Available online at www.sciencedirect.com
ScienceDirect
Int. J. Human-Computer Studies 67 (2009) 923–933
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Interactive sonification of complex data

Sandra Pauletto[a,*], Andy Hunt[b]

[a]Department of Theatre, Film and Television, The University of York, Heslington, York YO10 5DQ, United Kingdom
[b]Department of Electronics, The University of York, Heslington, York YO10 5DD, United Kingdom

Received 16 September 2008; received in revised form 26 February 2009; accepted 19 May 2009
Available online 30 May 2009

Abstract

In this paper we present two experiments on implementing interaction in sonification displays: the first focuses on recorded data (interactive navigation) and the second on data gathered in real time (auditory feedback).

Complex synthesised data are explored in the first experiment to evaluate how well the known characteristics present in the data are distinguished using different interaction methods, while real medical data (from physiotherapy) are used for the second.

The addition of interaction to the exploration of sonified recorded data improves the system usability (efficiency, effectiveness and user satisfaction), and the real-time sonification of complex physiotherapy data can produce sounds with timbral characteristics that audibly change when important characteristics present in the data vary.
© 2009 Elsevier Ltd. All rights reserved.

Keywords: Sonification; Audio feedback; Interactive navigation

1. Introduction

1.1. Sonification

Sonification is defined as "the use of non-speech audio to convey information" (Kramer, 1994). It is considered to be a timely topic, since just at the point in history where data is being produced in abundance, the tools for generating high-quality audio output are becoming widely available on everyday multimedia computing systems. The discipline of sonification concerns the exploration of methods for mapping the gathered data into a suitable sonic form that can be readily comprehended by human listeners.

1.2. Interactive sonification

Interactive sonification can be considered as a particular case of auditory display where the user is dynamically involved in the generation of the sound. Hermann and Hunt define it as

> "…the use of sound within a tightly closed human-computer interface where the auditory signal provides information about data under analysis, or about the interaction itself, which is useful for refining the activity." (Hermann and Hunt, 2005, p. 20)

Interaction in sonification can be implemented in two main ways.

- *Interactive sonification of recorded data*: previously stored data can be explored interactively while the sound is being generated.
- *Interactive sonification of data gathered in real time*: the sound can be produced *as* the data are gathered, i.e. in real time. In this case, changes in the data are instantly transformed into changes in sound and the user might be involved in producing the data.

In both cases the aim of the interaction is to allow the user to learn more, and more efficiently, about the information being displayed.

*Corresponding author. Tel.: +44 1904 433604; fax: +44 1904 432066.
E-mail addresses: sp148@york.ac.uk (S. Pauletto), adh2@ohm.york.ac.uk (A. Hunt).

1071-5819/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijhcs.2009.05.006

1.3. Review of research into interactive sonification

Interest in interaction has been shown by the auditory display community since the early 1990s. The concept of a *Sound Probe* (for pointing at a region of data of interest) was presented in Gröhn (1992) and then used again in Barrass and Zehner (2000) for the exploration of well logs. Fernström and McNamara (1998) describe the importance of *direct manipulation* (characterised by the continuous representation of objects of interest and by immediate feedback; Shneiderman and Plaisant, 2005) and apply it to an application for browsing musical tunes.

Saue (2000) presented a general model for the sonification of large spatial data sets in which "[…] *the interpreter is walking along paths in areas of the data set, listening to locally and globally defined sound objects*". This "virtual walking" could be done using a mouse or similar input device. The mouse has been the first choice of interface device for many researchers working in interactive sonification as it is the most common computer interface device in use to this day. Winberg and Hellström (2001) used the mouse as a *virtual microphone*. Hermann (2002) used it to interact with data spaces in his early examples of Model-Based Sonification (MBS)—a form of auditory display, which is akin to configuring the data under investigation as an *instrument*, which is "played" by the user in an intrinsically interactive way.

In the last few years, with the higher processing power of computers, more research on alternative interfaces can be found. For example, Beamish et al. (2003) presented a system that uses a *haptic turntable* for controlling the playback of digital audio. They argue that the system, initially intended for DJs, could also be used for the exploration of data as sound. Hermann and colleagues since 2001 have been exploring the use of novel interfaces to interact with MBS systems. In Hermann et al. (2001) a custom-built "hand box" interface is described where hand posture is analysed and reconstructed as a multi-joint hand model and used for exploration of sound and space. In Hermann et al. (2003) the *gesture desk* interface is introduced, which tracks the free movements of the hands and uses them to interact with data spaces. Their later *audio-haptic ball* is made of plasticine, and equipped with various sensors (acceleration and force-sensitive sensors) that send data to the computer when shaken, scratched, squeezed, rotated and hit. These interactions are then used to excite the MBS system, giving the user the impression that they are shaking a box to discover what is inside.

Milczynski et al. (2006) presented a malleable interaction surface for continuous and localised exploration of data using the fingers, and Bovermann et al. (2006) showed how moving an object (for example a stick) one could scan a data set which is virtually positioned around the stick.

A good source of information on the various aspects of interactive sonification can be found in the special issue of IEEE Multimedia, which was dedicated, in 2005, to interactive sonification (Hermann and Hunt, 2005).

1.4. Interactive sonification of medical data

The medical community has already used sound feedback for a number of applications. The stethoscope is still a fundamental medical tool, widely used for a range of tasks. Clinicians use such interactive sonic feedback as an everyday diagnostic technique, allowing them to detect complex time-based events (such as heartbeat and airflow), and hidden structural defects; such is the power of sound as an analytical aid. Sonification has the potential of being able to portray many more simultaneous data parameters than visual displays (Scaletti and Craig, 1990), whilst freeing the eyes and hands for visual–spatial tasks (such as surgery, or communicating with colleagues and patients). This requires the sound to be rendered in real time, and reacting instantly to changes in the data.

