Multimedia Systems 7: 23–31 (1999)

Multimedia Systems

Using sonification

Stephen Barrass[1], **Gregory Kramer**[2]

[1] GMD IMK – German National Research Centre for Information Technology, Schloss Birlinghoven, D-53754 St. Augustin, Germany;
e-mail: stephen.barrass@gmd.de
[2] Clarity/Santa Fe Institute, 310 NW Brynwood Lane, Portland, OR 97229, USA; e-mail: kramer@listen.com

Abstract. The idea behind sonification is that synthetic non-verbal sounds can represent numerical data and provide support for information processing activities of many different kinds. This article describes some of the ways that sonification has been used in assistive technologies, remote collaboration, engineering analyses, scientific visualisations, emergency services and aircraft cockpits. Approaches for designing sonifications are surveyed, and issues raised by the existing approaches and applications are outlined. Relations are drawn to other areas of knowledge where similar issues have also arisen, such as human-computer interaction, scientific visualisation, and computer music. At the end is a list of resources that will help you delve further into the topic.

Key words: Sonification – Visualisation – Multi-modal – Multimedia – Perceptual display – Human-computer interaction – Information design – Auditory display

Introduction

A "nano-guitar" the size of a single cell, with strings that actually vibrate, has been built by physicists at Cornell University (Craighead 1997). It could be a miracle cure – all we have to do is inject nano-guitars into the bloodstream of a sick person so the bacteria form garage bands that can be tracked down with a stethoscope. The nurses might need earplugs in the intensive care ward!

The idea of using sounds to diagnose illnesses, and even save lives, is not speculative or even unusual in hospital, where a stethoscope is a normal part of a doctors equipment. Medical students are taught to listen to tissues rubbing in the lungs, gasses bubbling in the intestines, and blood pumping through veins. Many other indicators, such as body temperature or blood CO_2 levels, are measured and shown as graphs. However, a graph can be distracting during visually demanding tasks in an operation, and it is possible to synthesise sounds to represent these indicators instead. Medical students performed better in a simulated operation when eight dynamic variables about the health of the patient were presented as sounds rather than graphs, and better with sounds alone than with both sounds and graphs combined (Fitch and Kramer 1994). Images produced by X-ray, Cat scans and magnetic resonance imagery (MRI) equipment are often used to look for symptoms of disease inside a patients body. However, it is very difficult to visually detect unhealthy regions of the brain in an MRI image, because of the nature of brain tissue. Unhealthy regions of the brain can be made more distinguishable by mapping image texture into sounds that can be heard by selecting a region of interest with a mouse (Martins et al. 1996). Listening to the data may help a doctor to diagnose a dangerous illness which might otherwise go undetected.

Medicine is not the only area where sounds can provide new insights into data relations and allow new and better ways to carry out a task. The next section describes a variety of other situations where sonifications have been tried and found to be useful.

Correspondence to: S. Barrass

Sonification is useful

Blind people rely on natural sounds even more than most of us. They learn to listen for useful sounds, and filter out those that are not. While walking through town, the sound of a constant line of traffic is useful for navigating a straight line and maintaining whereabouts, whilst the voices of passers-by are not (Swan 1996). Electronic aids have been developed which can assist a blind traveler to be more mobile and independent. One such system registers the traveler on a digital map with a global positioning satellite (GPS), while virtual sounds that seem to come from landmarks and buildings along a predefined route provide guidance (Loomis et al. 1994). Multimedia computer programs can make maps, diagrams and text more accessible to the visually impaired. "Audiograf" is a program that generates a sound from part of a diagram selected with a finger on a touch-screen. A line between two points sounds like a plucked string, and text selections are heard as speech (Kennel 1996). "Mathtalk" augments a text-to-speech translator with non-verbal cues to make it easier for a listener to understand written mathematical expressions. The cues provide an auditory overview of the organization of the expression by graphic symbols such as parentheses and subscripts – for example, an opening parenthesis has a rising tone and a closing parenthesis has a falling tone (Stevens et al. 1994). “Mercator” makes it easier for a blind person to use a graphical user interface (GUI) by mapping windows, menus, buttons and text fields into auditory navigation cues (Mynatt 1994a). Similarly, a web browser has been augmented with auditory cues about heading levels, layout, hyperlinks, and download times to make the internet more accessible for the visually impaired (James 1996). “Triangle” is a suite of sonified applications to support print-impaired students and professionals in maths, science, and engineering, which includes a word-processor, a gra phing calculator, x-y plots, tables, and tools for audio/braille figure drawing and reading (Gardner et al. 1996).

Sounds are often important when people work together in groups. Builders on a construction site coordinate their activities in a common project by listening to their workmates hammering, shovelling, and revving engines. The usefulness of sounds in collaborative activities was demonstrated by an experiment in which a pair of people worked together to produce as many bottles of coke as possible in a computer game-like simulation of a factory. The factory consisted of nine interconnected machines, such as a heater, bottler, and conveyors, with on/off and rate controls. Each person was seated in a separate room and could see and control half the factory, and talk to the other person by microphone. The coworkers produced more coke when they could also hear the status of the machines through the clanking of bottles, the boiling of water and other everyday sounds. The sounds helped them to track ongoing processes, monitor individual machines, maintain awareness of the overall condition of the factory, and talk about the factory more easily. The activity was also more enjoyable when the sounds were on (Gaver et al. 1991).

