

Mapping a single data stream to multiple auditory variables: A subjective approach to creating a compelling design

Greg Kramer, *Clarity and the Santa Fe Institute*

Introduction

Representing a single data variable changing in time via sonification, or using that data to control a sound in some way appears to be a simple problem but actually involves a significant degree of subjectivity. This paper is a response to my own focus on specific sonification tasks (Kramer 1990, 1993) (Fitch & Kramer, 1994), on broad theoretical concerns in auditory display (Kramer 1994a, 1994b, 1995), and on the representation of high-dimensional data sets (Kramer 1991a & Kramer & Ellison, 1991b). The design focus of this paper is partly a response to the others who, like myself, have primarily employed single fundamental acoustic variables such as pitch or loudness to represent single data streams. These simple representations have framed three challenges:

Behavioral and Cognitive Science—Can sonifications created with complex sounds changing simultaneously in several dimensions facilitate the formation of a stronger internal auditory image, or audiation, than would be produced by simpler sonifications?

Human Factors and Applications—Would such a stronger internal image of the data prove to be more useful from the standpoint of conveying information?

Technology and Design—How might these

richer displays be constructed?

This final question serves as a starting point for this paper. After years of cautious sonification research I wanted to explore the creation of more interesting and compelling representations.

My focus on a single data variable changing in time emerged from two sonification challenges. The first arose in my work with Dr. George Mawko, beginning in 1993, on the sonification of data from Radionuclide Ventriculography (RVG)—a non-invasive means for obtaining the blood volume change of the left ventricle (Mawko & Kramer, 1995). The essence of this problem was an eyes-busy task of watching the RVG "movie" (called a cine loop) while trying to integrate a simple XY plot presented alongside the movie (see Figure 1). Perceptually integrating the temporal correlation of these two images is difficult, primarily because analysts must glance from one image to the other as both change. The second challenge, which materialized largely through discussions with Dr. Larry Scadden and, later, Dr. John Gardener (see Gardner et al in these proceedings), was the presentation of XY plots and other low-D data to vision impaired persons.

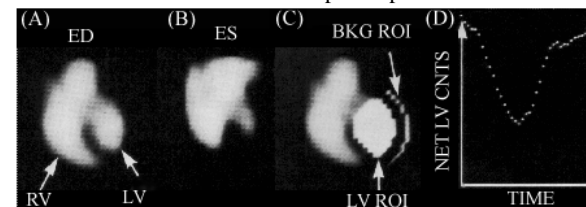


Figure 1

Both of the challenges mentioned above involved audibly representing a single data variable changing over time and presented an opportunity to approach sonification from a qualitatively different viewpoint than that explored in previous

work. The design task shifted from "How can we possibly represent all of these variables?" to "How can we effectively represent this one data variable in the most compelling way?" In rendering this data I hoped to not simply generate what might be called higher production values in auditory displays, or more attractive surface qualities for their own sake, but rather to represent the data in such a way as to enhance the information-conveying capacity of the sonifications. I also hoped to reduce listening fatigue and annoyance. The increased complexity would thus be in the service of functionality as well as aesthetics for their own sake.

Techniques

The data sets employed in the audio examples below include the heartwall (RVG) data and variations in the gravity of Venus gathered during the Magellan fly-by. In all of the designs presented, a combination of scaling and multiple mappings was employed. I defined multiple mapping as "...the routing of input data to more than one auditory variable" ([Kramer 1994b](#), p. 201). I defined scaling in this way: "Scaling the range of auditory variables may be accomplished by narrowing the upper and lower limits of that variable in the target sound generator or, in the case of mediating structures, narrowing the range at the input to those structures...This scaling technique can be thought of as effecting the number of just-noticeable-differences (JNDs) that are traversed by the auditory variable per change in the controlling data value."([p.200](#))

Caveat: This work is not intended to be psychoacoustically rigorous. Variables may interact to yield uneven results from a strictly psychoacoustic standpoint. I also have made little attempt to quantify the acoustic parameters. For

example, pulsing speed or amplitude envelope durations are described approximately and, as usual, timbre changes are very difficult to describe in words. Every attempt has been made to describe sound in experiential rather than technical terms. This work is presented as an exploration of an approach to design. The benefits may or may not outweigh the drawbacks in any particular task. It is my hope that the sound examples convey my intentions better than these verbal descriptions.