Recent examples of studies on auditory feedback in the medical field include Jovanov et al. (1998), who presented the use of audio as feedback for precise manual positioning of a surgical instrument and referred to it as *tactile audio*, Effenberg (2005) on the auditory feedback of movement, Hinterberger and Baier (2005) on auditory biofeedback of electroencephalography (EEG) data that allows self-regulation of the brain activity, Fox and Carlile (2005), who presented the *SoniMime* system with auditory feedback to assist the user in learning a gesture with minimal error and Ghez et al. (2000), who presented real-time auditory feedback of joints and muscles for patients who are unable to maintain "internal models" of their limbs and therefore monitor movements. In all these examples, the presence of audio was found to help self-regulatory abilities, making it reasonable to hypothesise that auditory feedback of physiotherapy data could produce interesting results.

1.5. Experiments to investigate interactive sonification

The focus of this paper is to investigate the two main scenarios for interactive sonification: recorded data and data gathered in real time. In both cases we are investigating if the addition of the real-time/interactive element helps the analysis of the information contained in the data. In both cases we are concentrating on large, complex data sets, containing more than one channel of data and with thousands of data points per channel. We are investigating whether overall characteristics and structures present in the data are clearly distinguishable in the sound when (in the first case) the data are explored with a high degree of interaction and (in the second case) when the data are sonified in real time, at the original sampling rate, producing a sound with a complex timbre. The data used in the first experiment were synthesised by the researcher so that the structures contained in the data would be known. In the second experiment the data are from the field of physiotherapy, a medical field where auditory biofeedback is of great interest for rehabilitation or diagnostic tools.

The aim of the first experiment is to compare the interactive navigation of complex data, carried out with methods that differ in their degree of interactivity. The first two methods use the mouse (due to its popularity), used in two different ways. The third method uses the jog wheel/shuttle interface (Contour Design, 2007), a popular device used in audio and video editing applications. The experimental task involves recognising the different data structures (e.g. a ramp, a noisy structure, a periodic structure, etc.) in the data set, which is a fundamental task in the field of analysis of complex data.

The second experiment proposes sonification of electromyography (EMG) data (data produced by the firing of the leg muscles when producing a contraction or movement) that can be produced in real time as the data are gathered. The focus of the experiment is to verify if some important characteristics of the data can be perceived in real time when all the six channels are simultaneously sonified. If this is possible we can deduce that meaningful auditory feedback of this data can be created and that useful information can be understood from it.

The combination of the results of these experiments will tell us if the addition of interaction to a sonification system in the two scenarios described above is useful in making the overall system more effective.

2. First experiment: the design, implementation and evaluation of low, medium and high interactive navigation methods

Here we compare three methods of navigating recorded data to see if, with the addition of interaction, the resulting sonification could become more effective and efficient. In an interactive sonification system, the user is allowed to "experience" the data, having a certain degree of control over the way the sound is generated from the data.

The main aspect to be evaluated is the *usability* of the system design, and typical measures of usability are (Shneiderman and Plaisant, 2005):

- The *number of errors* made while performing the task, i.e. the fewer the errors, the higher the effectiveness of the system.
- The *time to complete a task*, i.e. the shorter the time, the higher the efficiency of the system.
- *User satisfaction*, i.e. how much the users like the system. Data can be gathered via questionnaires, interviews and open comments.

These measures correspond also to three of Nielsen's criteria for the evaluation of system usability: errors, efficiency and satisfaction (Nielsen, 1993). Nielsen defines two additional usability criteria, learnability and memorability, which were not tested in this experiment given that the task is relatively simple and easy to learn. Further experiments, exploring more complex tasks, could test these two additional criteria.

2.1. The experiment

The aim of this experiment was to evaluate and compare three different methods for interacting with a data set as it is turned into sound. A usability test was carried out using the *Interactive Sonification Toolkit* (Pauletto and Hunt, 2004) as a high-fidelity prototype. The usability test focused on measuring the *efficiency* and *effectiveness* of the three interaction methods. At the end of the test, subjects answered a questionnaire about their perception of the interaction provided by different interaction methods. This gathered information about *user satisfaction*.

Here follows the description of the rationale behind the experiment design. The final goal of a software tool that uses sonification to display complex data sets is to support data analysis and exploration efficiently and effectively. A fundamental task in data analysis is to be able to identify particular structures present in the data under examination. Three separate factors can affect the efficiency and effectiveness of data analysis via sonification:

1. the specificity of the data set;
2. the sonification algorithm;
3. the interaction method.

Once a sonification method has been proven to be a good display for a particular data set (and once the user knows what kind of data structures can be found in the data set) the efficiency and the effectiveness with which a user analyses a data set and recognises the data structures it contains depend only on how the user is allowed to navigate the auditory display. In this experiment the independent variable is the interaction method, which is evaluated under three different conditions. The dependent variables, on the other hand, are the time spent to complete a task (measure of *efficiency*) and the number of data structure identification errors made during the execution of the task (measure of *effectiveness*).

2.2. Experiment description

Subjects were asked to navigate and listen to a data set undergoing sonification using three different interaction methods. Their task was to recognise which data structures were present in the data set and in which order. Before starting the test, the subjects were trained to listen to the particular kinds of data sets sonified in the test so that they had experience of how to analyse and recognise the structures present in the data sets. This training was given so that the experimenter did not need to consider the specific sonification algorithm and the type of data sets as variables in this experiment.

In the toolkit, the user is presented with an interaction area in the top half of the screen. This area represents the sonification from beginning (left) to end (right). The user can navigate the data by interacting with this screen area using two types of interface devices: the mouse and the jog wheel/shuttle interface.

2.3. The three interaction methods in detail

The three interaction methods studied in this experiment were:

(1) *Low interaction method*: a section of data can be selected with the mouse, and the playback duration can be fixed before playback (see Fig. 1);
(2) *Medium interaction method*: uses the jog wheel/shuttle interface to switch between different fixed playback speeds and direction (see Fig. 2);
(3) *High interaction method*: immediate change in direction and speed of playback is possible by moving the mouse over the interaction area as the point indicated by the mouse in the interaction area is sonified in real time.