Sounds can be very useful in circumstances where the need to move the eyes to acquire information is risky and a bottleneck for performance (Ballas 1994), such as driving an emergency vehicle or piloting a plane. In an experiment dating back to 1945, pilots took only an hour to learn to fly using a sonified instrument panel in which turning was heard by a sweeping pan, tilt by a change in pitch, and speed by variation in the rate of a “putt putt” sound (Kramer 1994a, p. 34). Radio beacons are used by rescue pilots to home-in on a tiny speck of a life-raft in the vast expanse of the ocean by listening to the strength of an audio signal over a set of radio headphones. Spatialized audio cues about runway layouts can reduce the risk of collisions on the ground by allowing the pilot to spend more time looking out the window while taxiing (Begault et al. 1996). Synthetic feedback sounds generated by instruments and tools may be particularly useful in situations where sounds cannot usually be heard, like when deep underwater in a diving suit, or out in space. When astronauts reported difficulties in tasks with power tools, the problem was fixed by equipping their spacesuits with an audio cue tied to the RPM of the power tool (Kramer 1994a, p. 35).

The soundtrack is an integral part of the modern movie, and audiences expect an ever more impressive 3D soundscape of voices, background sounds, music and special effects. The power of the movie soundtrack to engage and affect suggests that sounds may also be useful in presentations of other types of information to a general public. Most listeners can quickly and easily understand sonifications of simple functions, data distributions, and covariation between two variables (Flowers et al. 1996). The engagement of movie sound was coupled with a quantitative auditory display to produce a sonification of ozone levels in the Los Angeles basin in which a coughing sound builds to a coughing fit during rush-hour traffic. The sounds attract attention, can be understood quickly and correctly, and leave a lasting impression (Scaletti 1994). Another example is an environmental sonification in which a siren warns of sulphate concentrations in the atmosphere, a rain-like sound indicates rainfall readings, and an ominous wailing bemoans consequent deposits of acid-rain (Misenheimer and Landreth 1993).

Engineers use visualizations to analyze computer models and simulations of complex systems, An auditory display can be very useful when the simulation has an element which is acoustic. For example, the quality and completeness of a fluid flow simulation was analyzed by comparing the sonification of data generated by a simulated turbine with audio recordings of the actual turbine in motion (McCabe and Rangwalla 1994). Sounds can also make it easier to perceive cycles, rhythms, patterns and short events. A sonification was used to study a model of an artificial heart pump, in which a modulated tone indicated pressure on the pusher plate, a tapping sound indicated a blood cell entering threshold vorticity, and a drum identified the opening and closing of valves. McCabe and Rangwalla report that it seemed easier to determine the frequency of blood cells crossing threshold vorticity by listening than by looking for a change in color, and the sounds also made it easier to detect the opening and closing of the valves (McCabe and Rangwalla 1994). In a visualization for mine planning, designed by Chris Gunn from the Commonwealth Scientific and Industrial Research Organisation (CSIRO), seismic events caused by the fracturing of rock walls during a 3-month dig are depicted as red balls at the site of each event, as shown in Fig. 1.

However, the events are short compared to the duration of the data set and can flash by too quickly to notice, whilst multiple events near the same site overlap and hide each other. These problems were addressed by adding a short sound that varies in loudness with the magnitude of each event. The sounds draw attention to periods of greater seismic activity, and make it easier to detect multiple overlapping events that indicate areas of higher risk (Barrass 1998). Sounds are particularly useful when interesting variables do not appear together in the same image, or require demanding shifts in visual attention to scan. An example is a visualization designed by Simon Kravis at the CSIRO to assist hydrological engineers planning a water treatment works using a computer model. Chemical concentrations in the river are shown by colored segments that vary in response to rainfall over a 1-year period. The rainfall record is graphed in the top right of the image, as shown in Fig. 2. However, it is difficult to watch the rainfall graph and the river at the same time. The addition of a sonification of the rainfall allows visual attention to stay on the river, while quantities of rain are heard in a rain-like sound. After several repetitions,

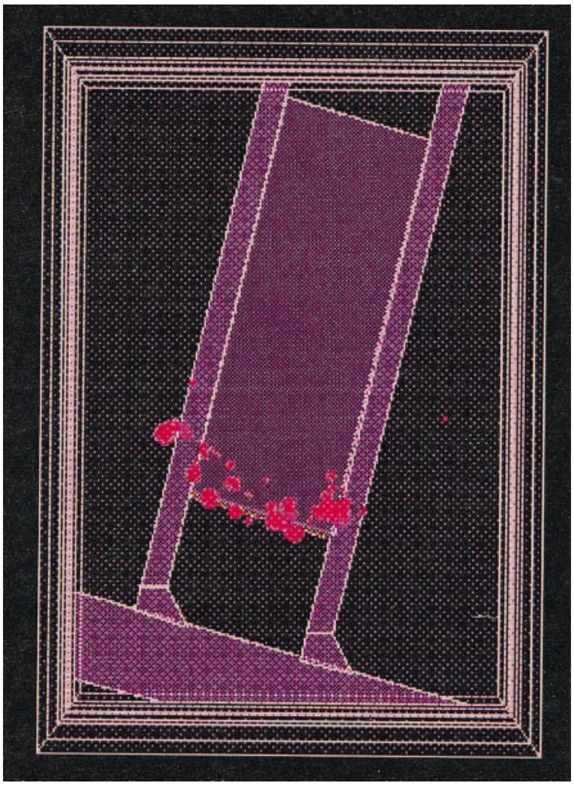

Fig. 1. A snapshot from a visualization for mine planning

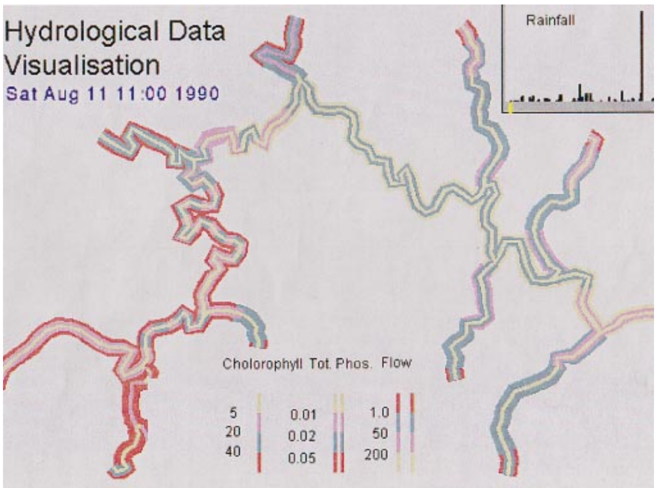

Fig. 2. A snapshot from a visualization for planning a new water-works

the listener learns the pattern of the rainfall over the year, and can anticipate a large downpour or a dry period. The soundtrack binds the set of visualizations together, making it easier to remember, integrate and understand relations between different chemicals that are never actually seen at the same time (Barrass 1998).

