bearing training with auditory feedback for patients after total hip arthroplasty resulted in better performance than without it.

In rowing, Sigrist *et al* [106] evaluated multimodal audiovisual and visuohaptic feedback in comparison to only visual feedback. In the audiovisual group, the oar movement was sonified in addition to the visual feedback. This condition resulted in better learning of the movement profile than in the visuohaptic condition. Dubus [36] also developed sonifications of rowing and evaluated them with the Swedish national rowing team [35], which showed that participants could efficiently determine basic characteristics of the data, but a balance between form and function in interactive sonification has to be aimed for in future research.

Ramezanzade *et al* [89] showed test subjects movement patterns of skilled basketball players. The visual group only saw the movement pattern, whereas the audiovisual group also heard a sonification of the elbow's angular velocity. When the subjects were asked to repeat the patterns themselves, analysis of the video material showed that the audiovisual group fared better in all except one test (“total time of elbow motion”).

Mezzarobba *et al* [76] showed videos of movement to Parkinson’s disease sufferers. One group were shown the videos with movement sonification, the other without. Only the group with movement sonification improved, suggesting that multisensory approaches could help patients to relearn gait movements and reduce freezing episodes.

This review indicates that real-time auditory feedback can help improve movement accuracy [89, 76] as well as with learning a new movement [106, 44, 14]. Dubus [35] points out that aesthetics should be taken into account. No directly comparable auditory feedback systems for exercise devices as the one described herein were found.

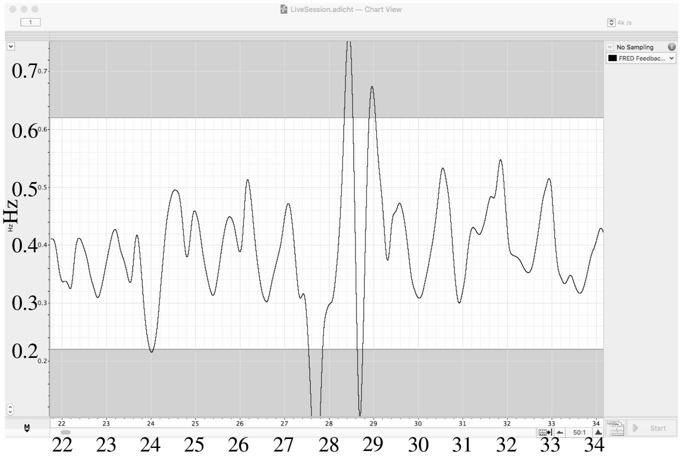


Fig. 6.2 Screen capture of the LabChart live scrolling window presented to FRED users (axis labels have been superimposed here to aid the reader).

6.3 Implementation

FRED is instrumented with a rotary encoder (RP6010, ifm Electronic GmbH, Essen, Germany [2]) with a resolution of 1000 pulses per rotation and three outputs, which are used to measure user performance: outputs 1 and 2 send a pulse for every 1/1000th rotation of the wheel. We only used the output of output 1 (in the following called the ‘tick sensor’), as output 2 is a 90°phase-shifted version of output 1. Output 3 sends a pulse each time the wheel completes a full rotation (the zero index, in the following called “top sensor”). With these, the movement speed and the exact position of the wheel can be determined relatively precisely (to 1/1000th of a rotation, provided the user doesn’t change direction of movement). The pulse stream

---

[2] See `https://www.ifm.com/gb/en/product/RP6010` for more information.

from the sensors is sampled at 4 kHz into LabChart [3]. The data is converted to frequency values (i.e., revolutions per second) for ease of display for the user (a patient). The resultant data stream is then smoothed using a triangular Bartlett filter. The feedback given to the user is the LabChart output of the tick sensor smoothed with a Bartlett window to remove the steps in the data presented at zoom level of 50:1, i.e., ca. 12 seconds are displayed at once (Fig. 6.2).

The display shows the optimal performance zone (the white area in Fig. 6.2) which, for the FRED configuration used here, is a speed of 0.4 Hz ± 0.2 Hz. The user is instructed to try to keep the plot line inside the white area. Previous work has determined that maintaining a speed of $0.2\text{ Hz} \leqslant f \leqslant 0.6\text{ Hz}$ leads to recruitment of the LM and TrA muscles thus producing therapeutic benefit, with optimal benefit occurring at $f = 0.4$ Hz [127].

The user can see how they are doing as the display updates at 20 Hz (every 50 ms). In non-expert users, the plot line varies a lot as each step has its unevenness due to the unconventional movement style required for smooth operation of FRED — users must shift weight to the rear heel to balance out the gravity-assisted forward step. Therefore, a curve such as is shown in Fig. 6.2 is generated. If a user were able to operate FRED perfectly there would be no variation in speed and the plot would show a flat line at 0.4 Hz.

6.3.1 Beach

Vickers came up with the idea to augment FRED with sonification. Considering the slow tempo, Vickers designed his sonification around a calming soundscape: waves and other sounds found on a beach (see `https://github.com/nuson/SoniFRED`). As mentioned above, the movement is very slow and unconventional, requiring concentration of the user and a calm, smooth movement.

As indication of the speed he used waves breaking on the shore. In his sonification, every two steps a wave breaks (ideally 0.4Hz/4 = 0.15Hz). This is slightly slow for your regular beach, but experiments showed that a wave a step was making for a very choppy sea. The wave speed is changing slowly according to the speed of the user, over the course of a second, and always stays inside the ideal range, acting both as calming feedback, and as reminder of the nearest in-range speed. This could be seen as a form of an adaptive metronome feedback.

If the user leaves the ideal zone, other sounds get triggered: too fast, a foghorn is sounding, too slow, the bells of a buoy can be heard. Each sound is triggered and then a recovery period of ≈1.5 s is triggered, in which no further warning sound takes place.

A running standard deviation is calculated and when it falls under a specific threshold, indicating the user is operating with very smooth speed, the sound of seagulls is added to the mix as an indicator. It stays in the mix until the standard deviation climbs higher than the threshold.[3]

6.3.2 Bike?

It was decided that another sonification design should be attempted, designed after the idea of the subject position based design framework (SPBDF) described in chapter 5.

Looking at activities that provide sonic feedback that can be understandable to the user and useful in its information content,cycling was chosen as an activity of leisurely transporting oneself on old, badly maintained bikes with internal-gear hubs on completely flat terrain as found in the Netherlands. These bikes, due to their gearing type and the way they were maintained, provide very exact feedback of cycling speed and suggest a relaxed transport experience.

The primary sound showing the speed would have been the ticking of the gear hub, the tires on the tarmac and the chain, as well as maybe a creaking crank. Other virtual cyclists at the ideal speed could be positioned around the user as a way to show if they are too slow or too fast. This could include sounds of their bikes, clothing, and breathing, and bicycle bells being rung as a warning sign before overtaking.

After discussing the idea with the FRED researcher, it was determined that this activity is not very applicable to a UK setting (as inner city cycling can be very stressful and cycling is often associated with sport and middle-aged men in lycra jumping red lights or roaming the countryside), and would not be understandable to the average user in North-East England, as well as being hard to perceive if the sounds are too subtle. Additionally, the movement profile is markedly different from the elliptical exercise device where users are supposed to dig their heels in.

6.3.3 Train

Kirsty Lindsay provided her idea of a steam train slowly chuffing up a slope. Though not a human motion, it provides amble opportunities for feedback sounds, and made sense as the design, look, and function of the exercise device can evoke parallels to trains, with its wheel being driven by a more vertical motion of the human on the device, in a similar way as a steam locomotive translates steam pressure into sideways motion and then into circular motion.

The image of a slow uphill steam train ride also was deemed a relaxed situation, for a majority of people only ever being experienced in a holiday setting and distinct enough from

---

[3] See `https://youtu.be/BDAyBtCzRck` for a video of this sonification with recorded data.

a train in full flight, which would be an image we want to avoid evoking here as the device needs precise, slow movements.

Whereas the beach sonification does not directly map the movement of the user to the wave sound, which would possibly disrupt its relaxing atmosphere, here the user directly *drives* the train, providing a more direct route to feedback about the regularity of their motion patterns.

The chuffing of a steam train is caused by exhaust steam, which is lead through a nozzle, the blast pipe, into the chimney. Each cylinder produces two chuffs, and normally a locomotive has at least two cylinders, which act a quarter of the wheel out of phase, so we would have four chuffs per rotation [1]. A comparison to videos of various steam trains[4] showed the chuffing frequency to be much higher to the one intrinsic to the speed of our "train wheel" at 0.4Hz (0.4Hz * 4 = 1.6Hz), but experimenting with higher speeds of 8 chuffs and 16 chuffs per rotation showed that these speeds resulted in a change of mood, as with the choppy sea described in the beach sonification design process, resulting in a soundscape that was too hectic.

The chuffing sound is directly coupled to the position of the wheel on the exercise device, so users — if they don't go backwards during one cycle — receive feedback at the same points in their walking cycle every time. This direct coupling is meant to lead to a smoother workout. The direct coupling causes irregularly spaced chuffs in inexperienced users, so a more smoothed out chuffing decoupled from the position of the wheel was discussed, but ultimately not implemented as with prolonged use the direct coupling was deemed to be more useful as feedback than a more pleasant experience at the start of the training.

Other sounds coupled with a steam train are: the whistle, used as warning sound or other signal, as it is possible to vary its output; the sound of the brakes; and the sound of the wheels on the tracks.

As the brakes indicate a slowing down of the train and the whistle can indicate a warning of some sorts, e.g., of a train approaching a level crossing, these sounds were chosen to act as warning sounds for leaving the zone. There was some discussion of the polarity of the warning: would the brake indicate the user was too fast and therefore should SLOW DOWN, or the user was "braking" and therefore too slow and should go faster? The same with the whistle - was the user going too fast and therefore the whistle is a warning to SLOW DOWN or for others to get out of the way, or was the user going too slow and should SPEED UP. We decided to go with the latter in the end.

---

[4] E.g. `https://youtu.be/fXcD6ZdPR9k?t=186`.

As feedback for being in the zone of desirable speed, a room simulation (reverb) was added after mixing all the above signals together. A bug in the code meant that this part of the feedback wasn't used in the study.[5]

6.3.4 Data input

Both sonifications were implemented in Max/MSP 7. As data input, Paul Vickers's sonification used static data files with a lower sampling rate provided by the FRED researcher.

Separate static data file reading facilities were implemented for the second sonification, which used the raw sensor data at the full sampling rate of 4000Hz and implemented all the data processing done in LabChart in Max 7. This enabled the integration of live data gathering into the sonification. This system produces the same smoothed curve output as the one visible on the live feedback. This was important as other forms of smoothing (essentially low-pass filtering the signal) would lead to other distortions in the output than the triangular (Bartlett) windowing used in Labchart, and could lead to out-of-zone signals being triggered at moments not visible in the Labchart output or vice versa, decoupling the visual from the sonic feedback.

To get a feed of the data from the machine, Prof Nick Caplan produced a BNC-to-Audio Jack adapter, that enables the raw electrical signal from the top and tick sensors of the exercise machine to be fed into two input channels of a standard external sound card used in music production, resulting in an uncomplicated setup for the final test and study.

The position of the wheel is computed from the raw top sensor and the tick sensor data and used in the train sonification for the chuffing and train wheel sounds.

The raw signal is then downsampled for performance reasons from the sound card's input sample rate (in our case 44100Hz as the same sound card was also used for sound output and the sample rate had to be identical for in- and output) to 4410Hz (1/10 of the audio sampling rate and roughly the same rate as used in LabChart, 4000 Hz in this case) and convolved with a 0.15s Bartlett window.

6.4 Study Methodology

A study was conducted with 45 volunteer healthy adult participants under 65 over a period of one week in Spring 2017 at the Centre for Life, Newcastle upon Tyne (UK), a science museum (A local newspaper report with video footage can be found at [47] and a video produced by the university is available at [83]). The study was designed in collaboration

---

[5] See `https://youtu.be/tFubibg4b80` for a video of this sonification with recorded data.

with Kirsty Lindsay and her supervisor Prof Nick Caplan. They designed the exercise routine and the way feedback was recorded. Lindsay, Caplan, Vickers and the author decided collaboratively on the different feedback conditions. The feedback scores were collected by Lindsay, Vickers, Caplan, and the author. The quantitative data from the FRED device was collected by Lindsay. The questionnaires were transcribed by Lindsay.