Fig. 1 shows the interaction area and how it is used in the Low interaction method. Fig. 2 shows the jog wheel/shuttle interface and its functions for the Medium interaction method.

2.4. The data

The data for this experiment were synthesised by the experimenter. On the basis of the experimenter's experience in working with data sonification, five main data structures were considered to be very basic and common in data sets produced by any type of process (e.g. natural, mechanical, etc.):

(1) noisy structure (random numbers);
(2) constant structure (sequence of constant values);
(3) linear structure (an ascending linear ramp);
(4) discontinuous structure (a sequence of constant values interrupted randomly by single different values);
(5) periodic structure (a sine wave).

Each data set used in this experiment included all of these structures and each data set channel contained the same number of data samples (220,500).

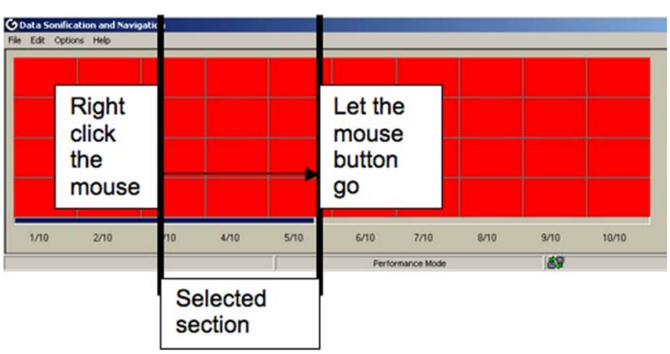

Fig. 1. Selection of data section (original in colour).

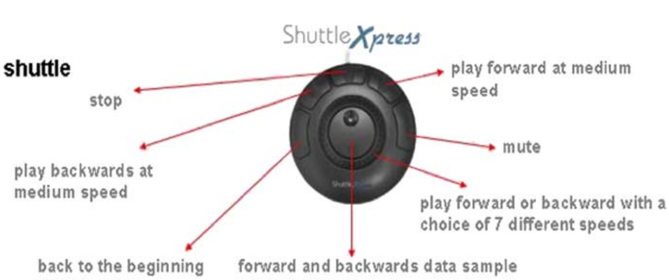

Fig. 2. The Shuttle XPress Interface.

2.5. The sonification method

The data sets were mapped to the amplitude of a sine oscillator (fundamental frequency 261.6 Hz, middle C). This type of mapping is simple enough so that people can very quickly learn and remember how the different data structures sound using this mapping.

For this experiment it was important that people could easily recognise the data structures, if they heard a simple data set. However, in order to be able to measure the effects of different interaction methods on the identification of the structures, it was important to create more challenging data sets that needed repeated listening to be understood. If repeated listening was not needed, the action of navigating the sonification would be pointless and obviously would not be measurable.

The strategy used to make the data sets at the same time simple, but needing repeated listening, was to construct them in such a way so that they would challenge the subject's hearing attention all the time. Experiments show (see Bregman, 1994, pp. 207–208) that if two different sequences of words are presented to subjects, one in the left ear and the other in the right ear, and afterwards the subjects are asked to repeat one sequence, they usually cannot report the words heard in the non-attended ear. This fact led to the idea, in this experiment, of playing different streams in the two ears simultaneously. Usually subjects would need repeated listening to switch attention from the left to the right ear and recognise all the elements in the two streams. For this reason, it was decided that each data set should contain two channels of data, one panned to the left and one to the right, each containing two different sequences of the five structures mentioned above.

Three different data sets were constructed. In each of them the order of the structures' sections and their length were different. Figs. 3–5 represent in a schematic way the three data sets containing two sequences of five different data structures.

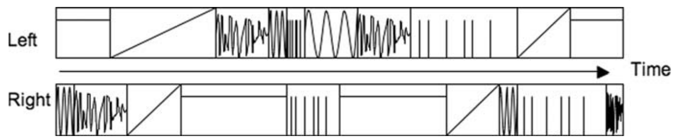

Fig. 3. Graphical representation of the first data set.

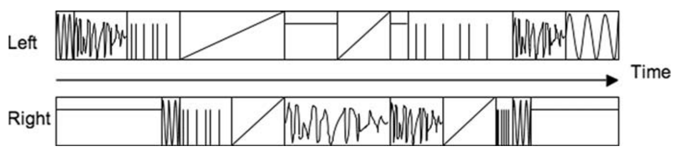

Fig. 4. Graphical representation of the second data set.

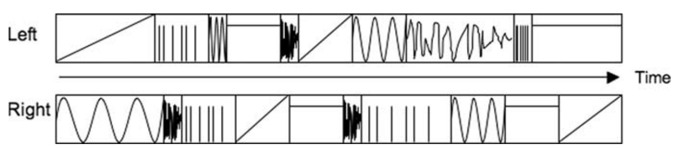

Fig. 5. Graphical representation of the third data set.

2.6. The subjects

This experiment has a within-subjects (or related) design; so the same group of subjects do the experiment under all the conditions. The order of presentation of conditions and data sets was randomised.

Eighteen subjects (15 men and 3 women, average age 28) participated. Sixteen subjects normally work with sound, while two do so sporadically. Again the test was carried out in a silent room with good quality headphones.

2.7. The experimental procedure

The subjects were given some training in using the interfaces and in recognising the data structures. When the experiment began the subjects were presented with a sheet of paper with two tables where they could write down the structures they heard in both ears. Fig. 6 shows the test sheet presented to the subjects.

The experimenter measured how long the subject took to recognise the sequences using a hand-held chronometer.

After the experiment each subject was asked to fill in a questionnaire and to rate from 1 (low) to 5 (high) the *pleasantness* of each interaction method, the *intuitiveness*, the *clarity* and the *quickness*. The questionnaire was designed to gather information about user satisfaction, and the perceived efficiency and effectiveness of the interaction methods. Then, the subjects were asked to select their *preferred* interaction method and to *comment* on why they chose it.