Sonifications can allow alternative perceptions and new insights into the data. During the Voyager-2 space mission in 1979, mission controllers became concerned by strange events as the craft began its traversal of the rings of Saturn, but could not pinpoint the problem from visual displays of the noisy data they were receiving. When the data was played through a music synthesizer a "machine-gunning" sound could be heard in the few seconds where the spacecraft had been in a region of dust concentration. This clue lead to the finding that the problems were caused by high-speed collisions with electromagnetically charged micro-meteoroids (Kramer 1994a, p. 35).

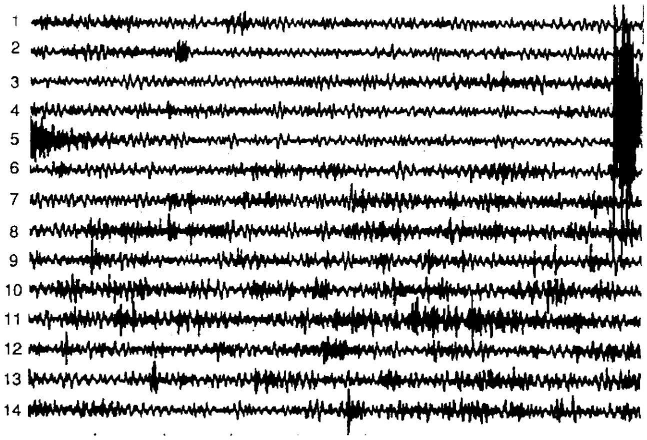

Fig. 3. A typical seismogram with contiguous horizontal 450-s segments (Hayward 1994, Fig. 10)

Large multidimensional data sets continue to be a visualization challenge. Seismic data sets are a prime example as they often contain billions of data samples. A typical graphic display of seismic data, shown in Fig. 3, may stretch over many hours or days of recording. It is difficult to get a visual overview of all this data, while at the same time retaining small but significant features and events. This problem led some seismologists to try listening to their data instead, using a direct playback technique called "audification". Audification involves speeding up the recordings 100–1600 times, so that the slow vibrations in rock shift to frequencies in the range of human hearing. Many hours of data can be heard in just a few minutes, and listeners can learn to discriminate nuclear bomb blasts from earthquakes with an accuracy of 90% (Hayward 1994). A range of seismic analysis tasks that benefit from audification are listed by Chris Hayward as overviewing large data sets, event recognition, signal detection, onset timing, model matching and education. The generality of these tasks suggests that these techniques could also be used in data mining and information visualization that likewise involve the human analysis of large, multidimensional data sets.

Our final example of sonification is the "quantum whistle". This is a sound that occurs at a size scale where you could imagine it as the lead instrument in a jam with some bacteria playing nano-guitars! For more than 30 years scientists have been searching for evidence of oscillations in superfluid gases that are predicted by quantum theory. After months of staring at oscilloscope traces in vain, James Davis and Richard Packard decided to listen to their experiment instead. They heard a faint whistling sound that is the first evidence of quantum oscillations occurring between chambers of superfluid helium (Davis and Packard 1997). This may also be notable as the first example of a historically important discovery made using sonification (Scaletti in email to the icad@santafe.edu listserver).

Although we have been speaking primarily about auditory-only displays, we expect that integrating sonification with other display modalities will be a key to their effectiveness and acceptance. Sonifications can be implemented alongside or integrated with visual, tactile, haptic, and other auditory displays such as speech. Each modality has certain strengths and each combination of modalities may produce different synergistic results. Combined visualizations and sonifications are common and most are ostensibly designed to benefit from the affordances unique to each modality (see, for example, Axen and Choi 1996; Martins et al. 1997). Like Kennel's combined tactile-auditory display or Kramer and Scott's work in progress at Stanford's Center for the Study of Language and Information (unpublished research proposal) other displays may combine the strengths of the tactile or haptic displays in spatial representation and interactivity with the strengths of temporal representations and auditory gestalts of the sonifications, producing a qualitatively different user experience.

Approaches to sonification

As you have gathered from the examples, there are many different ways to go about the design and realization of a sonification. Methods for designing an auditory display have been classified along a spectrum between analogic and symbolic by Kramer (Kramer 1994a, pp. 21–29). An analogic representation is one in which there is an immediate and intrinsic correspondence between the sort of structure being represented and the representation medium. There is a one-to-one mapping between points in the data space and points in representation space. Simple examples include the Geiger counter and auditory thermometer. A symbolic representation, on the other hand, categorically denotes the thing being represented and relations within the representation do not necessarily reflect intrinsic relationships in what is being represented. Examples include a computer beep and most automobile control notifications.

The semiotic distinctions (syntactic, semantic and lexical) have also been used to classify the best known methods of auditory display (earcons, auditory icons and parameter mapping) (Blattner et al. 1994). Earcons are short musical motifs using musical tones (Blattner et al. 1989), auditory icons are modeled on everyday sounds like a tap running or a door slamming (Gaver 1994), and parameter mappings present data variations in auditory variations such as duration, brightness, pitch etc. (Kramer 1994a; Scaletti 1994). Earcons are a syntactic method, auditory icons are semantic, and parameter mapping is lexical. This section describes these methods, along with strengths and weaknesses of each.