After a 3-minute warm-up with instructions from a researcher on the machine, participants were asked to use the functional reactive exercise device (FRED) with eight different feedback conditions in randomised order. Each feedback condition was used for one minute with a 30 second rest period in between the conditions. The eight feedback conditions comprised of visual-only feedback, metronome-only, metronome with visual feedback, music with visual feedback (where the same recording was used for every participant - a pop song chosen for being faster in tempo than the target tempo), the two sonifications with visual feedback, crowd noises with visual feedback (a recording of a coffee shop) and no feedback at all. Sonification without visual feedback was not tested as the sonification was meant to be used in addition to the visual feedback and the number of different conditions tested was already quite high.

The metronome condition was experientially determined best to be set at four times the speed of a cycle (i.e., 96 beats per minute = 0.4Hz $\times$ 4 = 1.6 Hz), i.e., two beats per step.

Participants wore closed professional headphones (Beyerdynamic DT 100) for all feedback conditions involving sound. Visual feedback was provided on a computer monitor on eye level for a 1.70m tall person (not adjustable), ca. one meter from the subject, in front of the exercise device. After each condition, participants were asked to rate their experience on a scale from 1 to 10, where 1 is "I hate this / not at all helpful" and 10 "I love this / completely helpful" using a visual analogue scale (see Appendix E.1). Performance data as well as experience scores were recorded in LabChart.

Each participant was asked to fill in a questionnaire before and after using the FRED device (see Appendix E for the questionnaires and Appendix D for the consent form). The author designed question 1.3, 1.4, 1.5 on Questionnaire 1 and part 2 of the second questionnaire.

The research questions for this study were:

1. How effective is sonified auditory feedback at helping participants on the Functional Re-adaptive Exercise Device achieve the target movement frequency, in comparison to visual feedback and simple auditory guidance (metronome)?

2. What is the effect of "directly related mapping" sonified auditory feedback, in comparison to "unrelated mapping" sonified auditory feedback, on the ability of participants to achieve the target movement frequency?
3. What is the effect of background environmental noise (e.g. music) on the effectiveness feedback in helping participants achieve the target movement velocity?

6.5 Study Results

Preliminary analysis of the recorded quantitative data was done by Kirsty Lindsay and replicated by the author. The author conducted all qualitative data analysis and the ANOVA below.

Reported in this section are all results related to the sonification conditions in comparison to the other feedback conditions. A detailed analysis of the remaining data will be published by Kirsty Lindsay. The method she used to evaluate the quantitative results [13], which is more widely used in sports science contexts, revealed no meaningful difference in result between the feedback conditions. The measure of meaningful difference enables a more practice-relevant analysis, as it indicates in this case if the feedback actually improved or diminished performance. The result indicates that feedback had no influence on the performance of the participants.

In the following sections, standard statistical tools (repeated-measure ANOVA) are used to conduct a comparison of the sonification feedback conditions to the other feedback conditions and between themselves. These provide a different measure of significance and might not indicate a meaningful difference from a sports science perspective, but might still indicate a tendency that is helpful in the further development of sonification in this area.

Afterwards, the results from questionnaires filled out by participants are presented and analysed.

The analysis, tables, and graphs in this section were created using IBM SPSS.

6.5.1 Per-condition results

For each participant and feedback condition the mean frequency and the standard deviation was calculated. The results of an repeated-measure ANOVA conducted on the mean frequency, standard deviation and score are reported in the following.

Mauchly's test indicated that the assumption of sphericity had been violated for all 3 measures: $\chi^2(27) = 43.291, p = .05; \chi^2(27) = 123.693, p = .0; \chi^2(27) = 276.826, p = .025$. Therefore, Greenhouse-Geisser corrected tests are reported ($\varepsilon = .797; \varepsilon = .527; \varepsilon =$

.250). The results show that the score was significantly affected by the feedback condition, $F(5.579, 234.317) = 8.531, p = .000, \omega^2 = 0.102$. The mean frequency was also significantly affected by the feedback condition, $F(3.686, 154.825) = 2.862, p = .029, \omega^2 = 0.018$. The standard deviation of the frequency was not significantly affected, $p > 0.05$.

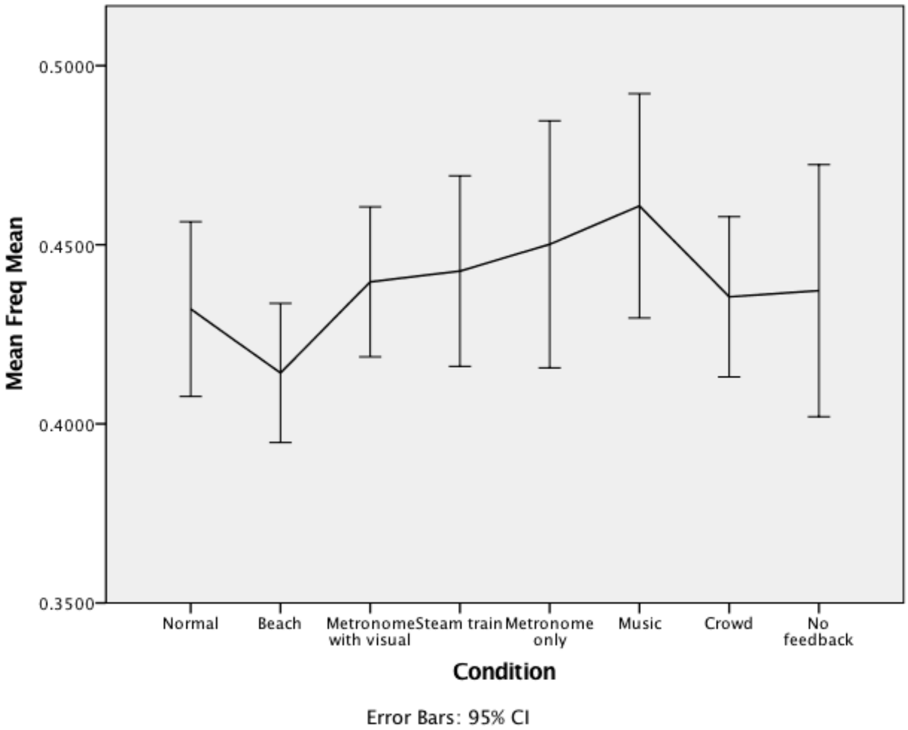


Fig. 6.3 Mean frequency mean

Planned comparisons for the mean frequency revealed that music feedback resulted in a significant increase in comparison to normal, sonified (beach or train), and metronome feedback (with and without visual feedback), $F(1,42) = 7.821, p = .008, r = .40$. Therefore music feedback made subjects go faster than normal, sonified or metronome feedback.

Furthermore, metronome with visual feedback also significantly increased the mean frequency in comparison to the sonified feedback conditions (beach and train), $F(1,42) = 4.112, p = .049, r = .30$. Subjects listening to the metronome feedback went faster than subjects on sonified feedback.

Looking at the interaction graph, the beach sonification lead to a significantly lower mean frequency than the train sonification, $F(1,42) = 11.458, p = .002, r = .46$.

Planned comparisons for the score revealed that the no feedback condition scored significantly lower than the other conditions, $F(1,42) = 12.667, p = .001, r = .48$.

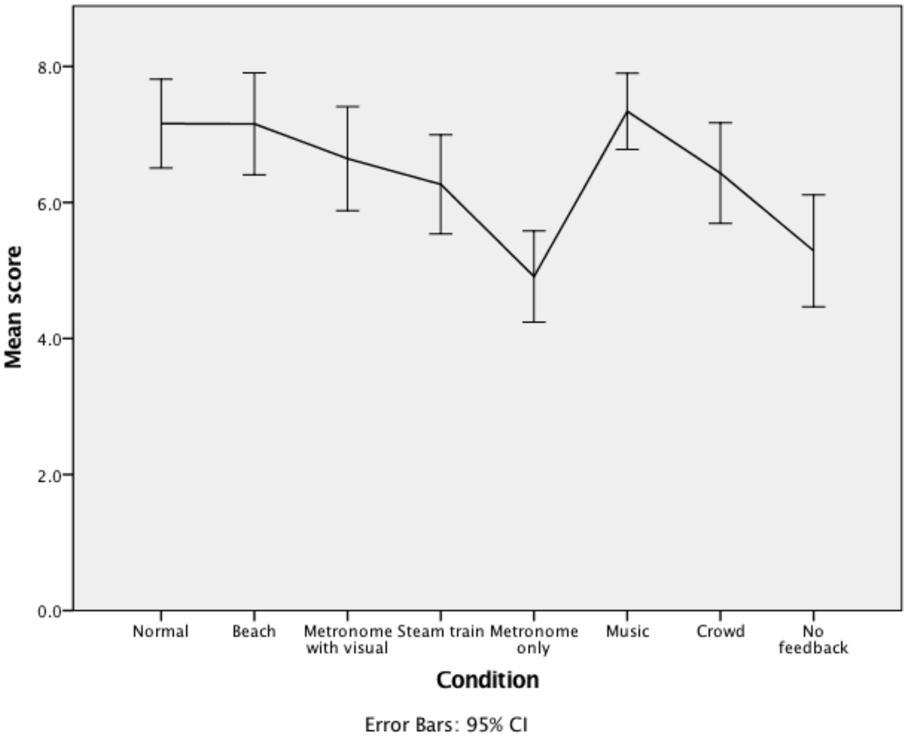


Fig. 6.4 Mean score

The music feedback condition scored significantly higher compared to the actual feedback conditions (conditions 1–5), $F(1,42) = 8.15, p = .007, r = .40$.

Metronome only feedback scored significantly lower compared to normal, sonified, or metronome with visual feedback conditions, L4: $F(1,42) = 29.08, p = .000, r = .64$.

Finally, the beach sonified feedback scored significantly higher compared to the steam train sonified feedback, L7: $F(1,42) = 5.33, p = .026, r = .34$.

6.5.2 Results from questionnaires

Only one of the participants (2.2 %) encountered sonified feedback/sonification/auditory display before taking part in this study. 46.7% of participants reported some musical training, though 26.7% of participants reported that the musical training happened more than 10 years ago.

When asked about their preferred type of feedback, where multiple answers were permitted, 29% preferred the Beach Sonification or the "Music with visualisation" option, 20% the Metronome with visualisation, 13.3% the Crowd noise with visualisation, 11.1% visualisation only, 8.9% reported no preference, 6.7% the steam train sonification, and 4.4% the metronome only option (see Table 6.5).

When asked which of the two sonifications they preferred, 61% of participants said they preferred the beach sonification, 13.3% preferred the steam train sonification, whereas 20% said they didn't like either and 4.4% liked both (see Table 6.2).

Sonification	Frequency	Percent
Beach	27	61.4
Both	2	4.5
Neither	9	20.5
Train	6	13.6

Table 6.2 Preferred sonification

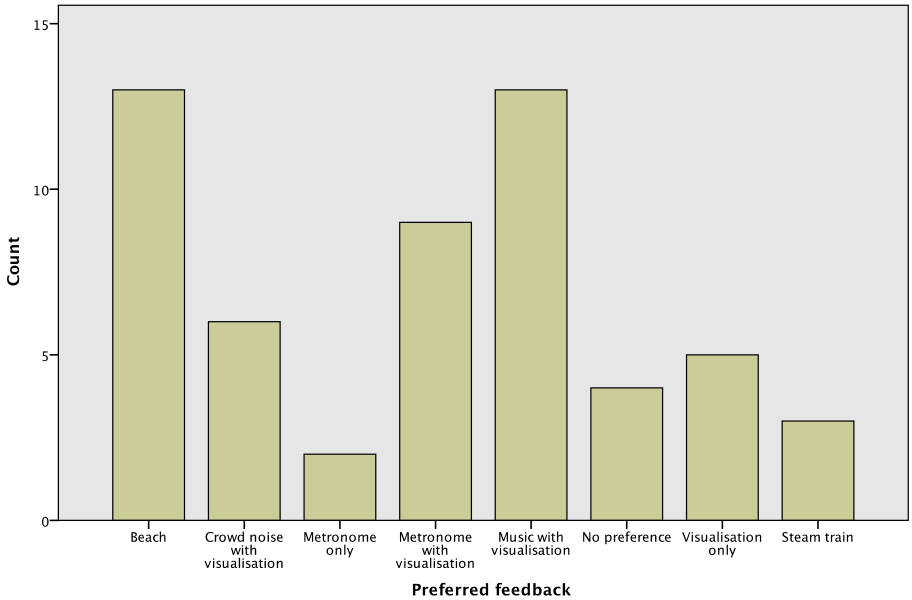


Fig. 6.5 Preferred feedback

When asked to rate the sonifications on a 5-point Likert scale in the categories helpful for pace, helpful for motivation, annoying, calming, and appropriate for the activity, the beach sonification was deemed calming by 80.5% of participants, whereas only 15.7% found the steam train calming. The beach sonification was also rated more positively in all other categories (see Table 6.3 and Figure 6.6).