2.8. Efficiency results

To analyse the results, non-parametric tests were performed as the data did not satisfy the conditions for

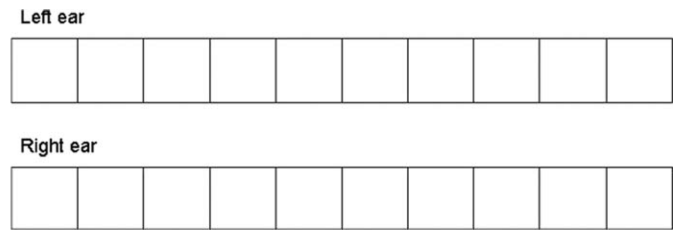

Fig. 6. Test sheet.

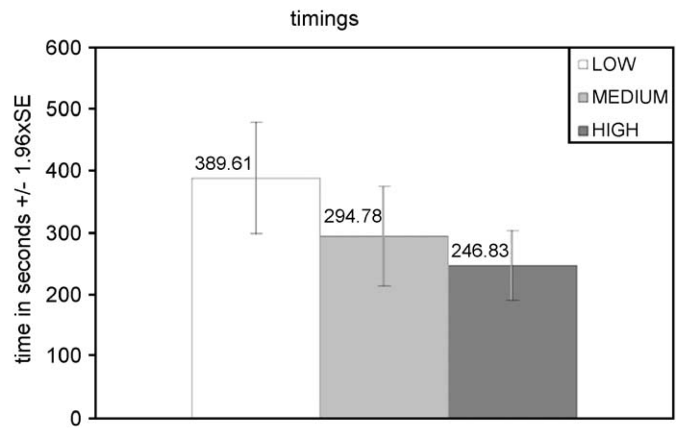

Fig. 7. Efficiency results.

parametric tests. Non-parametric tests such as Friedman's ANOVA (for more than two related conditions) and Wilcoxon signed rank (for two related conditions), with a Bonferroni correction applied to the significance level, were calculated to verify the significance of the differences between the three conditions.

The results for efficiency and the relative statistical tests indicate that the Low interaction method is slower than the High interaction method (p(one-tailed)<0.0167).

This means that exploring sonification by moving the mouse freely up and down the timeline is more efficient than selecting a section of the sonification, selecting the speed at which to hear it and then pressing play. The efficiency of the Medium interaction method is not significantly different from those of the other two methods. Fig. 7 shows the results for efficiency.

2.9. Effectiveness results

Subjects were asked to perform a task under three conditions, and the number of incorrect answers for each subject was recorded.

In this experiment, no significant difference in effectiveness was found between the three conditions. Fig. 8 shows the effectiveness results.

2.10. Verification of the equivalence of the data sets used in this test

The experiment is based on the assumption that the three data sets used in the test were equivalent to analyse, i.e. no

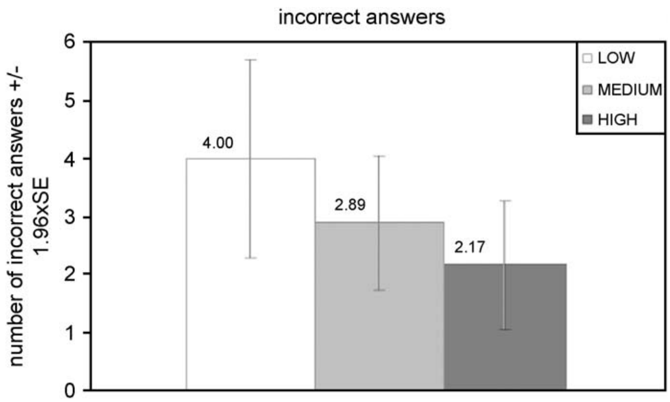

Fig. 8. Effectiveness results.

data set was easier to analyse than another data set. This assumption was verified using the test results. The test scores were grouped for each data set and it was verified that the differences in average score (for both effectiveness and efficiency) between the different data sets were not significant, i.e. no data set was easier to analyse than the others.

Moreover, no significant differences in efficiency and effectiveness were found between the average results obtained the first time the task was performed, the second and the third.

2.11. User satisfaction

The Medium interaction method was found to be the most pleasant ($p<0.0167$) followed by the High interaction method ($p<0.0167$) and then the Low interaction method ($p<0.0167$). This means that the subjects preferred using the jog wheel/shuttle interface, which allowed changing speed and direction of playback quickly and had a set number of constant playback speeds. The second most pleasant method was the mouse with a direct mapping between speed of movement of the mouse and playback speed. Selecting the speed of playback and pressing "play" was considered to be the least pleasant method of interaction.

The results for intuitiveness and clarity tell us that the Low interaction method was considered significantly less intuitive and clear ($p<0.0167$), than the other two methods. Differences between these last two methods are not significant.

The Low interaction method is also perceived to be significantly slower than the other interaction methods ($p<0.0167$). There is no significant difference in quickness between the Medium and High interaction methods.

The 18 test subjects were asked directly in the questionnaire which interaction method they preferred:

- 45% said the Medium interaction method;
- 33% said the High interaction method;
- 22% said the Low interaction method.

2.12. Summary of questionnaire results

The Medium interaction method was judged to be the most pleasant and was the preferred interaction method by the subjects. The High interaction method is the second best method and for intuitiveness, clarity and quickness it scores as well as the Medium interaction method. The Low interaction method is the worst in all the questionnaire results.

The Medium interaction method is considered appropriate, "professional and nice". It has various functionalities (different fixed speeds, back and forward playback, "return to zero" button, fine control of slow speed, etc.) and works often hands free (i.e. hands can be used to do other things while listening to the display). It is not necessary to look at the interaction area to know where one is in the data at any given time. This last aspect is at times considered confusing and can become also a disadvantage. The use of fixed playback speeds is considered both comforting and at times restrictive.