Let us begin with earcons. Earcons were developed to provide feedback about activities in a GUI. They are constructed by combining a lexicon of simple sounds to build more complex meanings. The lexicon may have elements that vary in rhythm, pitch, timbre, register, and dynamics. A tone "A" with pitch 440 Hz may mean "file" and tone "B" with pitch 600 Hz may mean "deleted". Combining A and B in series produces a rising tone "AB" that means "file deleted" (Blattner et al. 1989). Earcons have the following advantages:

- ease of production: earcons can be easily constructed and produced on almost any computer with tools that already exist for music and audio manipulation;
- abstract representation: earcon sounds do not have to correspond to the objects they represent, so objects that either make no sound or an unpleasant sound can still be represented.

A problem with earcons is learnability. Novices are able to learn 4–6 symbolic sounds within minutes, but further learning of up to ten can take hours. Beyond ten, the process is prolonged and some listeners may never learn the catalog completely (Patterson 1982).

Auditory icons were also originally designed to provide feedback about activities in a graphical user interface. The auditory icon approach is to map objects and events in the interface onto everyday sounds that represent reminiscent or conceptually related objects and events (Gaver 1994). Sounds in the real world are a model for sounds in the interface. For example, moving a file in a desktop GUI involves dragging it between windows. An auditory icon for this event is modeled on the sound of a real file being dragged across a real desktop. Auditory icons have the following advantages

- familiarity: everyday sounds are already familiar and may be understood very quickly;
- directness: everyday sounds can allow direct comparisons of length or size or other quantities.

Auditory icons can take advantage of experience with everyday sounds, but the most compelling sonic representations, such as a door knocking (Cohen 1994) or trash cans (Gaver 1994) may be in short supply. The development of a parametrized algorithm to mimic a real-world sound is not trivial, and recording and shaping sampled sounds is a professional skill. A key issue with auditory icons is the abstraction process inherent in representing a virtual event, such as a software operation, with a sound from a mechanical event. Such conceptual mappings may invoke learning demands not dissimilar to earcons. Lucas found no significant difference in the learnability of earcons and auditory icons measured by time taken to associate meanings, number of errors made in the process of learning, or in the improvement in these factors over two trials (Lucas 1994). An explanation of the rationale behind the sound design was the significant factor. The expectations, context and experience of the listener have significant effects on the recognition of recorded everyday sounds (Ballas 1994).

Parameter mapping is the usual approach taken to representing data as sound. Typically, a data dimension is mapped onto an auditory parameter such as duration, pitch, loudness, position, brightness, etc. Different variables can be mapped to different parameters at the same time to produce a complex sound. The parameter-mapping approach has the following advantages

- ease of production – existing tools allow mappings to many auditory parameters;
- multivariate representation – many data dimensions can be listened to at the same time.

One problem is that the sounds that are thus produced can be unpleasant. A parameter mapping to a musically interesting asynchronous granular synthesis (AGS) algorithm became irritating and unpleasant after many repetitions during evaluations (Smith et al. 1994). It is also difficult to predict how changes in a multivariate synthesis parameter will change the perception of the sound. Linear changes in AGS parameters can have complex, non-linear perceptual effects, and the range of variation can differ considerably with different parameters. Perceptual interactions between parameters can obscure data relations and confuse the listener, and a truly balanced multivariate parameter mapping may not be possible in practice (Kramer 1994b). For example, a simple bivariate display in which one variable is mapped to the position of a spectral peak and the other to attack time of a static harmonic tone will not be heard as a simple 2D space of perceptual variation. This is because both of these factors influence the perceived brightness of the sound and the listener will not be able to tell which variable is changing in some parts of the data space.

Earcons, auditory icons and a parameter mapping have all been demonstrated in the design of an interface for monitoring a steam turbine system (Albers et al. 1997). The resulting sonifications have very different sounds in them, and will sound very different. The earcons design has musical tones placed at different spatial locations that change in tempo or pitch in response to steam pressures, temperatures, and valve positions. The auditory icons design is modelled on the sounds generated by gurgling brooks, boiling liquids crackling fires, sizzling meat, and revolving fans. The parameter mapping design uses the perceptual grouping of simple tones into distinctive rhythmic streams by pitch, and sounds a lot like the bridge on the original Starship Enterprise (listen at http://neumann.computer.org/intelligent/extras/).

Issues in sonification

The range of methods with different advantages and disadvantages lead to the question of which to choose, and there is no known method for determining the best way to map data relations into sounds. When attendees at the inaugural International Conference on Auditory Display (ICAD 1992) were invited to design a sonification to support the discrimination of 6D multivariate data as either “gold” or “dirt”, the results varied widely across the submissions – ranging from significant to chance (Bly 1994). In her conclusions on this experiment, Sarah Bly called for the development of principles that address both the auditory perception and the structure of the data.

Knowledge about auditory perception can allow the designer to predict how the sonification will be heard by a human listener, and enables a theoretical evaluation of new, untried designs. The psychoacoustic theory of “perceptual streaming” (Bregman 1990), has been proposed as a basis for principles of sonification design by Sheila Williams (Williams 1994) who says “... a knowledge of the potential perceptual streams that may arise from a particular acoustic signal is essential in order to predict the interpretations for that signal ...”

However, psychoacoustic theories do not involve issues of representation that are central in sonification, where the listener needs to correctly understand data relations from the sounds. The need to consider data structure in mapping data relations into auditory relations is found in Chris Hayward’s description of why audification techniques work well for seismic data (Hayward 1994).

“One of the reasons that audification fails for arbitrary data, such as stock market figures or daily temperatures, is even a slow playback rate of 8000 samples/s requires many data points to make a sound of any duration. The resultant short sound does not reveal valuable information to the listener. Even for long sequences, the sounds do not resemble natural or familiar noises, because arbitrary data points seldom obey the physics of natural sounds transmitted through air. Seismic data, however, are an almost perfect case for audification. Seismic data sets are large. A seismic audification will sound like a recording of natural environmental sounds, because sounds transmitted through air (acoustic waves) have similar physics to seismic vibrations transmitted through the earth (elastic waves). The direct physically consistent playback can take advantage of human experience with natural sounds. A sharp explosion followed by decaying echoes includes information that is interpreted as the size and shape of the echo chamber. A set of echoes followed by the explosion is recognized as physically ridiculous or artificial.”