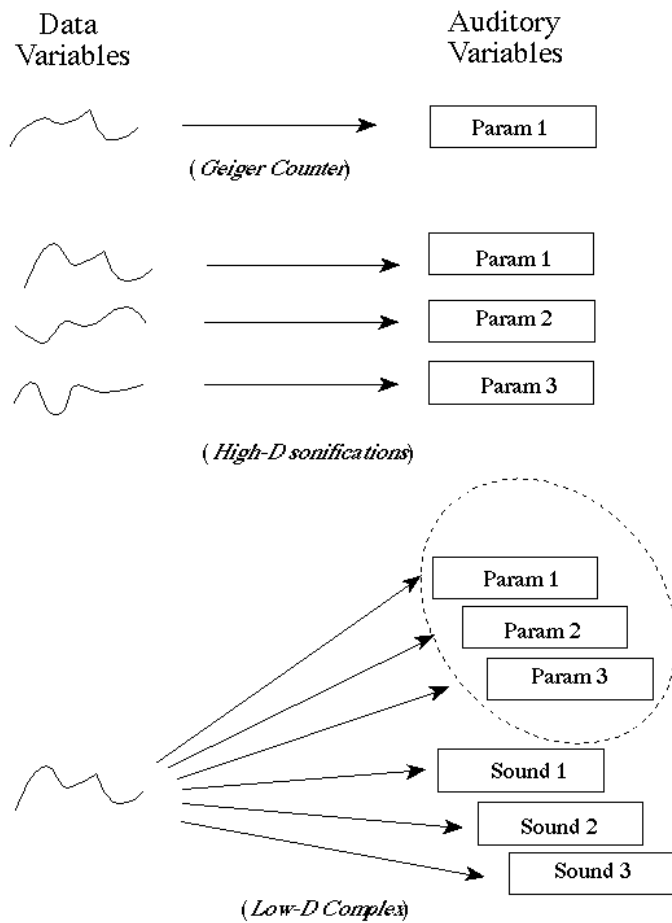


Figure 2

Building on insights from prior work

As a benchmark I decided to recreate, on current equipment, the RVG sonification I had produced in earlier work. This involved

using the single data stream to control the pulsing speed, brightness, pitch and loudness of a sound. In prior work these four parameters were often used to display four independent data variables (Figure 2). With minor technical differences necessitated by hardware changes this experiment went smoothly. By scaling these four variables to ranges that met our criteria of producing subjectively compelling and not displeasing results, we arrived at a sonification that evidently, by our informal assessment, facilitated the distinguishing from healthy and unhealthy hearts. The relatively simple sounds and variables produced results which, if not especially exciting, were encouraging as regards to accomplishing a task. (Sound examples 1 and 2)

[Sound 1](#): 4 simple variables—Healthy heart

[Sound 2](#): 4 simple variables—Unhealthy heart

Some design notes: Pitch changes of greater than an octave were annoying; an excursion of a major third to a fifth was adequate and smaller excursions were acceptable. Pulsing speed, with amplitude envelopes of about 140 ms, ranged from about 80 to 200 pulses per minute. Brightness was obtained by controlling a lowpass filter and ranged from the full spectrum sound (i.e. unfiltered=brightest) to passing the fundamental pitch but few harmonics. Loudness ranged from full scale of a comfortable listening level to about half of this level. For some of these variables, e.g. pitch and loudness, it seems that perhaps traversing too many JNDs with a sonification makes it more difficult to listen to despite the enhanced resolution a sensitive display has to offer.

This work was extended in a second rendering—created using similar variables (pitch, loudness, brightness) but with more spectrally complex sounds, filtering, and distortion processes. As before, the pitch excursion was minimal but important for enhancing one's ability to perceive the data's contour; i.e., it seemed to bring the sonification into "focus". Loudness was data controlled but also changed very little, while a 4-pole digital filter provided complex changes roughly akin to brightness. A digitally controllable distortion process served to add detail to this rendering, particularly on the mid to higher data values.

In Sound example 3 the RVG data from a healthy heart is used.

First we introduce the basic sound. Then we introduce the four controls/processes: first loudness, then pitch, then brightness, and finally the distortion process. In Sound example 4 we employ this same multiple mapping using the unhealthy heart data as the control source. In sound example 5 we use the Magellan Venus gravity data to control these same four parameters.