The High interaction method is considered to give total control because there is a direct relationship between interface device position and speed, and the playback position and speed. With this method, changes in position are immediate. The device used is very familiar and results can be obtained very quickly.

2.13. Summary of overall results

The clearest result of this experiment is that the Low interaction method is considered the worst under all aspects. The High and Medium interaction methods are considered better methods to use for navigating data and in particular the Medium interaction method is considered significantly more pleasant and was the preferred method of this group of subjects.

The main result from the objective measurements of efficiency and effectiveness is that the Low interaction method is less efficient than the other methods, while the interaction methods were not significantly different in terms of effectiveness, i.e. the analysis of the sonification can be done equally well using all the methods.

Finally, subjects indicated in this experiment that, even if there are no particular objective differences between the Medium and High interaction methods, they prefer and find more pleasant the Medium interaction method, which uses the jog wheel/shuttle interface. The reason for this preference could be the fact that the Medium interaction method provides, at the same time, very quick changes in playback speeds and direction, while allowing eyes- and hands-free moments in which the user can concentrate solely on the sound. The Medium interaction method does not require constant activity from the user (the High interaction method does) allowing a shift of attention rapidly between different tasks such as restarting the sound, changing playback speed, listening to the sound and analysing the data.

From this experiment we can conclude that the addition of a relatively high level of interaction to a sonification display improves the efficiency of the analysis of the sonified data. From the point of view of users' perception, the addition of a relatively high level of interaction greatly improves the overall auditory display, which is then perceived as more pleasant, clear, intuitive and quick to use. Interesting ideas for further work would be to explore in detail the reasons for the subjects' preference of the Medium interaction method and develop and test a hybrid interface device for interactive sonification that groups the qualities of the mouse and the shuttle interface.

3. Second experiment: real-time sonification of movement data

This section evaluates the effectiveness of a sonification display of multivariate data that could be produced in real time, interactively as the data are gathered. The particular data set used in this experiment contains electromyography data. Physiotherapists use EMG sensors to monitor the electrical activity of the muscles of patients. Electrodes attached to the skin of the subject detect the electrical signals from the muscles below the skin, and send it to a computer via an analogue-to-digital converter (ADC). The computer typically runs an application that receives the data, performs some basic statistics, and displays them in graphical form for analysis.

EMG signals are believed to be full of information about the muscle activity and it is hypothesised by our collaborators at the School of Heath and Social Care of Teesside University (UK) that current visual analysis does not exploit to the full the information contained in the data.

The fact that the technology allows us to display the data in real time, i.e. as they are gathered, creates the possibility for the patient to be involved in a deeper understanding of his/her movement. Two main obstacles prevent current visual EMG displays from working as biofeedback:

(1) most of the graphical display is meaningless to a non-expert;
(2) there is little shared and non-technical language that allows the transfer of clear and simple information between the physiotherapist and the patient.

It is believed that a sound display of EMG data could help overcome these two obstacles and create a real-time biofeedback tool.

If we could produce an auditory display of EMG data that maps correct movement to a clear characteristic of a sound, and incorrect movement to a clear, alternative characteristic of the sound, then the patient should gain a clear understanding of what a good movement sounds like and could use that sound as the target to aim for in rehabilitation when exercising the muscle's movement.

3.1. Current analysis

The two most important parameters in the study of EMG signals are amplitude and frequency (LeVeau and Andersson, 1992). The changes in these two parameters can be quantified and used to classify the electrical activity level that produces a certain muscular tension.

> The change in the myoelectric signal is based on the recruitment and firing rate of motor units within the muscle. In general, as more force is needed, more motor units are recruited, and the motor units already firing increase their frequency of firing. […] The interpretation of the changes in recruitment and the changes in firing rate can provide information concerning the muscle's level of force or its level of fatigue. (Ibid., p. 70)

Currently there are various non-real-time methods used to extract information about frequency and amplitude of the raw signal ranging from visually monitoring the raw signal to extracting the linear envelope or spectrum of the signal.

3.2. The data

The data used in this experiment were gathered by physiologist Dr John Dixon, lecturer at Teesside University (Middlesbrough), for his Ph.D. research work (Dixon, 2004). Dr Dixon measured the EMG data of three muscles of the leg: the vastus medialis, the vastus lateralis and the rectus femoris (see Fig. 9 for a schematic picture of the leg muscles).

All data were collected using a data acquisition system (BIOPAC Inc., USA), which includes a MP100 work-station with a high-level HLT100 transducer and dedicated analysis software (AcqKnowledge 3.5; BIOPAC, 2009).

To have an overall picture of the EMG activity on the leg muscles, 6 pairs of recording electrodes were used—two pairs per muscle. One electrode pair was placed in a low position with respect to the centre of the muscle and the other in a higher position. Fig. 10 shows the electrodes placements on the leg.

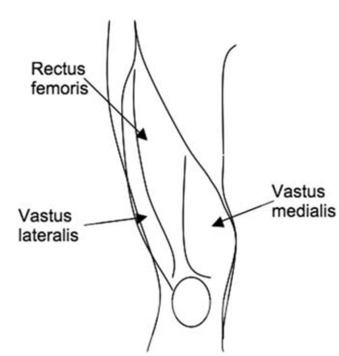

Fig. 9. This diagram sketches the positioning of the vastus lateralis, vastus medialis and the rectus femoris on a right leg.

Fig. 10. Positioning of the electrode pairs. Still image from video made by the author at Teesside University in collaboration with Dr Dixon and colleagues (original in colour).

The placement was standardised so that it could be reproduced in different subjects with different bodies. This standardisation allows for the data to be compared.

The subjects belonged to 3 particular groups: young (<45 years old) asymptomatic (i.e. not known to be exhibiting symptoms, in this case of osteoarthritis—OA) participants (23 subjects), old (>45 years old) asymptomatic participants (17 subjects) and old (>45 years old) patients (17 subjects) with symptoms of osteoarthritis of the knee.