To gage future directions or avenues for research into sonified feedback in exercise devices, we asked if and what other types of exercise devices would benefit from sonified

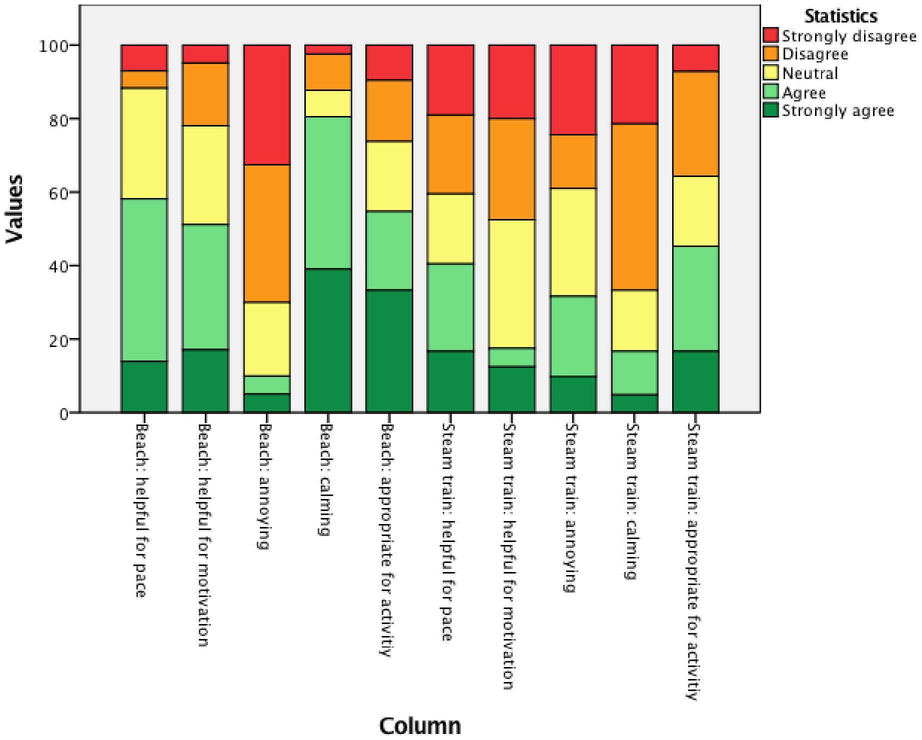


Fig. 6.6 Likert scale results

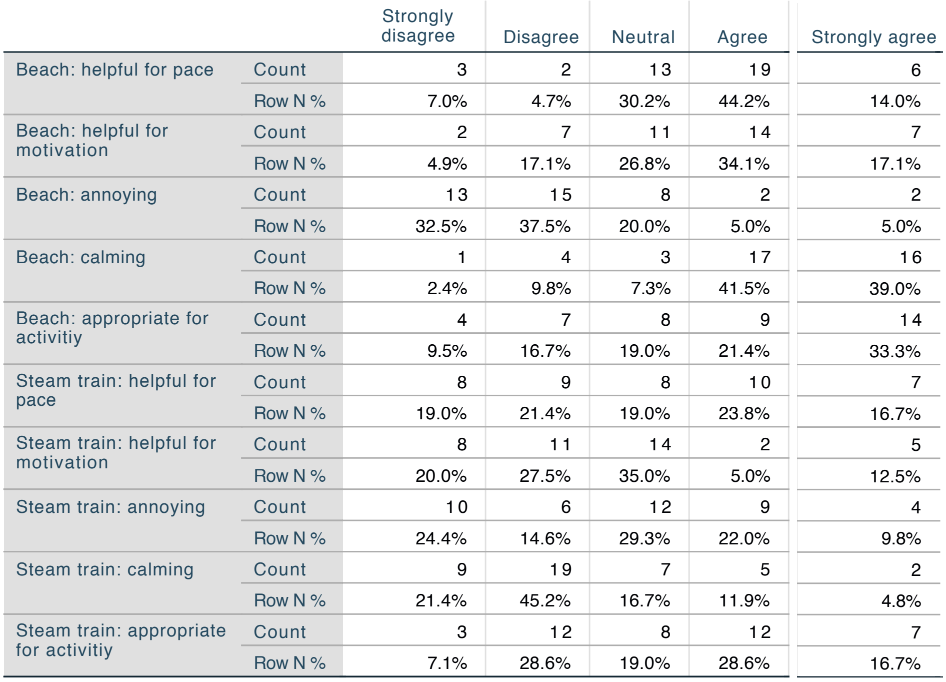
feedback. 35.6% of participants disagreed, whereas the other participants agreed and gave some suggestions, including bikes, wattbikes (a brand of exercise bikes) or cycling (9 mentions), rowing (6), treadmill or running (4), cross-trainers (2), golf swing, smith machine (a type of weight machine), swimming (all one mention) (see Table 6.4).

6.6 Discussion

The music feedback as well as the metronome with visual feedback resulted in a higher mean frequency. This is not surprising as the chosen track had a higher tempo than the target frequency and the metronome feedback can be confusing for people not used to it.

The track chosen seems to have resonated with the participants, as it scored higher than the actual feedback conditions (conditions 1–5). This could indicate that the sonification conditions have to work on their sound design to get to the level of a pop song.

The beach sonification was the feedback condition with the lowest mean frequency as well as a significantly lower mean frequency than the train sonification. This could be related to the more soothing soundscape chosen. In this context, a more relaxing atmosphere created could lead to a better result for this specific exercise.



		Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Beach: helpful for pace	Count	3	2	13	19	6
	Row N %	7.0%	4.7%	30.2%	44.2%	14.0%
Beach: helpful for motivation	Count	2	7	11	14	7
	Row N %	4.9%	17.1%	26.8%	34.1%	17.1%
Beach: annoying	Count	13	15	8	2	2
	Row N %	32.5%	37.5%	20.0%	5.0%	5.0%
Beach: calming	Count	1	4	3	17	16
	Row N %	2.4%	9.8%	7.3%	41.5%	39.0%
Beach: appropriate for activitiy	Count	4	7	8	9	14
	Row N %	9.5%	16.7%	19.0%	21.4%	33.3%
Steam train: helpful for pace	Count	8	9	8	10	7
	Row N %	19.0%	21.4%	19.0%	23.8%	16.7%
Steam train: helpful for motivation	Count	8	11	14	2	5
	Row N %	20.0%	27.5%	35.0%	5.0%	12.5%
Steam train: annoying	Count	10	6	12	9	4
	Row N %	24.4%	14.6%	29.3%	22.0%	9.8%
Steam train: calming	Count	9	19	7	5	2
	Row N %	21.4%	45.2%	16.7%	11.9%	4.8%
Steam train: appropriate for activitiy	Count	3	12	8	12	7
	Row N %	7.1%	28.6%	19.0%	28.6%	16.7%

Table 6.3 Likert scale results

In the questionnaires the beach sonification was generally rated more positively than and preferred to the steam train sonification. This reflects the effect observed in the quantitative data gathered: here, it was scored significantly higher than the other sonification.

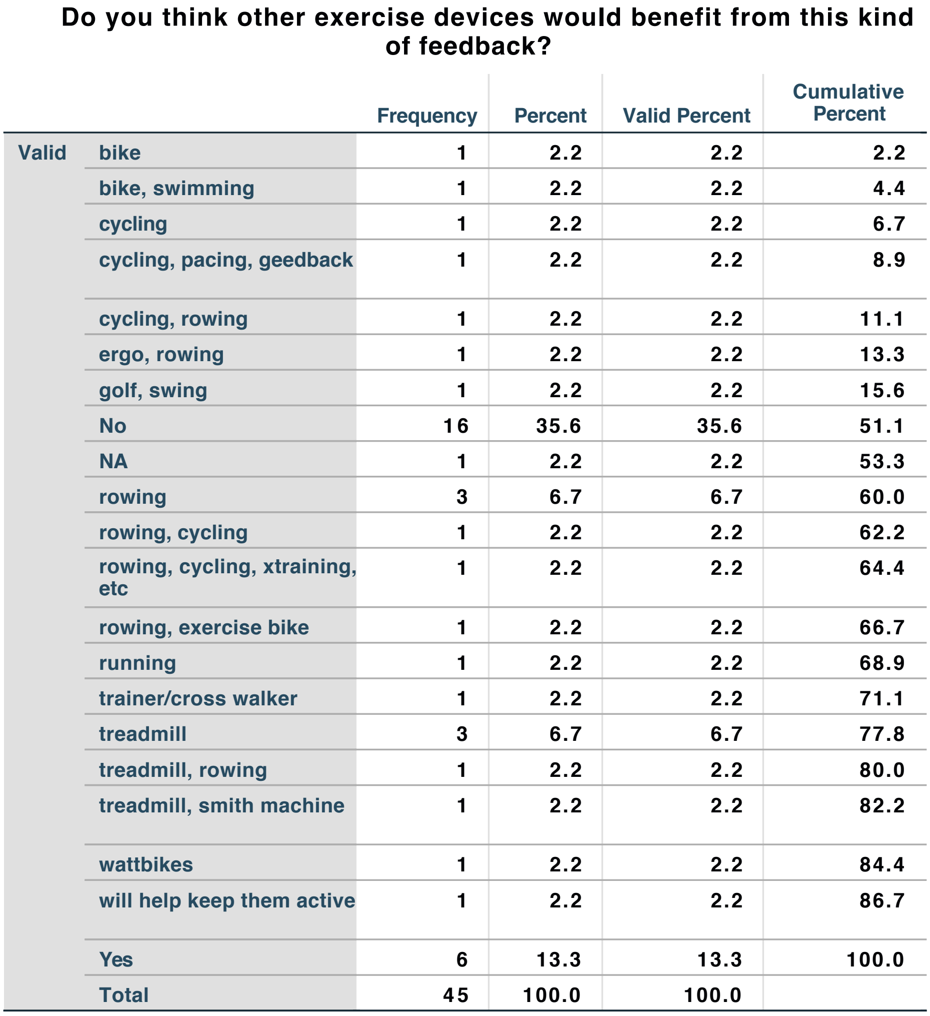
Only 2.2 % (one participant) had encountered an auditory display before the study, highlighting that the general public seems to be unaware of sonification as feedback. It is encouraging that a majority of participants (64.4 %) could imagine it being used more broadly in other exercise devices.

To conclude, the beach sonification showed promise as it addressed the need for a specific mood to facilitate the correct usage of the exercise device, whereas the steam train sonification might have been a more fitting metaphor but did not provide the same amount of calmness.

6.7 Conclusion

This chapter described sonified feedback for a novel exercise device. It provided an overview of other feedback systems. It outlined the process of sonification design using the SPBDF. A study was described and its results interpreted.

FRED is an exercise device to alleviate back pain. It only provided visual feedback in the form of a line graph display. The proposed SoniFRED system extends the capabilities of



Do you think other exercise devices would benefit from this kind of feedback?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	bike	1	2.2	2.2	2.2
	bike, swimming	1	2.2	2.2	4.4
	cycling	1	2.2	2.2	6.7
	cycling, pacing, geedback	1	2.2	2.2	8.9
	cycling, rowing	1	2.2	2.2	11.1
	ergo, rowing	1	2.2	2.2	13.3
	golf, swing	1	2.2	2.2	15.6
	No	16	35.6	35.6	51.1
	NA	1	2.2	2.2	53.3
	rowing	3	6.7	6.7	60.0
	rowing, cycling	1	2.2	2.2	62.2
	rowing, cycling, xtraining, etc	1	2.2	2.2	64.4
	rowing, exercise bike	1	2.2	2.2	66.7
	running	1	2.2	2.2	68.9
	trainer/cross walker	1	2.2	2.2	71.1
	treadmill	3	6.7	6.7	77.8
	treadmill, rowing	1	2.2	2.2	80.0
	treadmill, smith machine	1	2.2	2.2	82.2
	wattbikes	1	2.2	2.2	84.4
	will help keep them active	1	2.2	2.2	86.7
	Yes	6	13.3	13.3	100.0
	Total	45	100.0	100.0	

Table 6.4 Other devices that would benefit from sonified feedback

FRED to include sonic feedback. It hooks directly into the raw data stream from the device and provides a compact system that could be deployed relatively easily to support both the users and the therapists in their task — providing both process monitoring for therapists as well as feedback for users.