The need to consider the data structure was also raised by Stuart Smith in his list of obstacles to progress in sonification (Smith 1990).

“The first obstacle is the prevailing sonification model, which is simply to map data to sound parameters arbitrarily. The resulting sound is often unpleasant and usually lacks any natural connection to the data represented (one intuitively feels that medical images, for example, ought to somehow sound different from demographic data or satellite imagery). Models of sonification more sensitive to the kinds of data presented must be developed.”

An approach that links the structure of the data with the structure of heard sounds was proposed by Gary Kendall (Kendall 1991), when he observed that categorical data relations should sound categorical, and ordered data relations should sound ordered.

“Some classifications of sound events tend to be categorical. Excitation functions are typically of discrete types such as hitting, scraping, blowing, vocal glottis, etc. Some classifications of sounding objects are similarly categorical – metal, wood, hollow, solid, vocal tract, etc. These simple categorical distinctions can potentially be exploited in auditory presentations to communicate important distinctions in the data. Beyond these categorical distinctions, the essential goal is that perceptually continuous auditory attributes are scaled and mapped to data attributes in a way that is meaningful to the observer. Relevant changes in data should insure a change in what is perceived. Changes in what is perceived should signify meaningful changes in the data. The appropriate scaling functions will probably not exist a priori. Psychophysical scaling experiments may be needed in order to create perceptual scaling functions through which collections of auditory stimuli are mapped. This is made feasible only by utilizing a limited number of auditory tokens with well-understood perceptual properties. This suggests that sets of tokens be developed and scaled in advance.”

”Relevant changes in the data” is a phrase that points to the task as another important principle in designing a sonification. Different displays of the same information best support different tasks, and the usefulness of a display is a function of the task it is being used to support (Casner 1992). Gregory Kramer suggests that two broad types of tasks are important in auditory display (Kramer 1994a p. 15).

- Analysis – tasks where the user cannot anticipate what will be heard and is listening for "pop-out" effects, patterns, similarities and anomalies which indicate structural features and interesting relationships in the data.
- Monitoring – a "listening search" for familiar patterns in a limited and unambiguous set of sounds.

Tasks, knowledge and listening perception are all characterized and used as principles in Alty's goal-oriented approach for designing multimedia interfaces, summed up by four questions (Alty et al. Dolphin 1993):

- what is the goal?
- what task is needed to achieve it? * what knowledge is required?
- how is the knowledge characterized?

This approach was demonstrated in the design, realization and testing of a multimedia interface for the control system of a nuclear power plant. After a trial period, the system operators agreed that the sounds in the interface made handling alarms quicker and easier, and helped them to avoid mistakes related to the analysis of alarms. However, although the operators responded positively to the sounds, they did not take as much advantage of the sounds as was expected. They felt it took too much effort to learn the sounds, that there were too many sounds, and they did not like the low-quality voice output. In what is surely a lesson for sonification designers, the humorous air of fraud in a computer game is purposefully generated by juxtaposing low-tech sounds with the high-tech graphics (Tkaczevski 1996). The acceptance of a sonification may be influenced by the quality of the audio output, just as the perceived quality of a television set was influenced by the quality by the audio (Newman et al. 1991). This influence was recognized by engineers working on the multimedia Taligent Operating system, who were concerned that the audio be of compact disk quality after experience with the telephone quality sounds in an earlier system (Dougherty 1996). Jonathan Cohen describes some of the nastier pitfalls he encountered in designing sounds for background notification in an operating system as obnoxious sounds, ambiguous meanings, negative connotations, and incomprehensibility (Cohen 1994). He strongly suggests that an experienced sound designer should be involved in any such project. Kramer suggests that how one "feels" about the data should be reflected, where possible, in how one responds affectively to the auditory representation of the data (Kramer 1994b, pp. 214–216). Qualitative evaluations may be particularly important in designing a sonification for a commercial product.

One way to promote more widespread acceptance of sonifications is to provide easy-to-use tools and systems that allow non-experts to make their own sonifications tailored to their particular task, data, experience and expectations. However, even with the expanding universe of sound synthesis algorithms and control schemes, there is still a surprising gap when it comes to a practical sonification toolbox. Stuart Smith identifies the lack of tools as a major obstacle for sonification (Smith 1990).

"Sonification requires a general-purpose real-time sound synthesis capability, preferably at an affordable price ...[a] major obstacle is the nearly total absence of the kinds of models that allow design of computer graphics software systems that can run successfully on hardware made by many different manufacturers. The principle reasons for this situation are the lack of a satisfying comprehensive theory of timbre perception and the lack of an agreed upon theory of timbre generation. These translate directly into the situation we observe today: multiple incompatible sound-generation devices, each accompanied by its own suite of non-standard application packages."

There are problems associated with user interfacing, compatibility with different data formats, and with the perceptual scaling of auditory variables, particularly if one wishes to explore multivariate complex sounds as carriers of information. Transportability between sound devices is a problem shared with sound designers working in the computer games industry, who need to manually tweak global controls for every target soundcard system to reach a compromise that will be satisfactory across the board (Tkaczevski 1996). Although a compromise is a practical solution, it can only work when the designer knows what will be heard, and it compromises the capabilities of the better quality devices.

The need for tools and systems is also raised by Carla Scaletti in her research directions for progress in sonification (Scaletti 1994):

- Applications: further and more sophisticated examples of sonification applied to specific problems
- Sonification science: studies in the perception, cognition and neurophysiology of data-driven sound
- Systems: needs for future hardware and software include: integrated sonification/visualization languages, tools for getting from an imagined sound to a realized sound, the integration of sonification tools into mass market software like spreadsheets or statistical analysis packages, and more and better tools for exploratory sonification.