The subjects were asked to perform a Maximal Voluntary Isometric Contraction (MVIC). To perform this action, the patient is seated on a custom-built chair padded and comfortable, without his/her feet touching the floor. The leg is strapped to a fixed metal bar at the ankle. When the subject is asked to perform a MVIC, he/she will try to push his/her strapped foot forward as much as possible. The muscles of the thigh will therefore contract and produce firings detectable by the electrodes.

Each subject was asked to perform 5 contractions. Verbal encouragement was given using the phrase "push, push, push" in order to produce maximum contraction values. Between contractions the subjects were asked to rest for about 30 s (Dixon, 2004). The EMG data were gathered at a sampling rate of 2048 samples per second, i.e. one sample roughly every 0.5 ms. For the purpose of the experiment presented here, 30 such contractions were selected, each one containing 6 channels of data (2 channels per muscle).

We aimed to verify if trends that we knew existed in the data could be displayed using sonification. The reason for this choice is that if participants can hear in the sound a trend that *we know* exists then we can be sure that our display is effective in portraying important information.

The following characteristics are known to change with the age of the subjects, as reported by Dr Dixon:

1. the overall amplitude of the signal tends to reduce with age;
2. the slope or rise of the signal tends to reduce with age.

These two characteristics represent the muscles having less power with increased age and taking a longer time to reach the maximum power. The aim of the experiment described here is to verify that these two characteristics are clearly displayed by the chosen sonification algorithm.

3.3. The sonification method

For each subject, there exists a data file with 6 channels: 2 channels (representing 2 electrode pairs) for the vastus medials and vastus medialis oblique muscles, 2 channels for the vastus lateralis and 2 channels for the rectus femoris.

The sonification was designed following these criteria:

1. use the raw data, i.e. all the potentially useful information is maintained;
2. be a real-time sonification, i.e. the data must be fed into the sound synthesis algorithm at the same rate as they are gathered;
3. be easy to interpret.

Surface EMG signals do not contain frequencies above 500 Hz (Dixon, 2004). This means that if periodic patterns exist in the data they will vary at a rate that is either within a middle–low audio frequency range or at a rate that is lower than audio range, i.e. control rate.

Consequently, to create a real-time, easy to interpret sonification, we need to map the data to a sound parameter that can vary meaningfully at either audio rate or control rate. A sound parameter that has these characteristics is the amplitude of an oscillator (carrier) and the sound synthesis technique is *amplitude modulation* (AM). Table 1 shows the carrier frequency assigned to each channel of data and Fig. 11 shows diagrammatically the sonification algorithm.

All the channels were scaled so that the original relative relationship between channels' values was maintained. The carrier frequencies ranged between 200 and 1600 Hz for two main reasons:

1. the loudness/frequency dependency is not very strong in this frequency range;
2. auditory display designers tend to centre their display in the 200–5000 Hz frequency range because it corresponds to the typical range of various musical instruments and because it is the most effectively perceived frequency region (high sensitivity to frequency change; Walker and Kramer, 2004).

3.4. The resulting sound

Typically a contraction data set sonified in this way contains an overall tone (produced by small offsets present in the 6 channels of data) and an enveloped noisy part. In some cases the distinct tones of some oscillators are

Table 1
Electrode to carrier frequency assignment.

Electrode pair	Oscillator frequency (Hz)
RFUP	261.6 (mid C)
RFLOW	523.2
VLUP	784.8
VLLOW	1046.5
VMUP	1308.1
VMLOW	1569.7

where
RFUP = rectus femoris electrode pair positioned *high* with respect to the centre,
RFLOW = rectus femoris electrode pair positioned *low* with respect to the centre,
VLUP = vastus lateralis electrode pair positioned *high* with respect to the centre,
VLLOW = vastus lateralis electrode pair positioned *low* with respect to the centre,
VMUP = vastus medialis electrode pair positioned *high* with respect to the centre,
VMLOW = vastus medialis electrode pair positioned *low* with respect to the centre.

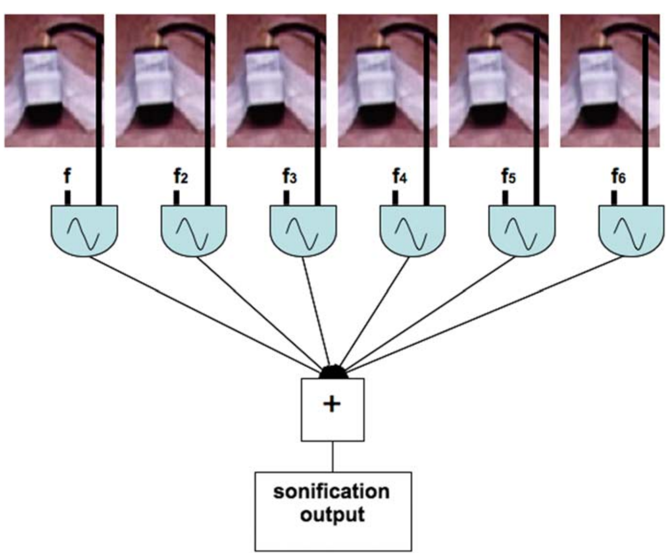

Fig. 11. EMG sonification algorithm.

perceived as separate from the overall tone and in some cases the overall envelope of the sound has a complex evolution in time (i.e. with various bumps and valleys), not simply an attack, a sustained and a release part. The resulting sound is normally perceived as a single timbre (instead of a mix of different timbres playing simultaneously). Each data channel contributes to the output sound by introducing sound energy (the AM side bands) in a frequency band around the frequency of the carrier of that particular channel.

The presence or absence of each of the six channels is perceived as a presence or absence of energy around the channel carrier frequency in the output sound.