A study showed that in comparison to other feedback conditions the beach sonifications increased task performance by decreasing the mean frequency of the participants. This sonification was not designed according to the principles of the framework proposed in Chapter 5, but still adhered to an ecological approach to sonification as outlined above. Its calming soundscape could be the reason for the decrease in mean frequency, which would point to increasing the focus on the mood of a soundscape.

The train sonification did not have an effect on task performance. As its direct chuffing feedback was designed to appeal not strictly to beginners, a more thorough long-term evaluation of the train vs. the beach sonifications would help to explore their long-term performance and could lead to improvements in SPBDT in case of negative results.

Chapter 7

Discussion

The previous chapters outlined three interactive sonification systems, their development and evaluation. Whereas the development of a system for image sonification described in Chapter 3 was meant to be well integrated into another project, organisational difficulties meant not enough access to domain experts.

In all the scenarios outlined, even when sonification systems were documented and available for the task domain, oftentimes they came without sound examples or code. This seems to be a general problem in sonification literature, probably exacerbated by the dominance of publishing formats that don't allow embedded content and the corresponding ephemerality of the world wide web, where university websites are restructured without thinking about old links.

In the context of Chapter 3 this led to a system that was designed around knowledge gained from literature and second-hand information. The task at hand was reasonably defined — highlight useful data in the images for the pathologist — but without a pathologist at hand, it is easy to make mistakes about what is useful data. The study of existing systems in image sonification and interactive sonification of medical data provided a blueprint around how the interaction could work. The developed prototype had the required functionality and used Python for feature extraction, Max for displaying the user interface and Supercollider for sound synthesis, therefore being not easily portable to other computers as three different software ecosystems have to be installed. The interface devised, a simple touch screen interaction with the loaded image, where users can touch parts of the image to highlight parts of the sonification, enabled interactive exploration of the images. The ideas in the system around using auditory glances but integrating them into the general interactive sonification mapping, stood up to evaluation in as far as they were deemed interesting, but not useful. The development of a system with a task description but without access to a user seems to be doomed to miss the mark.

The system presented therefore extends previous research in using a glance to preview the image, deriving both an interactive sonification and an overview from the same mapping, as well as trialling new ways of directing user attention in image sonification.

In sonifying three-dimensional structures of proteins (Chapter 4), a closer collaboration with domain experts was achieved. Integrating the sonification directly into an existing tool required some tinkering, but led to a robust system that is theoretically easily deployable. In dividing the responsibilities between a sound engine and the interface, a third part for data processing as in the first project was eliminated. The necessary data processing was simplified by the integration into the tool the plugin was written for (Chimera): all data needed was either readily available in the data structures provided by the plugin interface or could be generated by running internal commands.

The corresponding system binds sound objects directly to data points, which are themselves parts of the protein, having a spatial location. The location as well as other parameters are mirrored between Chimera and SuperCollider, with Chimera updating location data every time a user changes the viewpoint. Therefore this sonification also temporalises the static data by the spatial nature of the mapping, corresponding to the position of the user in virtual space in relation to the protein, which means points further from the user have less volume and other psychoacoustic effects provided by the spatialisation library.

The architecture of the system, where data is transformed by a function into sound objects, furthermore enables the rapid development of sonifications by rerunning the mapping function again, using Open Sound Control (OSC) communication to refresh both the SuperCollider and the Chimera side.

The resulting sonification relies on a travelling-wave paradigm for the interaction. Similar to the image sonification, the user is able to highlight different parts of the protein structure and receive the corresponding data mapped to pitch, but also the neighbouring data points are played radiating outwards from the point of interaction. This is another path-based temporalisation of static data, this time in three dimensions.

A prototype of a general sonification plugin was developed enabling the sonification of arbitrary data in the protein.

A very specific domain hindered the evaluation though as the task description turned out to be too vague to test against. The corresponding evaluation showed positive attitudes towards the ideas employed but the mapping from data to sound was deemed not very transparent. The development of a system without a clear, testable task but with access to users can produce interesting results but probably will still be useless in a general context.

The protein dynamics sonification represents the first sonification using Chimera as viewer and the first modular protein sonification system, defining an OSC-based protocol for communication between a protein viewer (in this case, Chimera) and a sonification application.

Drawing from these two projects, which failed in different ways, the Subject-Position Based Design Framework outlined in Chapter 5 is based on lessons learned from these as well as other sonifications and sonification design theories. It provides guidelines for focussed development of ‘useful’ interactive sonifications by focussing the design on finding the necessary subject position that should be inherent in the sonification. This includes finding a soundscape to base the sonification around that is relevant to the task domain itself; finding the most important data to map to the most salient feature of the soundscape; and (optionally) mapping more data to make the sonification more useful for long-term users. The aim is to design sonifications that map the data in such a rich and aesthetic way that their meaning is more easily grasped by the target audience of the sonification. To test the theory, existing sonifications were analysed as well as the protein and image sonifications described before. This provided good pointers for the next project as well as feeding back into the formulation of the framework.

The SPBDF extends e.g., TaDa [11], by adding a way to analyse existing sonifications and providing subject-position based guidance for choosing a soundscape.

Whereas both those projects lacked in definition, the SoniFRED project described in Chapter 6 both provided a clear task description and ample access to users.

A direct integration into the analysis system used was not possible, but in this case the data stream to be sonified could just be transformed into similar information in a standalone system programmed in Max. Everything was integrated into this Max patcher, providing relatively good portability and a relatively user-friendly interface.

Two different sonifications using SPBDF as their foundation were developed: a sonification using a beach soundscape with waves, seagulls, foghorns, and buoy bell sounds, which aimed to create a calm atmosphere suitable to the task; and a sonification using a steam train soundscape with chuffing, wheel noises, screeching brakes and train whistles, which tried to find a soundscape appropriate for the movement profile and provided more direct feedback of the user’s movements.

The sonifications were tested against other forms of feedback, both visual and audiovisual. The results of this evaluation showed a significant decrease of mean speed compared to the normal feedback condition when using the beach sonification, which is a desirable outcome as one of the main problems of using the exercise device at first is going too fast.

This points towards placing a greater focus on the mood of the soundscape than the fit to the task at hand in the SPBDF: the steam train provided an arguably better fit to the task and more direct feedback, which could be the preferable option in long-term usage, but the beach sonification provided the right mood or feeling for using the device, even for first time users: calming, it led to more measured movement.

As outlined in the introduction (see Section 1.3.1), in each of these cases existing tools were studied where possible: the lack of contact with domain experts meant that in the case of the image sonification, literature had to be consulted; the protein viewer used by the domain expert was investigated and led to the development of a general sonification plugin based on a general visualisation plugin provided by the program; the current system and its visual feedback was studied with the collaborators and the exercise device used with it.

This was followed by looking at where sonifications would be beneficial in the context: as another sensory channel to represent data and mitigate fatigue and user error; as additional representation of specific data axes; and as another form of feedback providing less staring at screens by users and therapists.

Each chapter then outlines existing sonifications in these scenarios, outlines potential solutions, and performs evaluations on the solutions.

These all fed back into the development of SPBDF as outlined in Chapter 5, which ultimately led to SoniFRED's steam train sonification being developed according to this framework (Chapter 6), and its evaluation.

Specifically the evaluation of SoniFRED showed that interactive sonification can provide valuable feedback and decrease the mean speed of the user (which contributes to task performance as users in general walk too fast when they are new to the exercise device and its application) when, in the design of the sonification, the soundscape employed invokes an appropriate emotional response in the user.

In response to the research questions outlined in the introduction, interactive sonification can be used to enhance visual feedback in the three projects explored in different ways:

Image sonification: by providing an interface with an overview (auditory glance) of the sonified data first, and then enabling the user to explore parts of the image they find interesting.

Protein dynamics sonification: by coupling the sound to the spatial structure of the protein and giving the user the option to explore specific parts of the structure while highlighting the most important parts of the data and structure at the same time.

SoniFRED: by designing a sonification based around a soundscape, where the most important data is mapped to the most salient part of the soundscape.

An ecological approach to perception can be used to design interactive sonifications as worked out in the SPBDF by finding appropriate soundscapes leading the user to perceive the data mapping in a way that is most helpful to them.

Chapter 8

Conclusion

8.1 Thesis summary

This thesis explores questions around designing interactive sonifications that supplement existing visual systems and focused on developing the theories of sonification first put forward by Vickers [117, 120].

In the field of *pathology* the application was drawing on Bouridane's previous work in diagnostic imaging [112] to investigate how sonification could assist in the detection and recognition of medical image features and therefore support diagnostic imaging. In the course of the development of an interactive sonification system for use by pathologists in the triage stage of colorectal cancer diagnostics different image feature extraction algorithms were implemented and diverse sonification strategies developed. The system displays magnified biopsy samples alongside an auditory representation of extracted features, which is played once when a new image is loaded to provide an *auditory glance* of the image. After this overview, the user can trigger the sonification by probing parts of the image on a touch screen or with the mouse. A discussion of the prototype with a pathologist proved very valuable, but various problems with the project this research was associated with prompted a pivot towards other fields described below.

Molecular visualisation software plays a major role in *structural biology* and structure-based drugs design. To render certain attributes of the molecules on display as sonification a plug-in for interactive sonification was developed for one of the leading software packages for interactive visualisation and analysis of macromolecular complexes (protein-protein or protein-ligand complexes, DNA) [86]. This plugin comprised of: a simple interface for mapping features of the protein to interactive sonification, where users can probe the protein in the protein viewer and a spatialised sonification of the chosen mapping is triggered; a framework for building different sonification methods in Python and an OSC-controlled sound producing software (in this case, SuperCollider); and sonification methods for B-factors representing the dynamics of parts of a protein, forces influencing docking of ligands to proteins, and potential docking sites for protein-protein docking. An evaluation of the B-factor representation with domain experts yielded mostly positive feedback.

Through these different perspectives and drawing from research into design theory undertaken by Vickers, a new theory of sonification design based around the idea of the subject position in film studies developed. A majority of current sonification design theories are focussed on psychoacoustically optimised mapping or elaborate transformations from data into sound. This thesis explores an alternative conception of design: working from a frame of reference encompassing a non-expert audience will lead towards sonifications that are more easily understood.

This theory was put to the test in a collaborative research project with *sports scientists* at Northumbria University. Sonified feedback for an exercise device designed for back pain rehabilitation was developed, which originally only provided feedback to the user through a graph on a screen. Exploring the question if sonification could provide a useful additional feedback channel both for the user and the physiotherapist, as it would enable the both to focus less on the screen in front of them and provide another means of emphasising the critical aspects of usage of the device, led to a fruitful collaboration and a large scale study with two sonifications, one developed by Vickers, both described in this thesis and using the design theory outlined in Chapter 5.

8.2 Thesis contributions

The thesis describes a sound object-based interactive sonification framework for spatial data. Using the OSC protocol to communicate between the spatial data provider and the sonification provider, it provides a simple way to bind elements of a sonification to spatial points. These elements of the sonification are spatialised in relation to the listener position using Ambisonics. With this framework an interactive sonification for protein dynamics was implemented, representing the inherent flexibility of parts of the protein. While the most flexible parts of the protein are always present in the sonification, the user can interact with the protein by clicking on components and receives their sonic representation as well as visual feedback.

This sonification is implemented as a plug-in for an open-source molecular visualisation software (UCSF Chimera), released as open source software. An evaluation showed general interest in domain experts, but the mapping was deemed not very transparent. Furthermore, a general sonification plug-in derived from this sonification was implemented, enabling the mapping of arbitrary data coupled to the protein's components to pitch.

It introduces a design framework for sonification called *SPBDF*, developed as a response to ideas set out by Vickers [117, 120] and expanded with Hogg and Worrall in [121]. It extends their research into design theory, concretising it into a process to design as well as analyse sonifications. It provides guidance for designing sonifications with a specific subject position helpful to the user and the task. This design framework was used to analyse existing sonifications and make predictions around their evaluation. When evaluations were available, the analyses showed a good match with their results.