This section has highlighted some of the major issues in sonification raised by experienced researchers in the field, which can be summarised as follows:

- veridicality: the need to ensure that relations in the data can be heard correctly and confidently in the sounds,
- usefulness: the effect that a sonification has on a task,
- usability: the amount of usage required before a sonification becomes useful,
- acceptance: how much a sonification is actually used in practice
- tools: support for sonification by people who are not necessarily experts.

A multidisciplinary approach to issues in sonification

The issues raised in sonification have also been raised in other areas, such as scientific visualization, human-computer interaction, sound design for commercial products and computer music.

Veridicality is an important issue in graphic information design and scientific visualization. Principles of psychological directness, data characterization and perceptual scaling have been developed to help designers produce visual displays that are veridical for different types of data (Bertin 1981; Cleveland 1985; Casner 1992). These principles capture expert knowledge that can allow less-than-experienced designers to produce more effective displays, and also allow design decisions to be made without the need for a costly experiment. A first step toward a principled approach to sonification design has been developed by combining visualization principles for veridical display of data types with psychoacoustic observations (Barrass 1997). The approach is modelled on the application of perceptually scaled color spaces in scientific visualization, which provide the capability to specify in device-independent (transportable) terms, and to independently manipulate perceptual variables (Robertson 1988). As Smith has pointed out, the problem with sound is complicated by the lack of a small set of orthogonal parameters that adequately span hearing perception. However, the problem has been approached by targeting the space specifically to the veridical representation of categorical and continuous data types by categorical and continuous sound variations (Barrass 1998). A drawback is that the calibration of a new device requires a time-consuming series of perceptual measurements.The ongoing development of sonification principles that include a consideration of how a listener can correctly hear data structures may contribute to, and be informed by, future directions in psychoacoustic research.

Usefulness and usability are major issues in human-computer interaction (HCI) research. Methods for requirements analysis, user-centered design, and usability evaluation have already been applied in sonification (Brewster 1994; Lucas 1994; Mynatt 1994b). No doubt HCI will continue to be a strong influence. Meanwhile, sonification research continues to be informed by quantitative psychoacoustic research, and qualitative evaluations of display effectiveness will play a major role in the iterative design process.

The need to build tools for sonification that are easy to use and apply can draw on research in the computer music community into interfaces for intuitive real-time interaction with sounds. An example is a composition tool that allows the user to change a sound by grabbing it and moving it in a virtual space (Horry 1994). Also contributing to the issue of tools and systems is the more general computer science research into digital signal processing, transducer design, scheduling, compression and data formats which influence the real-time interaction and generation of synthetic sounds and soundscapes. Beyond providing tools, the wider acceptance of sonifications will be influenced by culture, style, aesthetics, affect, and production values. These are primary issues in commercial sound design, sound art, and music composition. Sonification designers can learn valuable techniques from these other sound professionals. There is also much to learn from the discourse of sound culture and sound art, and a need for sonification to actively contribute.

The overlap with these other areas points to the essentially interdisciplinary nature of this field of research. Indeed, we would suggest that multidisciplinary teams are essential to substantial progress in sonification research.

The future

There are many ways that sonification has already proven useful. The future will bring more use of sonifications in medicine, assistive technologies, computer interfaces, information communication, engineering analyzes, and cockpits. More widespread acceptance will lead to courses in sonification in the general education curriculum. Beyond more use in the same areas, it is likely that sonifications will replace some types of visual displays altogether. People will use a sonification instead of a map to drive around a strange city. Sonifications will be used in court rooms to present forensic evidence to a jury, in television news reports about the environment, and in boardrooms where decision makers analyze social and economic trends. People will listen to their own internal health, and the health of others around them. The common use of sonifications will lead to a more multimodal way of thinking that will have far-reaching effects on society as a whole. Sonification is a key technology for a multimedia society, extending sound from its accompaniment role to that of information conveyance.

Resources

If you are interested in sonification a good place to start delving further is the web site of the International Community for Auditory Display (ICAD) <http://www.santafe.edu/ icad>. ICAD is a forum for presenting research on the use of sound to display data, monitor systems, and provide enhanced user interfaces for computers and virtual reality systems. It is unique in its singular focus on auditory displays and the array of perception, technology, and application areas that this encompasses.

A history of sonification and the papers from the inaugural ICAD conference are contained in the book "Auditory Display, Sonification, Audification, and Auditory Interfaces" (Kramer 1994b).

Many examples of useful sonification can be found in the proceedings of the ACM Conference on Assistive Technologies (ASSETS)<http://www.acm.org/sigcaph/conferences/>. ACM SigSound provides a page of links <http://datura.cerl.uiuc.edu/netstuff/sigsoundLinks.html> to DSP, auditory research, computer music, sonification demos, organisations, courses of study and many other sound-related resources.

The Acoustic Society of America (ASA) is a scientific society dedicated to increasing and diffusing the knowledge of acoustics and its practical applications, and is a good place to find fundamental research on psychoacoustics <http://asa.aip.org>.

Tools for handling sounds in new ways, the perception of sounds in virtual reality, and musical aesthetics are topics in the Computer Music Journal <http://mitpress.mit.edu/e-journals/Computer-Music-Journal/>. Sound perception, art and culture are topics of the Journal of Organised Sound <http://www.cip.cam.ac.uk/Journals/JNLSCAT95/oso/oso .html, Leonardo the journal of the International Society for the Arts, Sciences and Technology <http://mitpress.mit.edu/ e-journals/Leonardo/>, and SoundSite the online journal of Sound Theory, Philosophy of Sound and Sound Art <http://sysx.apana.org.au/soundsite/>

Acknowledgements. Thanks to Simon Kravis, Chris Gunn and Chris Hayward for allowing us to use images that they generated in their work.