[Sound 3](#): Basic parameters, complex sound—Healthy heart

[Sound 4](#): Basic parameters, complex sound—Unhealthy heart

[Sound 5](#): Basic parameters, complex sound—Gravity of Venus

Controllable Complexity and Layering

In early unpublished sonification experiments I have worked with different waveforms being mixed together under the control of data ([Kramer et. al., 1991b](#)). That is, one data stream would move a window through a collection of sounds. These sounds were presumably arranged such that going from sound 1 to sound 2 (s1 to s2) would produce a timbre change that would be extended as the window moved onto s3 then s4. The amplitude curves of the window were designed to produce a roughly consistent power output as different waveforms were accessed thus providing a continuously changing timbre that was perceived as one auditory stream but constructed from many acoustic sources. This approach as we implemented it was not terribly successful and lay fallow until now.

In sound examples 6, 7, and 8 below we layered five different sounds to achieve controllable complexity. The loudness of all five sounds was controlled by the data, causing the auditory rendering to fade in and out with changes in heartwall pressure, gravity, etc. However the sounds were designed with loudness shapes that resulted in different fade-in times and rates. Superimposed upon this layering technique, pitch and loudness of the sounds were both co-varied. It is interesting to note that when identical parameters of different sound sources co-vary, the gestalt principle of common fate tie them together, or causes streaming.

Highlighting the peaks

Another instance of subjectivity in our design process is that rather than seeking a one to one correspondence between the display space and the data, we took the license to highlight higher data values (as had been done with distortion in the previous example). Employing a modified version of the cross-fading technique from our earlier work, lower data values caused duller sounds to fade in and disproportionately brighter/rougher sounds were used to represent higher data values. This can be likened to a designer using white to represent the peaks of a mountain range on a map while lower elevations are simply represented by changes in green, yellow and brown tints. The suggestion of snow in such hypsometric tints supports the discontinuity. For a further look at the use of color in highlighting visual displays, see ([Rogowitz and Treinish, 1993](#)).

The design process described above was quite subjective and in this way stands in contrast to Stephen Barrass's excellent and methodical work using John Grey's timbres to achieve perceptually consistent mappings between changes in a data space and corresponding changes in perceptual space ([Barrass, 1996](#)). In some cases we sought a blending of sounds due to similar timbres. In other components, e.g. the cymbal sound, we took advantage of the differentiation due to a greater concentration of upper partials and higher noise content to create the highlighting effect. As we built the sound up we listened to the sonifications, keeping in mind the overall sound of the display and the types of data we intended to represent.

Sound example 6 will demonstrate building of the sound one component at a time. The loudness of all the waveforms is controlled by the RVG data from a healthy heart. First you hear a filtered orthogonal waveform. Then we add an identical waveform at a slightly different pitch which adds a warble to the sound to animate it. We then add a noise- modulated oscillator which, with the segment of vocal sample that follows lends body and roughness. Finally we add the cymbal sound to highlight the highest data values. The pitch of these individual sounds is then controlled by the same RVG data. This supports the streaming and

amplifies the contour in the data.

[Sound 6](#): Healthy heart—layered sound

[Sound 7](#): Unhealthy heart—layered sound

[Sound 8](#): Gravity of Venus—layered sound

Discussion

There are three important assumptions that guided this work. The first is what might be called a "faux-Gibsonian logic". Such a logic might ask: "Since our auditory systems evolved to hear complex sounds, might complex sonifications convey data more effectively than those employing pure waveforms". Additionally might not some more complex sonifications be easier to listen to? My personal experience, and that of my studio colleagues, is that the complex sonifications were richer and easier to listen to than those created using purer, simpler sounds. At this point it is not possible to say whether the new sonifications conveyed the data more effectively, but we agreed that they were more suggestive of a complex phenomena such as heart wall pressure changes and therefore involved us more readily in the display's metaphor. The second assumption guiding this work was that complexity is an expected and necessary feature in creating a successful overall acoustic ecology. Truax (1984) suggests that an acoustical ecology requires variety, complexity and balance. Because one sound is used in each sonification, variety may not seem very relevant to the illustrations. However, if a display user is working with multiple data sets and/or working over an extended period of time (as I often have), variety between sonifications can be most welcome. Complexity in a single sonification display may be introduced by the complexity of the display's timbre and its evolution over time. Balance as defined by Truax (spatial, temporal, social and cultural) was not considered in this work.