Lastly, it introduces a real-time feedback system for a physiotherapeutic exercise device designed using SPBDF. SoniFRED provides the user with a sonification of data directly derived from the exercise device's sensors. In a study, the designed sonification around a steam train soundscape did not influence task performance and the other sonification based around sounds found on a beach was preferred.

8.3 Challenges and solutions

As described above, in the image sonification project there was not enough contact with domain experts possible, which led to contacting a local domain expert who volunteered their time and provided some basic evaluation of the prototype and essentially demonstrated its lack of practical basis.

In the protein sonification a domain expert was at hand but a task definition concrete enough for a larger study was not found. In this case the target domain seemed at the same time too big and too specific to concretise into a useful task description.

In both projects, the sonification of a static data set without a time axis led to the development of interactive solutions to the problem, which were made possible by the spatial nature of the data points.

In SoniFRED data was hard to obtain from the tool used by the researchers, so the raw, analogue, data feed from the exercise itself was fed into a sound card and the data processing of the tool had to be rebuilt in another software package based on this "audio" signal.

8.4 Future research directions

This work and its outcomes have highlighted further work that could be done on each of the projects outlined above.

The field of pathology seems to tend into the direction of automated classification in the long run, so further efforts to provide useful interfaces using sonification might not be worth the time and effort, also exacerbated by not enough pathologists. The described prototype could be further developed into an outreach tool or artistic installation though.

In the field of structure-based drug design it would be beneficial to find a task that is possible to be evaluated in a narrow study and develop a sonification for such a task. Furthermore, with a general sonification plugin outlined, this could be extended to support more reasonable sonification options than simple mappings to pitch. Integrating it into Chimera directly or providing it as a third-party plugin would enable longitudinal studies of its use by domain experts.

Exercise devices become more and more important in the wake of gyms popping up everywhere and people becoming more health-conscious, along with the epidemic proportion of people having back pain issues this should provide ample opportunity to improve feedback and therefore exercise results through sonification. The only problem will be designing it in a way that people can integrate it into their routine and maybe media consumption when using exercise devices (i.e., their podcast app or media player).

The analysis, recreation and/or preservation of sonifications could be a valuable activity to further sonification design and practice, as well as providing a database of sonifications that is accessible to researchers. Probably this is not very easily justifiable to funding bodies though, considering the lack of prestige compared to “new” work.

The SPBDF itself could be further refined in its analysis and other ways it works to shape it into a useful tool for designing sonifications. It should also be evaluated further by designing sonifications for tasks in simple target domains that are easily evaluated with many participants, making predictions about the outcomes of the evaluations, and then evaluating them with users to see (1) how the results compare to the predicted results, and (2) how good the resulting sonifications fulfil the requirements of the task.

References

[1] Steam locomotive - Wikipedia. https://en.wikipedia.org/wiki/Steam_locomotive#Steam_circuit.

[2] Otsu's method. *Wikipedia*, June 2018. Page Version ID: 844681806.

[3] AD Instruments. https://www.adinstruments.com/products/labchart, 2017.

[4] Vincent Akkermans, Frederic Font Corbera, Jordi Funollet, Bram De Jong, Gerard Roma Trepat, Stelios Togias, and Xavier Serra. Freesound 2: An improved platform for sharing audio clips. In *Klapuri A, Leider C, Editors. ISMIR 2011: Proceedings of the 12th International Society for Music Information Retrieval Conference; 2011 October 24-28; Miami, Florida (USA). Miami: University of Miami; 2011.* International Society for Music Information Retrieval (ISMIR), 2011.

[5] Joseph Anderson. Introducing... the Ambisonic Toolkit. In *Proceedings of the Ambisonics Symposium 2009*, 2009.

[6] Ulrike Axen and Insook Choi. Using additive sound synthesis to analyze simplicial complexes. In S Smith and G Kramer, editors, *Proceedings of the Second International Conference on Auditory Display (Santa Fe (New Mexico), USA, 7-9 November, 1994)*, pages 31–43. Georgia Institute of Technology, 1994.

[7] Ulrike Axen and Insook Choi. Investigating geometric data with sound. In *Proceedings of the Third International Conference on Auditory Display (Palo Alto (California), USA, 4-6 November, 1996)*, pages 25–28, 1996.

[8] Holger Ballweg, Agnieszka K Bronowska, and Paul Vickers. Interactive Sonification for Structural Biology and Structure-based Drug Design. In Thomas Hermann, Roberto Bresin, and Jiajun Yang, editors, *ISon 2016, 5th Interactive Sonification Workshop*, pages 41–47, Bielefeld, Germany, 2016. CITEC, Bielefeld University.

[9] Michael Banf and Volker Blanz. A modular computer vision sonification model for the visually impaired. In *Proceedings of the 18th International Conference on Auditory Display (Atlanta (GA), USA, 18-21 June, 2012)*, pages 121–128, 2012.

[10] Michael Banf and Volker Blanz. Sonification of images for the visually impaired using a multi-level approach. In *Proceedings of the 4th Augmented Human International Conference (Stuttgart, Germany, 7-8 March, 2013)*, pages 162–169. ACM, 2013.

[11] Stephen Barrass. TaDa! Demonstrations of Auditory Information Design. In Steven P. Frysinger and Gregory Kramer, editors, *ICAD '96 Third International Conference on Auditory Display*, pages 17–24, Palo Alto, 1996. Xerox PARC, Palo Alto, CA 94304.

[12] Natasha Barrett. Interactive Spatial Sonification of Multidimensional Data for Composition and Auditory Display. *Computer Music Journal*, 40:2:47–69, 2016.

[13] Alan M Batterham and William G Hopkins. Making meaningful inferences about magnitudes. *International journal of sports physiology and performance*, 1(1):50–57, 2006. ISSN 1555-0265.

[14] Ludovic Baudry, David Leroy, Ràgis Thouvarecq, and Didier Chollet. Auditory concurrent feedback benefits on the circle performed in gymnastics. *Journal of Sports Sciences*, 24(2):149–156, February 2006. ISSN 0264-0414. doi: 10.1080/02640410500130979.

[15] Oded Ben-Tal, Jonathan Berger, Bryan Cook, Michelle Daniels, Gary Scavone, and Perry Cook. SonART: The sonification application research toolbox. In *Proceedings of the 8th International Conference on Auditory Display (Kyoto, Japan, 2-5 July, 2002)*, pages 151–153, 2002.

[16] Thomas Boggs. Spectral Python (SPy). 2011. [Online; accessed 2015-09-09].

[17] Richard Boulanger, Paris Smaragdis, and John Fitch. Scanned Synthesis: An introduction and demonstration of a new synthesis and signal processing technique. In *Proceedings of International Computer Music Conference (Berlin, Germany, 27 August-1 September, 2000)*, pages 372–375. ICMA, 2000.

[18] Albert S Bregman. *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT press, 1994.

[19] Emily Caddick-Bourne. Verbal response at "*Hearing-in: Philosophical perspectives on sounds and sonification*". Event at University of London's Institute of Philosophy, 10 October (transcribed by Paul Vickers). 2014.

[20] Simon Carlile. Psychoacoustics. In Thomas Hermann, Andy Hunt, and John G. Neuhoff, editors, *The Sonification Handbook*, chapter 3, pages 41–61. Logos Publishing House, Berlin, Germany, 2011. ISBN 978-3-8325-2819-5.

[21] Ryan J Cassidy, Jonathan Berger, Kyogu Lee, Mauro Maggioni, and Ronald R Coifman. Analysis of hyperspectral colon tissue images using vocal synthesis models. In *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on (Pacific Grove (CA), USA, 7-10 November, 2004)*, volume 2, pages 1611–1615. IEEE, 2004.

[22] Ryan J Cassidy, Jonathan Berger, Kyogu Lee, Mauro Maggioni, and Ronald R Coifman. Auditory display of hyperspectral colon tissue images using vocal synthesis models. In *Proceedings of the 10th International Conference on Auditory Display (Sydney, Australia, 6-9 July, 2004)*, page No pages. Georgia Institute of Technology, 2004.

[23] A Chaddad, C Tanougast, A Dandache, A Al Houseini, and A Bouridane. Improving of colon cancer cells detection based on Haralick's features on segmented histopathological images. In *Computer Applications and Industrial Electronics, IEEE International Conference on (Penang, Malaysia, 4-7 December, 2011)*, pages 87–90. IEEE, 2011.

[24] Insook Choi. Interactive Sonification Exploring Emergent Behavior Applying Models for Biological Information and Listening. *Frontiers in Neuroscience*, 12, 2018. ISSN 1662-453X. doi: 10.3389/fnins.2018.00197.

[25] Eric F Clarke. *Ways of Listening: An Ecological Approach to the Perception of Musical Meaning*. Oxford University Press, 2005.

[26] Assaf Cohen, Ehud Rivlin, Ilan Shimshoni, and Edmond Sabo. Memory based active contour algorithm using pixel-level classified images for colon crypt segmentation. *Computerized Medical Imaging and Graphics*, 2015.

[27] Allan D. Coop. Sonification, musification, and synthesis of absolute program music. International Community on Auditory Display, 2016.

[28] T Claire Davies and Catherine M Burns. Advances in cognitive work analysis and the design of ecological visual and auditory displays. *Cognitive Technology*, 13(2):17–23, 2008.

[29] Alberto de Campo. *Science by Ear: An Interdisciplinary Approach to Sonifying Scientific Data*. na, 2009.

[30] Alberto de Campo. Toward a Data Sonification Design Space Map. In *13th International Conference on Auditory Display*, pages 342–347, Montréal, Canada, 26.

[31] Dorothée Debuse, Olivia Birch, Alan St Clair Gibson, and Nick Caplan. Low impact weight-bearing exercise in an upright posture increases the activation of two key local muscles of the lumbo-pelvic region. *Physiotherapy theory and practice*, 29(1):51–60, 2013.

[32] Franco Delogu, Massimiliano Palmiero, Stefano Federici, Catherine Plaisant, Haixia Zhao, and Olivetti Belardinelli. Non-visual exploration of geographic maps: Does sonification help? *Disability & Rehabilitation: Assistive Technology*, 5(3):164–174, 2010.

[33] Florian Dombois. Using audification in planetary seismology. Georgia Institute of Technology, 2001.

[34] Florian Dombois and Gerhard Eckel. Audification. In *The Sonification Handbook*. 2011.

[35] Gaël Dubus. Evaluation of four models for the sonification of elite rowing. *Journal on Multimodal User Interfaces*, 5(3-4):143–156, May 2012. ISSN 1783-7677, 1783-8738. doi: 10.1007/s12193-011-0085-1.

[36] Gaël Dubus. Interactive sonification of motion: Design, implementation and control of expressive auditory feedback with mobile devices. 2013.

[37] John Dunn and Mary Anne Clark. Life music: The sonification of proteins. *Leonardo*, 32(1):25–32, 1999.

[38] Alistair DN Edwards, Geneviève Hines, and Andy Hunt. Segmentation of biological cell images for sonification. In *Image and Signal Processing, 2008. Congress on (Sanya, China, 27-30 May, 2008)*, volume 2, pages 128–132. IEEE, 2008.

[39] Alistair DN Edwards, Andy Hunt, Geneviève Hines, Vanessa Jackson, Alyte Podvoiskis, Richard Roseblade, and Jon Stammers. Sonification strategies for examination of biological cells. In *Proceedings of the 16th International Conference on Auditory Display (Washington (DC), USA. 9-15 June, 2010)*, page No pages, 2010.

[40] Martin Eriksson, Kjartan A. Halvorsen, and Lennart Gullstrand. Immediate effect of visual and auditory feedback to control the running mechanics of well-trained athletes. *Journal of Sports Sciences*, 29(3):253–262, February 2011. ISSN 0264-0414. doi: 10.1080/02640414.2010.523088.

[41] Nicolas Férey, Julien Nelson, Christine Martin, Lorenzo Picinali, Guillaume Bouyer, A Tek, Patrick Bourdot, Jean-Marie Burkhardt, Brian FG Katz, Mehdi Ammi, et al. Multisensory VR interaction for protein-docking in the CoRSAIRe project. *Virtual Reality*, 13(4):273–293, 2009.

[42] Jamie Ferguson and Stephen A. Brewster. Evaluation of Psychoacoustic Sound Parameters for Sonification. http://eprints.gla.ac.uk/151492/, 2017.