References

1. Albers M, Barrass S, Brewster S, Mynatt B (1997) Dissonance on Audio Interfaces. IEEE Expert, September; <http://www.computer.org/pubs/expert/>
2. Alty JL (1995) Multimedia Interface Design Tutorial Notes. In: Australian Conference on Computer-Human Interaction OZCHI'95, University of Wollongong, Wollongong, Australia
3. Alty JL, Bergan M, Craufurd P, Dolphin C (1993) Experiments using Multimedia Interfaces in Process Control: Some Initial Results. Comput Graphics 17(3)
4. Axen U, Choi I (1996) Investigating Geometric Data with Sound. In: Frysinger S, Kramer G (eds) Proceedings of the Third International Conference on Auditory Display, 1996, Palo Alto, Calif., Web-proceedings <http://www.santafe.edu/˜icad>
5. Ballass JA (1994) Delivery of Information Through Sound. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces, SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison-Wesley, Reading, Mass., pp 79–94
6. Barrass S (1998) Auditory Information Design. Ph.D. Thesis. Australian National University, Canberra, Australia <http://viswiz.gmd.de/ barrass/thesis/index.html>
7. Barrass S (1998) Some Golden Rules for Designing Auditory Displays. In: Boulanger R (ed) Csound Textbook. MIT Press, Cambridge, Mass. (in press)
8. Begault DR, Wenzel EM, Shrum R, Miller J (1996) A Virtual Audio Guidance and Alert System for Commercial Aircraft. In: Frysinger S, Kramer G (eds) Proceedings of the Third International Conference on Auditory Display ICAD'96, Palo Alto, Calif., Web-proceedings <http://www.santafe.edu/˜icad>
9. Bertin J (1981) Graphics and Graphic Information Processing. Walter de Gruyter, Berlin
10. Blattner M, Sumikawa D, Greenberg R (1989) Earcons and Icons: Their Structure and Common Design Principles. Hum Comput Interaction 4(1): 11–44. Lawrence Erlbaum Associater. Mahawah, New Jersey
11. Blattner MM, Papp AL, Glinert EP (1994) Sonic Enhancement of Two-Dimensional Graphics Displays. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 447–470
12. Bly S (1994) Multivariate data analysis. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 405–416
13. Bregman AS (1990) Auditory Scene Analysis. MIT Press, Cambridge, Mass.
14. Brewster SA, Wright PC, Edwards ADN (1994) A Detailed Investigation into the Effectiveness of Earcons. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 471–498
15. Casner S (1992) A Task Analytic Approach to the Automated Design of Graphic Presentations. ACM Trans Graphics 10: 111–151
16. Cleveland WS (1985) The Elements of Graphing Data. Wadsworth Advanced Books and Software, Monterey, Calif.
17. Cohen J (1994) Monitoring Background Activities. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 499–532
18. Craighead H (1997) Silicon Guitar; <http://www.npr.org/news/tech/970724.guitar.html>
19. Davis JC, Packard R (1997) Quantum oscillations between two weakly coupled reservoirs of superfluid He-3. Nature July 31 1997 <http://physics1.berkeley.edu/research/jcdavis/press.html>
20. Dougherty T (1996) What Does Pink Sound Like? Designing the Audio Interface for the TalOS. In: Frysinger S, Kramer G (eds) Proceedings of the Third International Conference on Auditory Display ICAD'96, Palo Alto, Calif., Web-proceedings <http://www.santafe.edu/˜icad>
21. Fitch WT, Kramer G (1994) Sonifying ther Body Electric: Superiority of an Auditory over a Visual Display in a Complex Multivariate System. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., Web-proceedings <http://www.santafe.edu/˜icad>
22. Flowers JH, Buhman DC, Turnage KD (1996) Data Sonification from the Desktop: Should Sound be part of Standard Data Analysis Software. In: Frysinger S, Kramer G (eds) Proceeedings of the Third International Conference on Auditory Display ICAD'96, Palo Alto, Calif.
23. Gardner JA, Lundquist R, Sahyun S (1996) Triangle, a Practical Application of Non-Speech Audio for Imparting Information. In: Frysinger S, Kramer G (eds) Proceedings of the Third International Conference on Auditory Display ICAD'96, Palo Alto, Calif., Web-proceedings <http://www.santafe.edu/˜icad>
24. Gaver WW (1994) Using and Creating Auditory Icons. In: Kramer G (ed) (1994) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 417–446
25. Gaver WW, Smith RB, O'Shea T (1991) Effective Sounds in Complex Systems: The Arkola Simulation, Proceedings of CHI '91 Human Factors in Computing Systems, New Orleans, April 28–May 2, 1991. ACM Press, New York, pp 85–90
26. Hayward C (1994) Listening to the Earth Sing. In: Kramer G (ed) (1994) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 369–404
27. Horry Y (1994) A Graphical User Interface for MIDI Signal Generation and Sound Synthesis. In: Proceedings of the International Computer Music Conference ICMC '94, September 12–17, 1994, Aarhus, Denmark. International Computer Music Association, San Francisco, Calif., pp 276–279
28. James F (1996) Presenting HTML Structure in Audio: User Satisfaction with Audio Hypertext. In: Frysinger S, Kramer G (eds) Proceedings of the Third International Conference on Auditory Display ICAD'96, Palo Alto, Calif., Web-proceedings <http://www.santafe.edu/˜icad>
29. Kendall GS (1991) Visualization by Ear: Auditory Imagery for Scientific Visualization and Virtual Reality. Comput Music J 15(4): 70–73
30. Kennel AR (1996) AudioGraf: A Diagram reader for Blind People. In: Proceedings of ASSETS'96 Second Annual ACM Conference on Assistive Technologies, April 11–12, 1996, Vancouver, Canada. ACM Press, New York, pp 51–56
31. Kramer G (1994a) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass.
32. Kramer G (1994b) Some Organizing Principles for Representing Data With Sound. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 185–222
33. Loomis JM, Gollege RG, Klatzky RL, Spiegle JM, Teitz J (1994) Personal Guidance System for the Visually Impaired. In: Proceedings of ASSETS'94 First Annual ACM Conference on Assistive Technologies, Oct 31–Nov 1, 1994. Los Angeles, Calif. ACM Press, New York, pp 85–91
34. Lucas PA (1994) An Evaluation of the Communicative Ability of Auditory Icons and Earcons. In: Kramer G, Smith S (eds) Proceedings of the Second International Conference on Auditory Display ICAD '94, Santa Fe, New Mexico. 7–9 Nov, 1994
35. Martins ACG, Rangayyan RM, Portelo LA, Amaro E, Ruschioni RA