The overall sound can also be said to sound ‘‘rough’’. This fact is not surprising as the sensation of roughness can be caused by a signal with AM when the modulation frequency is in a range between 15 and 300 Hz (Zwicker and Fastl, 1999) and we know that surface EMG signals do not contain frequencies above 500 Hz (Dixon, 2004). This experiment will also verify if the perception of roughness in the sound can be related to any characteristic of the data.

3.5. The experimental procedure

After having transformed each contraction's data set into a sound using the Interactive Sonification Toolkit (Pauletto and Hunt, 2004), a program was developed in Pure Data (Pure Data, 2002) in order to run the experiment and gather most of the experimental results automatically.

Subjects were presented with a screen containing 30 buttons. Each button, if clicked, played the sonification of one data set (the order of data sets was randomised for each subject). Subjects were asked to listen to all the sounds at least once before starting the test so that they were aware of the sound variability. They were then asked to score from 1 to 5 (low to high) the following characteristics of the sound of each data set:

(1) roughness;
(2) overall loudness;
(3) speed of sound's attack.

3.6. Reasons for choosing to test the above characteristics

Roughness can be considered as a descriptor that indicates an overall quality of the sound's timbre. It was noticed that the sounds could be described using this descriptor and, therefore, it was decided to test if it varied with, for instance, age or group (i.e. symptomatic or asymptomatic group). Also, the word ‘‘roughness’’ is a general descriptor that *non-technical* people can understand. If a relationship is found between this descriptor and a parameter such as age or group, then this term could be used to communicate information between patients and physiotherapists in a non-technical manner. For example, when physiotherapists say to their patients: ‘‘try to make the sound rougher’’; the patients, while trying to control the roughness of the sound, will actually change the way their muscles fire.

Overall loudness and *speed of sound's attack* vary with age and thus they were the main variables to be tested in order to evaluate the effectiveness of sonification.

The scores of each subject were recorded automatically in the computer. Each subject was presented with the data in a random order so that biases due to order of presentation were cancelled. Listening conditions were as for the first experiment.

3.7. The subjects

This experiment has a within-subjects or related design since the same group of listening subjects scored all of the sounds. This is also a correlational study because we primarily want to know if the perception of certain sonification parameters is correlated with the age of the subjects producing the EMG data.

Twenty-one subjects (18 men and 3 women, average age 29) performed the test. Nineteen participants were used for working with sound, and 2 were researchers in physiotherapy with sufficient understanding of sound. All subjects declared having normal hearing.

3.8. The results

For each one of the 30 sounds and for each of the characteristics, an average was calculated over the number of listening subjects. The roughness, attack speed and loudness average scores rank correlated with the age of the patients providing the EMG data (p(one-tailed)<0.01) with the following Spearman's correlations: -0.54 for roughness, -0.58 for loudness and -0.75 for attack speed.

The overall loudness and the speed of the sound attack were perceived to decrease with age. This confirms our hypothesis and therefore allows us to say that this type of EMG sonification does portray effectively some characteristics of the data as expected.

The Spearman's correlation of the sound attack speed with age is much higher than for roughness and overall loudness. This seems to be the characteristic that best portrays the age of the person providing the EMG data.

The roughness average scores are highly correlated with the loudness average scores, Spearman's correlation 0.91, and with the attack speed average scores, Spearman's correlation 0.73.

This means that people perceive the sound resulting from this sonification to be rough when the loudness is high and the attack speed is high. Since these two characteristics are now known to be related to age then, on the basis of this correlation, we can expect a rough sound to belong to a young person and a less rough sound to belong to an older person. The descriptor "roughness" is an example of one descriptor that could be used in the communication between a patient and a physiotherapist. For instance, if a physiotherapist needs to tell a patient to try to make a younger and healthier muscle contraction, he/she could say "Can you make the sound rougher?". The online version of this paper contains three sound examples of the EMG sonification for subjects 24, 43 and 71 years old.

3.9. A step further: real-time implementation of EMG sonification

This type of sonification can be created in real time because the synthesis algorithm accepts the EMG data at the producing sampling rate. A demo real-time implementation was created routing the EMG signals from the BIOPAC system to an ADC converter by PicoTech Ltd., which then sent them to the PD program running the sonification algorithm (a video of this demo can be found in the online version of this paper). Fig. 12 shows the equipment connections involved in the demo audio biofeedback. Further work is needed to test this audio biofeedback system with patients.

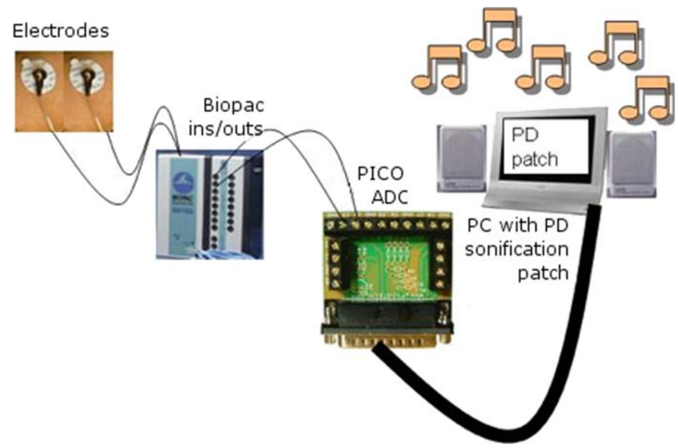

Fig. 12. Connections for the real-time implementation of the EMG sonification (original in colour).

4. Conclusions

In this paper two studies have been presented that investigate the two main approaches to interactive sonification: the interactive sonification of recorded data and real-time data. We have found that the overall usability of a sonification display of complex data increases if the user is allowed to navigate the display interactively. We have found that real-time sonification of large EMG data sets is possible and gives useful information that can be represented by a sound descriptor that can be easily understood by non-experts (e.g. patients). In particular, the timbral sound parameter – roughness – was found to be correlated to the age of the person making the movement, i.e. the rougher the sound, the younger the person. It is hypothesised that, if the system is used as an audio biofeedback, this sound descriptor could be used as measure of how young and healthy the limb movement is, making the system a useful rehabilitation tool.