[43] Keith M Franklin and Jonathan C Roberts. A path based model for sonification. In *Information Visualisation. Proceedings. Eighth International Conference on (London, UK, 14-16 July, 2004)*, pages 865–870. IEEE, 2004.

[44] Emma Frid, Jonas Moll, Roberto Bresin, and Eva-Lotta Sallnäs Pysander. Haptic feedback combined with movement sonification using a friction sound improves task performance in a virtual throwing task. *Journal on Multimodal User Interfaces*, pages 1–12, May 2018. ISSN 1783-7677, 1783-8738. doi: 10.1007/s12193-018-0264-4.

[45] Steven P Frysinger. A brief history of auditory data representation to the 1980s. In *Proceedings of the First Symposium on Auditory Graphs (Limerick, Ireland, 10 July, 2005)*, pages 410–413, 2005.

[46] Miguel Angel Garcia-Ruiz and Jorge Rafael Gutierrez-Pulido. An overview of auditory display to assist comprehension of molecular information. *Interacting with Computers*, 18(4):853–868, 2006.

[47] Hannah Graham. What officer workers have in common with astronauts. http://www.chroniclelive.co.uk/news/north-east-news/office-workers-could-risk-health-13165834, June 2017.

[48] Florian Grond and Fabio Dall'Antonia. SUMO: A sonification utility for molecules. In *Proceedings of the 14th International Conference on Auditory Display, Paris, France, June 24-27, 2008*, page No pages. International Community for Auditory Display, 2008.

[49] Jan Hammerschmidt and Thomas Hermann. Slowification: An in-vehicle auditory display providing speed guidance through spatial panning. 2016.

[50] Robert M. Haralick and Karthikeyan Shanmugam. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621, 1973.

[51] Kenshi Hayashi and Nobuo Munakata. Basically musical. *Nature*, 310(5973):96, 1984.

[52] Chris Hayward. Listening to the earth sing. In *SANTA FE INSTITUTE STUDIES IN THE SCIENCES OF COMPLEXITY-PROCEEDINGS VOLUME-*, volume 18, pages 369–369. ADDISON-WESLEY PUBLISHING CO, 1994.

[53] Thomas Hermann. *Sonification for Exploratory Data Analysis*. PhD thesis, Bielefeld University, 2002.

[54] Thomas Hermann. Taxonomy and definitions for sonification and auditory display. In *Proceedings of the 14th International Conference on Auditory Display (Paris, France, 24-27 June, 2008)*, page No pages, 2008.

[55] Thomas Hermann. Wave Space Sonification. 2018.

[56] Thomas Hermann and Andy Hunt. Interactive Sonification. In Thomas Hermann, Andy Hunt, and John G Neuhoff, editors, *The Sonification Handbook*, pages 273–298. Logos Verlag, 2011.

[57] Thomas Hermann and Helge Ritter. Listen to your data: Model-based sonification for data analysis. *Advances in intelligent computing and multimedia systems*, 8:189–194, 1999.

[58] Thomas Hermann, T Nattkemper, Walter Schubert, and Helge Ritter. Sonification of multi-channel image data. In *Proceedings of the Mathematical and Engineering Techniques in Medical and Biological Sciences (Las Vegas (Nevada), USA, 26-29 June, 2000)*, pages 745–750, 2000.

[59] Thomas Hermann, Gerold Baier, Ulrich Stephani, and Helge Ritter. Vocal sonification of pathologic EEG features. In *Proceedings of the 12th International Conference on Auditory Display (London, UK, 20-23 June, 2006)*, volume 6, pages 158–163, 2006.

[60] Thomas Hermann, Andy Hunt, and John G Neuhoff. *The Sonification Handbook*. Logos Verlag Berlin, 2011.

[61] Julia Herold, Christian Loyek, and Tim W. Nattkemper. Multivariate image mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1(1):2–13, 2011. ISSN 1942-4795. doi: 10.1002/widm.4.

[62] Henri L. Hurkmans, Johannes B. Bussmann, Eric Benda, Jan A. Verhaar, and Henk J. Stam. Effectiveness of Audio Feedback for Partial Weight-Bearing in and Outside the Hospital: A Randomized Controlled Trial. *Archives of Physical Medicine and Rehabilitation*, 93(4):565–570, April 2012. ISSN 0003-9993, 1532-821X. doi: 10.1016/j.apmr.2011.11.019.

[63] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python. http://www.scipy.org/, 2001. [Online; accessed 2015-09-09].

[64] Emil Jovanov, Dusan Starcevic, Vlada Radivojevic, Aleksandar Samardzic, and Vladimir Simeunovic. Perceptualization of biomedical data. An experimental environment for visualization and sonification of brain electrical activity. *Engineering in Medicine and Biology Magazine, IEEE*, 18(1):50–55, 1999.

[65] T. Kagawa, H. Kudo, S. Tanoue, H. Kiyosue, H. Mori, H. Nishino, and K. Utsumiya. A Supporting Method of Medical Imaging Diagnosis with Sonification. In *Complex, Intelligent and Software Intensive Systems, Sixth International Conference on (Palermo, Italy, 4-6 July, 2012)*, pages 699–704, July 2012. doi: 10.1109/CISIS.2012.144.

[66] T. Kagawa, H. Nishino, S. Tanoue, H. Kiyosue, and H. Mori. A medical imaging diagnosis supporting method with sound effects. In *Awareness Science and Technology and Ubi-Media Computing, International Joint Conference on (Aizu-Wakamatsu City, Japan, 2-4 November, 2013)*, pages 338–344, November 2013. doi: 10.1109/ICAwST.2013.6765462.

[67] Robert Kosara. New Sister Site: EagerEars.org. *EagerEyes blog*, April 2009.

[68] Gregory Kramer and Bruce N. Walker. Sound science: Marking ten international conferences on auditory display. *ACM Transactions on Applied Perception (TAP)*, 2 (4):383–388, 2005.

[69] Gregory Kramer, Bruce Walker, Terri Bonebright, Perry Cook, John H Flowers, Nadine Miner, and John Neuhoff. Sonification report: Status of the field and research agenda. *Prepared for the National Science Foundation by members of the International Community for Auditory Display*, 1997.

[70] David G Lowe. Object recognition from local scale-invariant features. In *Proceedings of the Seventh IEEE International Conference on Computer Vision (Kerkyra, Greece, 20-27 September, 1999)*, volume 2, pages 1150–1157. IEEE, 1999.

[71] Donncha O Maidin and Mikael Fernström. The best of two worlds: Retrieving and browsing. In *Proceedings of the COST G-6 Conference on Digital Audio Effects (Verona, Italy, 7-9 December, 2000)*, page No pages, 2000.

[72] Antonio Cesar Germano Martins and Rangaraj Mandayam Rangayyan. Experimental evaluation of auditory display and sonification of textured images. In *Proceedings of the Fourth International Conference on Auditory Display (Palo Alto (California), USA, 2-5 November, 1996)*, pages 129–134, 1997.

[73] Antonio Cesar Germano Martins and Rangaraj Mandayam Rangayyan. Complex cepstral filtering of images and echo removal in the Radon domain. *Pattern Recognition*, 30(11):1931–1938, 1997.

[74] James McCartney. Rethinking the computer music language: SuperCollider. *Computer Music Journal*, 26(4):61–68, 2002.

[75] Peter BL Meijer. An experimental system for auditory image representations. *IEEE Transactions on Biomedical Engineering*, 39(2):112–121, 1992.

[76] Susanna Mezzarobba, Michele Grassi, Lorella Pellegrini, Mauro Catalan, Bjorn Kruger, Giovanni Furlanis, Paolo Manganotti, and Paolo Bernardis. Action Observation Plus Sonification. A Novel Therapeutic Protocol for Parkinson's Patient with Freezing of Gait. *Frontiers in Neurology*, 8, 2018. ISSN 1664-2295. doi: 10.3389/fneur.2017.00723.

[77] Rosane Minghim and A Robin Forrest. An illustrated analysis of sonification for scientific visualisation. In *Proceedings of the 6th IEEE Visualization Conference (Atlanta (GA), USA, 29 October-3 November, 1995)*, pages 110–117. IEEE Computer Society, 1995.

[78] Konstantin Mischaikow, Marian Mrozek, and Pawel Pilarczyk. Graph approach to the computation of the homology of continuous maps. *Foundations of Computational Mathematics*, 5(2):199–229, 2005.

[79] Kazuaki Nakane, Yasunari Tsuchihashi, and Nariaki Matsuura. A simple mathematical model utilizing a topological invariant for automatic detection of tumor areas in digital tissue images. *Diagn Pathol*, 8:1–4, 2013.

[80] Kazuaki Nakane, Akihiro Takiyama, Seiji Mori, and Nariaki Matsuura. Homology-based method for detecting regions of interest in colonic digital images. *Diagnostic pathology*, 10(1):36, 2015.

[81] Tim W Nattkemper. Multivariate image analysis in biomedicine. *Journal of Biomedical Informatics*, 37(5):380–391, 2004.

[82] Tim W Nattkemper, Thomas Hermann, Walter Schubert, and Helge Ritter. Look & listen: Sonification and visualization of multiparameter micrographs. In *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE (Cancun, Mexico, 17-21 September, 2003)*, volume 2, pages 1311–1314. IEEE, 2003.

[83] Northumbria University. Train like an Astronaut. https://vimeo.com/219668024, June 2017.

[84] Timo Ojala, Matti Pietikäinen, and David Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29 (1):51–59, 1996.

[85] E Paterson, PM Sanderson, NAB Paterson, D Liu, and RG Loeb. The effectiveness of pulse oximetry sonification enhanced with tremolo and brightness for distinguishing clinically important oxygen saturation ranges: A laboratory study. *Anaesthesia*, No pages, 2016.

[86] Eric F Pettersen, Thomas D Goddard, Conrad C Huang, Gregory S Couch, Daniel M Greenblatt, Elaine C Meng, and Thomas E Ferrin. UCSF Chimera – a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13):1605–1612, 2004.

[87] Remy Peyret, Ahmed Bouridane, Fouad Khelifi, Muhammad Atif Tahir, and Somaya Al-Maadeed. Automatic classification of colorectal and prostatic histologic tumor images using multiscale multispectral local binary pattern texture features and stacked generalization. *Neurocomputing*, 275:83–93, 2018.

[88] Joseph Plazak, Simon Drouin, Louis Collins, and Marta Kersten-Oertel. Distance sonification in image-guided neurosurgery. *Healthcare technology letters*, 4(5):199, 2017.

[89] Hesam Ramezanzade, Behrouz Abdoli, Alireza Farsi, and Mohammad Ali Sanjari. The Effects Of Sonification Of Visual Modeling On Relative Timing: The Role In Motor Learning. *ACTA KINESIOLOGICA*, 11:17–27, 2017.

[90] Saima Rathore, Muhammad Aksam Iftikhar, Mutawarra Hussain, and Abdul Jalil. A novel approach for ensemble clustering of colon biopsy images. In *Frontiers of Information Technology (FIT), 2013 11th International Conference On*, pages 25–30. IEEE, 2013.

[91] Benjamin Rau, Florian Frieb, Michael Krone, Christoph Muller, and Thomas Ertl. Enhancing visualization of molecular simulations using sonification. In *Virtual and Augmented Reality for Molecular Science (VARMS@ IEEEVR), 2015 IEEE 1st International Workshop On*, pages 25–30. IEEE, 2015.

[92] Stephen Roddy. *Embodied Sonification*. PhD Thesis, Trinity College Dublin, 2016.

[93] Stephen Roddy and Dermot Furlong. Embodied cognition in auditory display. In *Proceedings of the 19th International Conference on Auditory Display (ICAD 2013)*, 2013.

[94] Stephen Roddy and Dermot Furlong. Rethinking the Transmission Medium in Live Computer Music Performance. 2013.

[95] Stephen Roddy and Dermot Furlong. Embodied Aesthetics in Auditory Display. *Organised Sound*, 19(01):70–77, 2014.

[96] Agnieszka Rogińska, Kent Friedman, and H Mohanraj. Exploring sonification for augmenting brain scan data. In *Proceedings of the 19th International Conference on Auditory Display (Lodz, Poland, 6-9 July, 2013)*, pages 95–101, 2013.

[97] Sophia Roosth. Screaming Yeast: Sonocytology, Cytoplasmic Milieus, and Cellular Subjectivities. *Critical Inquiry*, 35(2):332–350, 2009.