(1996) Auditory Display and Sonification of Textured Images. In: Frysinger S, Kramer G (eds) Proceedings of the Third International Conference on Auditory Display ICAD'96, Palo Alto, Calif. Webproceedings <http://www.santafe.edu/˜icad>
36. McCabe K, Rangwalla A (1994) Auditory Display of Computational Fluid Dynamics Data. In: Kramer G (ed) Auditory Display : Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 321–340
37. Misenheimer R, Landreth C (1993) Caustic Sky. In: Bargar R, Evans B, Scaletti C (eds) An Introduction to Data Sonification, Siggraph'93, August 1–6, 1993, Anaheim, Calif. ACM Press
38. Mynatt ED (1994a) Auditory Presentation of Graphical User Interfaces. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 533–555
39. Mynatt ED (1994b) Designing with Auditory Icons. In: Kramer G, Smith S (eds) Proceedings of the Second International Conference on Auditory Display ICAD '94, Santa Fe Institute, Santa Fe, New Mexico. 7–9 Nov, 1994
40. Newman WR, Krickler A, Bove BM (1991) Television, Sound, and Viewer Perceptions. In: Proceedings Joint IEEE and Audio Engineering Society Meeting, 1991, Detroit, Mich. February 1–2, 1994
41. Patterson RD (1982) Guidelines for Auditory Warning Systems on Civil Aircraft. Paper No. 82017, Civil Aviation Authority, London, UK
42. Robertson PK (1988) Visualizing Color Gamuts: A User Interface for the Effective Use of Perceptual Color Spaces in Data Displays. IEEE Comput Graphics Appl 8(5): 50–64
43. Scaletti C (1994) Sound Synthesis Algorithms for Auditory Data Representations. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 223–252
44. Smith S (1990) Representing Data with Sound. In: Proceedings of IEEE Visualization 1990. IEEE CS Press, Piscataway, N.J.
45. Smith S, Levkowitz H, Pickett RM, Torpey M (1994) A System for Psychometric Testing of Auditory Representations of Scientific Data. In: Kramer G, Smith S (eds) Proceedings of the Second International Conference on Auditory Display ICAD'94, Santa Fe, New Mexico
46. Smith S, Picket RM, Williams MG (1994) Environments for Exploring Auditory Representations of Multidimensional Data. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 167–184
47. Stevens RD, Brewster SA, Wright PC, Evans ADN (1994) Design and Evaluation of an Auditory Glance at Algebra for Blind readers. In: Kramer G, Smith S (eds) Proceedings of the Second International Conference on Auditory Display ICAD '94, Santa Fe Institute, Santa Fe, New Mexico. 7–9 Nov, 1994
48. Swan N (1996) Ways of Seeing. The Health Report. Radio National Transcripts, Monday, 19 February 1996; <http://www.abc.net.au/rn/talks/8.30/helthrpt/hstories/hr190201.htm>
49. Tkaczevski A (1996) Auditory Interface Problems and Solution for Commercial Multimedia Products. In: Frysinger S, Kramer G (eds) Proceedings of the Third International Conference on Auditory Display ICAD'96, Palo Alto, Calif., Web-proceedings <http://www.santafe.edu/˜icad>
50. Williams SM (1994) Perceptual Principles in Sound Grouping. In: Kramer G (ed) Auditory Display: Sonification, Audification and Auditory Interfaces. SFI Studies in the Sciences of Complexity, Proceedings Volume XVIII. Addison Wesley, Reading, Mass., pp 95–126

STEPHEN BARRASS is a research scientist at the German National Research Centre for Information Technology, where he is exploring interactive sonification on the virtual workbench and in the CyberStage immersive environment. He obtained a Ph.D in Information Technology from the Australian National University in 1998, and a B.E. in Electrical Engineering from the University of N.S.W. in 1987. He is a member of IEEE and ACM.

GREGORY KRAMER inaugurated the International Conference on Auditory Display, chaired the ICAD 92 and ICAD 94 conferences and co-chaired ICAD '96. He co-edited the ICAD '94 and '96 proceedings. He also chaired a National Science Foundation workshop convened to report on the state of sonification research and recommend a research a research agenda. The White Paper emerging from this effort ist helping to further define the field of data sonificatin. He also founded the non-proft corporation International Community for Auditory Display and is now Chairman Emeritus. He edited the first book published in this area, "Auditory Display: Sonification, Audification and Auditory Interfaces" (Addison Wesley). Kramer lectures on this topic at research institutions worldwide. He has been a member of the Santa Fe Institute since 1989 where his area of concentration has been sonification of high dimensional systems and understanding how we comprehend complexity. Kramer is on the Editorial Board of the MIT journal Presence. A National Endowment for the Arts Composition Fellow, and Assistant Professor at New York University from 1975–79, Dr. Kramer also founded the non-profit arts facility studio PASS in New York City. Currently Dr. Kramer is the president of Clarity, an auditory display research and development company, where his areas of concentration include applications of auditory display to monitoring and analysis applications. Kramer holds patents in audio signal processing and sonification and directs ongoing efforts to license Clarity's patents to the music and PC industries. A Vipassana meditation instructor, Greg lives in Portland, Oregon with his wife and three sons.