Overall these results tell us that displays using sound to portray complex data can be effective and that the addition of user interaction increases the usability of the system, and can add functionalities to the system, as in the case of the audio biofeedback.

Acknowledgements

The work presented here is the result of the research conducted during the EPSRC funded project *Data mining through an interactive sonic approach* based in the Electronics Department of the University of York.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ijhcs.2009.05.006.

References

Barrass, S., Zehner, B., 2000. Responsive sonification of well logs. In: Proceedings of the International Conference on Auditory Display, Atlanta, USA.

Beamish, T., van Den Doel, K. Maclean, K., Fels, S. 2003. D'groove: a haptic turntable for digital audio control. In: Proceedings of the International Conference on Auditory Display, Boston, USA.

BIOPAC, 2009. Biopac Systems, Inc. (online) available at: 〈http://www.biopac.com〉 (accessed 26 February 2009)

Bovermann, T., Hermann, T., Ritter, H., 2006. Tangible data scanning sonification model. In: Proceedings of the International Conference on Auditory Display, London, UK.

Bregman, A.S., 1994. Auditory Scene Analysis. MIT Press, Cambridge, MA.

Contour Design, 2007. Contour Design (online) available at: 〈http://www.contourdesign.com〉 (accessed 26 February 2009).

Dixon, J., 2004. An electromyographic analysis of quadriceps femoris in patients with osteoarthritis of the knee. Ph.D. Thesis, School of Health, Teesside University, Middlesborough, 2004.

Effenberg, A.O., 2005. Movement sonification: effects on perception and action. IEEE Multimedia 12 (2), 53–59.

Fernström, M., McNamara, C., 1998. After direct manipulation—direct sonification. In: Proceedings of the International Conference on Auditory Display, British Computer Society, Glasgow, UK

Fox, J., Carlile, J., 2005. SoniMime: sonification of fine motor skills. In: Proceedings of the International Conference on Auditory Display, Limerick, Ireland.

Ghez, C., Rikakis, T., Duois, R.L., Cook, P.R., 2000. An auditory display system for aiding interjoint coordination. In: Proceedings of the International Conference on Auditory Display, Atlanta, USA.

Gröhn, M., 1992. Sound probe: an interactive sonification tool. In: Proceedings of the International Conference on Auditory Display, Santa Fe, USA.

Hermann, T., Nölker, C., Ritter, H., 2001. Hand postures for sonification control, in gesture and sign language in human–computer interaction. In: Wachsmuth, Ipke, Sowa, Timo (Eds.), Proceedings of the International Gesture Workshop GW2001, 2002. Springer, Berlin, pp. 307–316.

Hermann, T., Henning, T., Ritter, H., 2003. Gesture desk—an integrated multi-modal workplace for interactive sonification. In: Proceedings of the International Gesture Workshop, Genova, Italy.

Hermann, T., 2002. Sonification for exploratory data analysis. Ph.D. Thesis, Faculty of Technology, Bielefeld University, Germany.

Hermann, T., Hunt, A., 2005. An introduction to interactive sonification. IEEE Multimedia 12 (2), 20–24 (April–June).

Hinterberger, T., Baier, G., 2005. Parametric orchestral sonification of EEG in real time. IEEE Multimedia 12 (2), 70–79 (April–June 2005).

Jovanov, E., Starcevic, D., Wegner, K., Karron, D., Radivojevic, V., 1998. Acoustic rendering as support for sustained attention during biomedical processing. In: Proc. of the International Conference on Auditory Display, British Computer Society, Glasgow, UK.

Kramer, G., 1994. An Introduction to auditory display. In: Kramer, G. (Ed.), Auditory Display: Sonification, Audification, and Auditory Interfaces, Santa Fe Institute Studies in the Sciences of Complexity, Proceedings vol. XVIII. Addison-Wesley, Reading, MA.

Leveau, B., Andersson, G.B.J., 1992. Output forms: data analysis and applications. In: Soderberg, G.L. (Ed.), Selected Topics in Surface Electromyography for Use in the Occupational Setting: Expert Perspectives, US Department of Health and Human Services, Public Health Service, Centers for Diseases Control, National Institute for Occupational Safety and Health, 1992, pp. 69–102.

Milczynski, M., Hermann, T., Bovermann, T., Ritter, H., 2006. A malleable device with applications to sonification-based data exploration. In: Proceedings of the International Conference on Auditory Display, London, UK.

Nielsen, J., 1993. Usability Engineering. Academic Press, Boston.

Pauletto, S., Hunt, A., 2004. A toolkit for interactive sonification. In: Proceedings of the International Conference on Auditory Display, Sydney, Australia.

PURE DATA programming environment, 2002. PURE DATA (online) available at: 〈http://pd.iem.at/〉 (accessed 26 February 2009).

Saue, S., 2000. A model for interaction in exploratory sonification displays. In: Proceedings of the International conference on Auditory Display, Atlanta, US.

Shneiderman, B., Plaisant, C., 2005. Designing the User Interface, fourth ed. Addison-Wesley, Reading, MA.

Scaletti, C., Craig, A.B., 1990. Using sound to extract meaning from complex data. In: Farrel, E. (Ed.), Extracting Meaning from Complex Data: Processing, Display, Interaction, Proceedings of SPIE, vol. 1459, pp.207–219.

Walker, B.N., Kramer, G., 2004. Ecological psychoacoustics and auditory displays: hearing, grouping and meaning making. In: Neuhoff, J. (Ed.), Ecological Psychoacoustics. Academic Press, New York, pp. 150–175.

Winberg, F., Hellström, S.O., 2001. Qualitative aspects of auditory direct manipulation: a case study of the towers of Hanoi. In: Proceedings of the International Conference on Auditory Display, Helsinki, Finland.

Zwicker, E., Fastl, H., 1999. Psychoacoustics: Facts and Models. Springer, Berlin, New York.