[98] Mohammed Ali Roula, Ahmed Bouridane, and Fatih Kurugollu. An evolutionary snake algorithm for the segmentation of nuclei in histopathological images. In *Image Processing. International Conference on (Singapore, Singapore, 24-27 October, 2004)*, volume 1, pages 127–130. IEEE, 2004.

[99] Penelope Sanderson, Janet Anderson, and Marcus Watson. Extending ecological interface design to auditory displays. In *Proceedings of the 10th Australasian Conference on Computer-Human Interaction*, pages 259–266, 2000.

[100] Penelope M Sanderson and Marcus O Watson. From information content to auditory display with ecological interface design: Prospects and challenges. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 49, pages 259–263. SAGE Publications, 2005.

[101] Rajib Sarkar, Sambit Bakshi, and Pankaj K Sa. Review on image sonification: A non-visual scene representation. In *Recent Advances in Information Technology, 1st International Conference on (Dhanbad, India, 15-17 March, 2012)*, pages 86–90. IEEE, 2012.

[102] Carla Scaletti. Why Sonification is a Joke. Keynote talk presented at ICAD 2017: The 23rd International Conference on Auditory Display, June 2017.

[103] Simon Shelley, Miguel Alonso, Dik Hermes, and Armin Kohlrausch. On the use of sound for representing geometrical information of virtual objects. In *Proceedings of the 14th International Conference on Auditory Display (Paris, France, 24-27 June, 2008)*, page No pages, 2008.

[104] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of the IEEE Symposium on Visual Languages (Boulder (CO), USA, 3-6 September, 1996)*, pages 336–343. IEEE, 1996.

[105] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review. *Psychonomic bulletin & review*, 20(1):21–53, 2013.

[106] Roland Sigrist, Georg Rauter, Laura Marchal-Crespo, Robert Riener, and Peter Wolf. Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning. *Experimental brain research*, 233(3):909–925, 2015.

[107] Denis Smalley. Spectromorphology: Explaining sound-shapes. *Organised sound*, 2 (2):107–126, 1997.

[108] Rudolf Steiner. *Knowledge of the Higher Worlds and Its Attainment*. Book Tree, 2007.

[109] Jelle Stienstra. Augmented Speed-skate Experience. Technical report, 2009.

[110] Jelle Stienstra, Kees Overbeeke, and Stephan Wensveen. Embodying Complexity Through Movement Sonification: Case Study on Empowering the Speed-skater. In *Proceedings of the 9th ACM SIGCHI Italian Chapter International Conference on Computer-Human Interaction: Facing Complexity*, CHItaly, pages 39–44, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0876-2. doi: 10.1145/2037296.2037310.

[111] Alexandra Supper. Sound information: Sonification in the age of complex data and digital audio. *Information & Culture*, 50(4):441–464, 2015.

[112] M.A. Tahir, Fei Yan, P. Koniusz, M. Awais, M. Barnard, K. Mikolajczyk, A. Bouridane, and J. Kittler. A Robust and Scalable Visual Category and Action Recognition System Using Kernel Discriminant Analysis With Spectral Regression. *Multimedia, IEEE Transactions on*, 15(7):1653–1664, November 2013. ISSN 1520-9210. doi: 10.1109/TMM.2013.2264927.

[113] Stephen Taylor. From Program Music to Sonification: Representation and the Evolution of Music and Language. Georgia Institute of Technology, 2017.

[114] Alex Tek, Benoist Laurent, Marc Piuzzi, Zhihan Lu, Matthieu Chavent, Marc Baaden, Olivier Delalande, Christine Martin, Lorenzo Piccinali, Brian Katz, et al. *Advances in Human-Protein Interaction-Interactive and Immersive Molecular Simulations*. InTech, 2012.

[115] Stephen Todd, Peter Todd, Frederic Fol Leymarie, William Latham, Lawrence A Kelley, Michael Sternberg, Jim Hugues, and Stephen Taylor. FoldSynth: Interactive 2D/3D visualisation platform for molecular strands. In *Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine*, pages 41–50. Eurographics Association, 2015.

[116] Stefan Van Der Walt, S Chris Colbert, and Gael Varoquaux. The NumPy array: A structure for efficient numerical computation. *Computing in Science & Engineering*, 13(2):22–30, 2011.

[117] Paul Vickers. Sonification for process monitoring. In Thomas Hermann, Andy Hunt, and John G Neuhoff, editors, *The Sonification Handbook*, pages 455–491. Logos Verlag, 2011.

[118] Paul Vickers. Sonification and Music, Music and Sonification. In Marcel Cobussen, Vincent Meelberg, and Barry Truax, editors, *The Routledge Companion to Sounding Art*, pages 135–144. Routledge, Oxford, 2017. ISBN 978-1-138-78061-3.

[119] Paul Vickers and Bennett Hogg. Sonification Abstraite/Sonification Concrète: An 'Aesthetic Perspective Space' for Classifying Auditory Displays in the Ars Musica Domain. In Tony Stockman, Louise Valgerður Nickerson, Christopher Frauenberger, Alistair D. N. Edwards, and Derek Brock, editors, *ICAD 2006 - The 12th Meeting of the International Conference on Auditory Display*, pages 210–216, London, UK, 20. ISBN 0-902-23821-3 (CDROM) ISBN 0-902-23820-5 (online).

[120] Paul Vickers, Joe Faith, and Nick Rossiter. Understanding Visualization: A Formal Approach Using Category Theory and Semiotics. *Visualization and Computer Graphics, IEEE Transactions on*, 19(6):1048–1061, 2011.

[121] Paul Vickers, Bennett Hogg, and David Worrall. Aesthetics of sonification: Taking the subject-position. In Clemens Wöllner, editor, *Body, Sound and Space in Music and Beyond: Multimodal Explorations*, SEMPRE Studies in The Psychology of Music, chapter 6, pages 89–109. Routledge, 2017. ISBN 978-1-4724-8540-3.

[122] Axel Volmar. Sonic facts for sound arguments: Medicine, experimental physiology, and the auditory construction of knowledge in the 19th century. *Journal of Sonic Studies*, 4(1), 2013.

[123] Bruce N. Walker and Gregory Kramer. Ecological psychoacoustics and auditory displays: Hearing, grouping, and meaning making. *Ecological psychoacoustics*, pages 150–175, 2004.

[124] Lilian Lee Yen Wei and Ag Asri Ag Ibrahim. Sonification of Geo-Visualization for Complex Dengue Data: An Investigation Using Ambient Sound Framework. *Advanced Science Letters*, 24(3):1892–1895, 2018.

[125] John Williamson, Roderick Murray-Smith, and Stephen Hughes. Shoogle: Excitatory multimodal interaction on mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 121–124. ACM, 2007.

[126] Scott Wilson, David Cottle, and Nick Collins. *The SuperCollider Book*. The MIT Press, 2011. ISBN 0-262-23269-3 978-0-262-23269-2.

[127] Andrew Winnard, Mona Nasser, Dorothee Debuse, Maria Stokes, Simon Evetts, Mick Wilkinson, Julie Hides, and Nick Caplan. Systematic review of countermeasures to minimise physiological changes and risk of injury to the lumbopelvic area following long-term microgravity. *Musculoskeletal Science and Practice*, 27:S5–S14, 2017.

[128] Kristina Wolfe. Sonification and the Mysticism of Negation. *Organised Sound*, 19 (03):304–309, 2014.

[129] David Worrall. *Sonification and Information: Concepts, Instruments and Techniques*. PhD thesis, 2009.

[130] Woon Seung Yeo and Jonathan Berger. Raster Scanning: A New Approach to Image Sonification, Sound Visualization, Sound Analysis, and Synthesis. In *Proceedings of the International Computer Music Conference (New Orleans (LA), USA, 6-11 November, 2006)*, pages 34–41. Department of Music, Stanford University, 2006.

[131] Woon Seung Yeo, Jonathan Berger, and Zune Lee. SonART: A framework for data sonification, visualization and networked multimedia applications. In *Proceedings of the International Computer Music Conference (Miami, USA, 1-6 November, 2004)*, page No pages, 2004.

[132] Tsubasa Yoshida, Kris M. Kitani, Hideki Koike, Serge Belongie, and Kevin Schlei. EdgeSonic: Image Feature Sonification for the Visually Impaired. In *Proceedings of the 2nd Augmented Human International Conference (Tokyo, Japan, 13 March, 2011)*, AH '11, pages 11:1–11:4, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0426-9. doi: 10.1145/1959826.1959837.

[133] Haixia Zhao. *Interactive Sonification of Abstract Data-Framework, Design Space, Evaluation, and User Tool*. PhD thesis, University of Maryland, 2006.

[134] Haixia Zhao, Catherine Plaisant, Ben Shneiderman, and Jonathan Lazar. Data sonification for users with visual impairment: A case study with georeferenced data. *ACM Transactions on Computer-Human Interaction*, 15(1):4, 2008.

Appendix A

Glossary

Audification – Auditory display where the data is directly played (e.g., by speeding up and listening to seismographic recordings), therefore listening to the raw data points.

Auditory Display – Any expression of data through the medium of sound (sonification, musification, audification).

Earcon – Brief, distinctive sound mapped to a specific meaning (e.g., the sound played by an operating system denoting an error). Earcons therefore represent a type of auditory display.

Musification – A type of auditory display more focussed on the musical aesthetics than the communication of data.

Programme music – Type of music that strives to express an extra-musical narrative (e.g., Vivaldi's *Four Seasons*).

Sonification – A type of auditory display where non-speech audio is used to convey information.

Soundscape – An auditory landscape - e.g., all the sounds a listener would perceive in a specific location (forest, café, car factory).

Work song – A song linked to a form of work, either about the work or sung while working (usually to coordinate steps that need to happen in a specific rhythm).

Appendix B

ChiSon consent form

Information Sheet

Sonification to Support Structural Biology and Structure-based Drug Design

Researcher	**Holger Ballweg**
Data Controller	**Northumbria University** Newcastle City Campus Ellison Place 2 Newcastle-upon-Tyne NE1 8ST 0191 232 6002 **Holger Ballweg** [contact details redacted]
Programme of study	**Computer Science and Digital Technologies PGR**
Supervisor's name	**Dr. Paul Vickers**

What is this project about?

The research project looks into if and how sonification can support biologists and chemists in tasks related to structural biology and structure-based drug design.

Which information is required of me?

Initially, we will ask you to fill in a short questionnaire about your workplace and your working habits. This will help us plan the software and make it as comfortable and useful for you as possible.
During the user study, we will provide a short training session and then you will be asked to use the software to work on tasks from the field of structural biology and structure-based drug design and provide feedback via another questionnaire afterwards. You will be either using the software with or without sonification, depending on which group you will be in. With your permission, you will be filmed during the use of the software to provide us with more data about usability issues. The software will also log every interaction. With your help, we will be able to assess if and how sonification can support you in your tasks.

Are there any risks?

There is a risk of damaging your hearing by high volume sounds, but the maximum volume will be calibrated before the study and monitored during. You will be able to modify the volume of the sound system, but won't be able to increase it over the maximum volume fixed beforehand.

What about anonymity?

Information obtained through the user studies and questionnaires will be used anonymised. Video recordings will either not contain your face or be anonymised afterwards and only retained in this form.

How will the information be published?
It will be published anonymised in the dissertation and in scientific papers detailing the findings. We will send you copies of the publications if you want.

How long will the study take?
The initial questionnaire shouldn't take more than 15 minutes.
The user study will likely take approximately 2-3 hours. You will be asked to fill in a questionnaire directly after the user study so the experience is still fresh in your mind, but if you have further comments you can contact the researchers any time.
This second questionnaire should not take more than 30 minutes to fill in.

What happens with the data when the research project ends?
The data will be stored by the university for at least 3 years after the study is completed and then disposed of.

Can I withdraw my permission?
Under the Data Protection Act 1998 you may withdraw your permission or ask to access your information at any time. Please contact the data controller – Northumbria University or Holger Ballweg – by any means, preferably in writing.

Appendix C

ChiSon questionnaires

Questionnaire

Sonification to Support Structural Biology and Structure-based Drug Design

Researcher	**Holger Ballweg**
Data Controller	**Northumbria University** Newcastle City Campus Ellison Place 2 Newcastle-upon-Tyne NE1 8ST 0191 232 6002 **Holger Ballweg** [contact details redacted]
Programme of study	**Computer Science and Digital Technologies PGR**
Supervisor’s name	**Dr. Paul Vickers**

Workplace

How noisy is your normal work environment?