The third assumption of this work relates to metaphor as employed in our selection of data to sound parameter mappings. In all the cases demonstrated an increase in data values was mapped to higher pitch, brighter sound, and generally more complexity, such as the warble, modulated noise and breath sample in sound example 6. While my work with Bruce Walker preliminarily

indicates that intuitive selection of metaphors in auditory displays may not be a consistently productive approach (see Walker and Kramer in these proceedings), I found it useful, when working with these multiple mappings, to use metaphor in an intuitively consistent manner.

Closing Thoughts

While we hesitate to say that they produced more accurate renderings of the data, our experience with the enhanced-complexity sounds we generated was that they were subjectively richer and easier to listen to. As we worked with the data we could tell when the display resolution was too low and when the auditory image was stronger and more focused. My suggestion to the designer in combining sounds to create a complex sonification is to be wary of unwanted streaming effects, particularly if the sounds are timbrally disparate or occupy different frequency ranges. I also suggest watching for streaming based upon frequent radical changes in the data. For sonification researchers without a musical background it may be helpful to collaborate with someone who is familiar with the task of working with sound "on its own terms". It is worth noting that if you create sonifications with sonic detail you'll need a playback system of sufficiently high resolution. Stereo imaging, frequency range, and lack of distortion will be important to the communication of the data.

Acknowledgments

I'd like to acknowledge the creative input and facilities access provided by Clark Salisbury and David Fulton and editorial input from R. Jonathan Kramer.

References

- Barrass, S. (1996). Sculpting a Sound Space with Information Properties. *Organised Sound*(pp.1-2). Cambridge University Press.
- Fitch, T., and Kramer, G.(1994). Sonifying the Body Electric: Superiority of an Auditory over a Visual Display in a Complex, Multi-Variate System. In G. Kramer (Ed.)*Auditory Display: Sonification, Audification and Auditory Interfaces*Santa Fe Institute, Studies in the Sciences of Complexity Proceedings (Vol.

XVIII). Reading, M: Addison-Wesley Publishing Company.

Kramer, G. (1990). Audification of the ACOT Predator/Prey Model. Unpublished research report prepared for Apple Computer's Advanced Technology Group, Apple Classrooms of Tomorrow.

Kramer, G. (1991a). Audification: Using Sound to Understand Complex Systems and Navigate Large Data Sets. Proceedings of the Santa Fe Institute Science Board, Santa Fe Institute.

Kramer, G., and S. Ellison. (1991b). Audification: The Use of Sound to Display Multivariate Data. Proceedings of the International Computer Music Conference, pp. 214-221.

Kramer, G. (1993). Sonification of Financial Data: An Overview of Spreadsheet and Database Sonification, The Proceedings of Virtual Reality Systems '93, SIG Advanced Applications, New York, NY.

Kramer, G. (1994a). An Introduction to Auditory Display, In G. Kramer (Ed.), Auditory Display: Sonification, Audification and Auditory Interfaces, SFI Studies in the Sciences of Complexity Proceedings (Vol. XVIII), Reading, MA: Addison-Wesley Publishing Company.

Kramer, G. (1994b). Some Organizing Principles for Representing Data With Sound, In G. Kramer (Ed.), Auditory Display: Sonification, Audification and Auditory Interfaces, SFI Studies in the Sciences of Complexity Proceedings (Vol. XVIII), Reading, MA: Addison-Wesley Publishing Company.

Kramer, G. (1995). Sound and Communication in Virtual Reality. In F. Biocca @ M. Levy (Eds.) Communication In The Age Of Virtual Reality, Lawrence Erlbaum Assoc.

Mawko, G. and Kramer, G. (1995). Sonification for assessment of regional function in Radionuclide Ventriculography, unpublished proposal.

Rogowitz B.E. and Treinish L.A. (1993). An Architecture for Rule Based Visualization, In Proceedings of IEEE Visualization '93,

San Jose, California.

Truax, B. (1984). Acoustic Communication, 43. Norwood, NJ: Ablex Publishing Corp.

Author

Greg Kramer
Clarity/Santa Fe Institute
310 NW Brynwood Lane
Portland, OR 97229
(503)292-8559
fax (503)292-4982
kramer@listen.com