(1) I could hear a needle fall
(2) …
(3) …
(4) …
(5) …
(6) …
(7) I have to speak loudly on the phone and press the receiver on my ear to hear

Working habits

What tasks do you use molecular visualising application (e.g., Chimera) for?

How many hours a day do you use molecular visualising application on average?
[number]

Do you feel like using a molecular visualising application is more or less fatiguing than your other work?

(1) More fatiguing
(2) …
(3) Equally fatiguing
(4) …
(5) Less fatiguing

Do you see any deficits in using a molecular visualising application?

Sound, Music, Sonification

Did you receive musical training?

(1) Yes, still do

(2) Yes, in the last 5 years

(3) Yes, in the last 10 years

(4) Yes, longer than 10 years ago

(5) No

Do you have a hearing difficulty? (tick what applies)

[ ] No

[ ] Tinnitus

[ ] Partial hearing loss

[ ] Full hearing loss

How often do you listen to music while working?

(1) Always

(2) Regularly

(3) Half the day

(4) Once a day

(5) Once a week

(6) Once a month

(7) Never

Do you use sound as an auxiliary device in your normal working day (e.g., pulse-oxymeter)?

If you work on molecules, do you imagine sounds associated with them? If yes, could you describe them?

Questionnaire 2

Sonification to Support Structural Biology and Structure-based Drug Design

Researcher	**Holger Ballweg**
Data Controller	**Northumbria University** Newcastle City Campus Ellison Place 2 Newcastle-upon-Tyne NE1 8ST 0191 232 6002 **Holger Ballweg** [contact details redacted]
Programme of study	**Computer Science and Digital Technologies PGR**
Supervisor’s name	**Dr. Paul Vickers**

How would you describe your experience with the sonifications you heard?

Beneficial	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Musical	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Annoying	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Helpful	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Fatiguing	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Transparent	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree

Would you be interested in using this sonification in your work?

absolutely (1) (2) (3) (4) (5) not at all

If so, where would you use it?

[ ] Research

[ ] Teaching

[ ] Other: _______________

Would you be interested in using sonification in general in your work?

absolutely (1) (2) (3) (4) (5) not at all

Did you attitude to sonification change during this experiment?

absolutely (1) (2) (3) (4) (5) not at all

Was the mapping transparent?

absolutely (1) (2) (3) (4) (5) not at all

Other comments

Please add any further comments or feedback you have here.

Appendix D

SoniFRED consent form

northumbria
UNIVERSITY NEWCASTLE

Faculty of Health & Life Sciences

Aerospace Medicine and Rehabilitation Laboratory
Northumbria University

INFORMED CONSENT FORM

Project Title: SoniFRED

Principal Investigators: Kirsty Lindsay and Holger Ballweg

	please tick or initial where applicable
I have carefully read and understood the Participant Information Sheet.	☐
I have had an opportunity to ask questions and discuss this study and I have received satisfactory answers.	☐
I understand I am free to withdraw from the study at any time, without having to give a reason for withdrawing, and without prejudice.	☐
I have answered ‘no’ to all of the health screening questions and feel that I am physically fit enough to take part in the SoniFRED study. I understand I must inform the research team immediately if I feel unwell at any time during the study.	☐
I agree to take part in this study.	☐

Signature of participant.. Date..........................

(NAME IN BLOCK LETTERS)...

Signature of researcher.. Date..........................

(NAME IN BLOCK LETTERS)...

northumbria
UNIVERSITY NEWCASTLE
Faculty of Health & Life Sciences

Aerospace Medicine and Rehabilitation Laboratory
Northumbria University

PHOTOGRAPHS/VIDEOS/TAPE RECORDINGS CONSENT FORM

Project title: SoniFRED and Mission X Train Like and Astronaut

Principal Investigator: Kirsty Lindsay and Holger Ballweg

I hereby confirm that I give consent for the following recordings to be made:

Recording	Purpose	Consent
Photography of activity, including facial photograph	SoniFRED- to illustrate scientific literature Mission X- for use in publicity material	
video of activity, including bodily movement	SoniFRED- to illustrate scientific literature and provide insight into FRED use Mission X- for use in publicity material	

I understand that the recording(s) may be published in an appropriate journal/textbook or on an appropriate Northumbria University webpage, **which would automatically mean that the recordings would potentially be available worldwide**. My name or other personal information will never be associated with the recording(s). I understand that I have the right to withdraw consent at any time prior to publication, but that once the recording(s) are in the public domain there may be no opportunity for the effective withdrawal of consent

Tick or initial the box to indicate your consent ☐

Signature of participant... Date.....……………….

Signature of Parent / Guardian in the case of a minor

.... Date.....……………….

Signature of researcher... Date.....……………….

Appendix E

SoniFRED questionnaires

Participant ID: ....

Questionnaire 1: SoniFRED

Researchers	**Kirsty Lindsay**	**Holger Ballweg**
Data Controller	**Northumbria University** Newcastle City Campus Ellison Place 2 Newcastle-upon-Tyne	NE1 8ST 0191 232 6002
	Kirsty Lindsay [contact details redacted]	**Holger Ballweg** [contact details redacted]
Programmes of study	**Sport Exercise and Rehabilitation PGR**	**Computer Science and Digital Technologies PGR**
Supervisors' names	**Dr Nick Caplan**	**Dr Paul Vickers**

Before attending our exhibition and/or taking part in the FRED research trial...

1.1 ... what level was your understanding of how spinal muscles contribute to postural control?

low (1) (2) (3) (4) (5) (6) (7) high

1.2 ... what level was your understanding of how the spinal postural muscles decondition during spaceflight, with low back pain and/or as we age?

low (1) (2) (3) (4) (5) (6) (7) high

1.3 Do you have a vision impairment?

( ) No

( ) Yes, I am short-/farsighted

( ) Yes, namely: ___________________________

1.4 Did you receive musical training or taught/teach yourself?

(1) Yes, still do

(2) Yes, in the last 5 years

(3) Yes, in the last 10 years

(4) Yes, longer than 10 years ago

(5) No

1.5 Do you have a hearing difficulty? (tick what applies)

[ ] No [ ] Tinnitus [ ] Partial hearing loss [ ] Full hearing loss

Participant ID: ....

Questionnaire: SoniFRED

Researchers	**Kirsty Lindsay**	**Holger Ballweg**
Data Controller	**Northumbria University** Newcastle City Campus Ellison Place 2 Newcastle-upon-Tyne	NE1 8ST 0191 232 6002
	Kirsty Lindsay [contact details redacted]	**Holger Ballweg** [contact details redacted]
Programmes of study	**Sport Exercise and Rehabilitation PGR**	**Computer Science and Digital Technologies PGR**
Supervisors' names	**Dr Nick Caplan**	**Dr Paul Vickers**

1. General

1.1 If you were to kick a ball, would you prefer to use your left or right foot?

( ) Left foot ( ) Right foot

1.2 Do you currently experience low back pain (from below the ribs to the buttocks)?

( ) Yes ( ) No

1.3 Have you experienced low back pain in the past?

( ) Yes, ...ago ( ) No

1.4 Did you experience any leg or foot discomfort/pain while walking on FRED?

( ) Yes, right leg/foot
( ) Yes, left leg/foot
( ) Yes, both legs/feet
( ) No

1.5 Did you have a favourite/preferred type of feedback?

( ) Standard visual feedback
( ) Metronome only
( ) Metronome and visual feedback
( ) Sonified feedback (Beach)
( ) Sonified feedback (Steam train)
( ) Non music sound and visual feedback
( ) Music and visual feedback
( ) *No preference*

Before attending our exhibition and/or taking part in the FRED research trial...

1.6 ... what level was your understanding of how spinal muscles contribute to postural control?

low (1) (2) (3) (4) (5) (6) (7) high

1.7 ... what level was your understanding of how the spinal postural muscles decondition during spaceflight, with low back pain and/or as we age?

low (1) (2) (3) (4) (5) (6) (7) high

After attending our exhibition and/or taking part in the FRED research trial...

1.8 ... what level is your understanding of how spinal muscles contribute to postural control?

low (1) (2) (3) (4) (5) (6) (7) high

1.9 ... what level was your understanding of how the spinal postural muscles decondition during spaceflight, with low back pain and/or as we age?

low (1) (2) (3) (4) (5) (6) (7) high

1.10 Did you receive musical training or taught/teach yourself?

(1) Yes, still do

(2) Yes, in the last 5 years

(3) Yes, in the last 10 years

(4) Yes, longer than 10 years ago

(5) No

2. Sonified feedback

2.1 Did you encounter sonified feedback/sonification/auditory display before?

( ) Yes, ... (*please elaborate*)
( ) No

2.2 How did you experience the beach sonification (waves, seagulls, etc.)?

Helpful for keeping pace	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Helpful for motivation	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Annoying	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Calming	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Appropriate for the activity	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree

2.3 How did you experience the steam train sonification (chuffing, wheel sounds, etc.)?

Helpful for keeping pace	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Helpful for motivation	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree

Annoying	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Calming	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree
Appropriate for the activity	fully agree	(1)	(2)	(3)	(4)	(5)	not at all agree

2.4 Which sonified feedback did you prefer?

( ) Beach ( ) Steam train ( ) No preference ( ) No sonified feedback

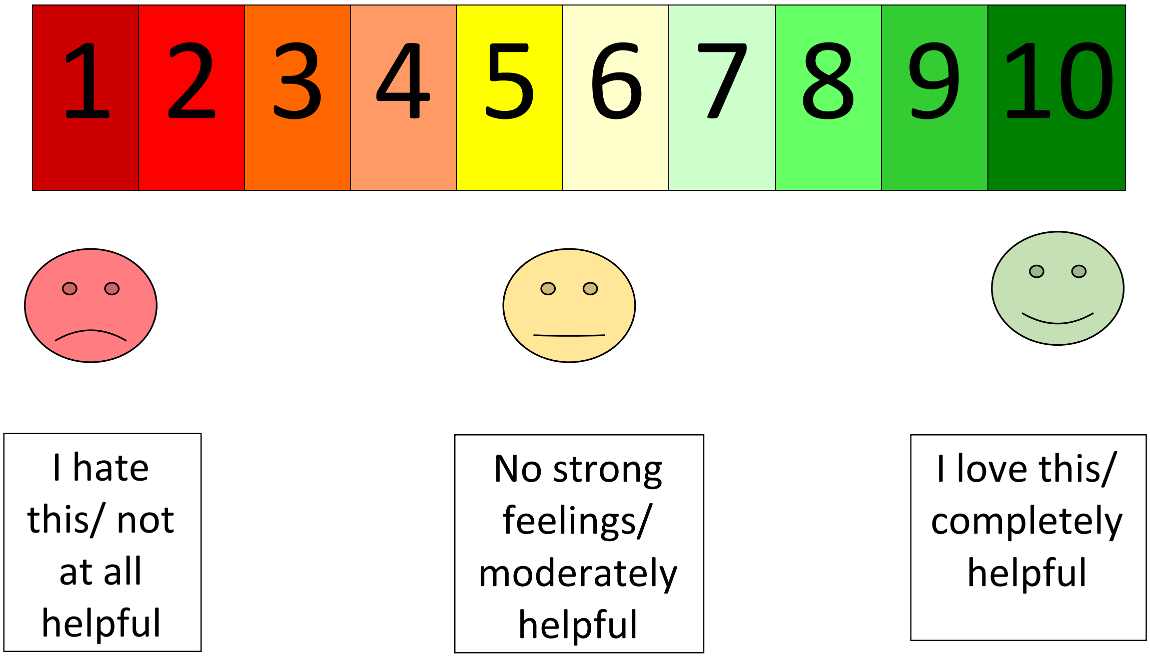
2.5 Do you think other exercise devices would profit from this kind of feedback?

( ) Yes, *(please specify*) .... ( ) No

3. Other comments

3.1 If you have any other feedback you would like us to give, please feel to add them here:

E.1 Visual analogue scale



Appendix F

Publications

Ballweg: 'Sonification to Support Image Analysis and Diagnostic Imaging', *Student Think Tank at the 21st International Conference on Auditory Display*. Graz, 2015

Ballweg, Bronowska, Vickers: 'Interactive Sonification for Structural Biology and Structure-based Drug Design', *Proceedings of the Interactive Sonification Workshop (ISon) 2016*. Bielefeld, 2017.