ARTSCILAB

Ways of Listening for Information

A Vague Taxonomy of Sonification Techniques

“Listen” by Hsin-Chien Huang (used with Permission)

Author: Scot Gresham-Lancaster https://orcid.org/0000-0002-8400-6041

Co-Authors: Sharath Chandra Ram https://orcid.org/0000-0001-7984-9442

Roger Malina https://orcid.org/0000-0003-3399-3865

Art Sci Lab Grey Paper #9 ATEC UTDALLAS Richardson, TX

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

Summary

It is becoming more and more difficult for anyone to make decisions based on the growing amount of big data. The dominant methods today are based on visualized information. In this work we present a vague taxonomy of ways to turn data into sound — data sonification — and argue that better decisions can be made based on visualization augmented by sonification or by sonification alone.

This paper will summarize the discoveries and results of ongoing in-depth exploration into this area of research. We have identified 5 major areas of categorization: Alerts and Alarms, Audification, 1st order sonification, 2nd order sonification, and Machine Listening. This last category “Machine Listening” identifies a hot new area of research and development in the field; machines automatically listen to data encoded signals to mimic human auditory perception using emerging methods in AI, audio, and speech signal processing. These techniques enable us to analyze information rich signals before they have even been turned to data, in what we call ‘pre-data sonification’

All of these varied approaches to sonification necessitate a “vague” taxonomy to deal with the fuzzy conceptual boundaries of application that can vary considerably according to context or conditions. Additionally, it is the case that some sonification involve more than

one of these types, and some types overlap with others as a complex 'Venn diagram' of shared attributes.

The purpose here is to clarify the classifications of techniques and their interactions with such a "vague" taxonomy. Many types have been created and researched by the team over the last decade. To illustrate this point are some detailed descriptions of some of the various techniques implemented in our work in a variety of computer music languages that will be described for each technique and classification. To further summarize our work, we will be using some of our modules based on this taxonomy to listen to data and allow the listener and users of machines make better decisions quickly and to facilitate the distribution of these various techniques on a large scale.

Table of Contents

1. Introduction - Historical Context

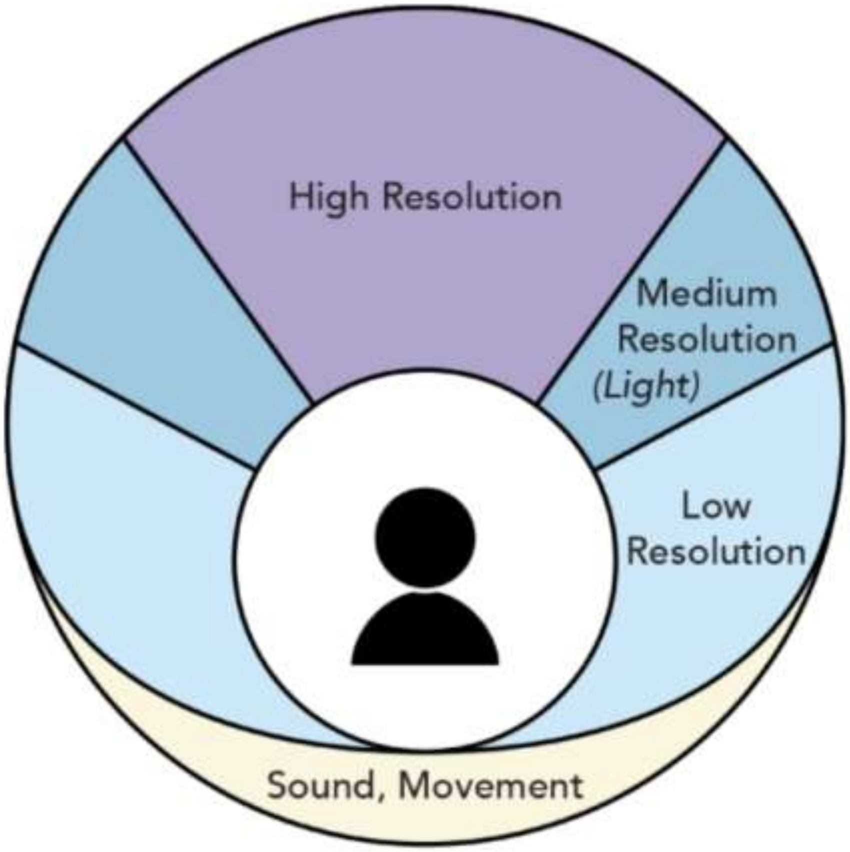
Cognitive overload was an early concern of human computer interface (HCI) innovators. The relationship to "real world" ubiquitous computing, now referred to as IoT, (Internet of Things) was being explored in depth in many early papers. Most cited is the work at Xerox PARC of researchers Mark Weiser, Rich Gold and John Seely Brown in the early 1990's.[1] [2] This way of thinking about stepping outside of the standard computer interface and integrating the user interface was followed by a measured HCI examination of the cognitive load of the user experience.

---

[1] Mark Weiser: The world is not a desktop. Interactions 1(1): 7-8 (1994)

[2] Mark Weiser, Rich Gold, John Seely Brown: The Origins of Ubiquitous Computing Research at PARC in the Late 1980s. IBM Systems Journal 38(4): 693-696 (1999)

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA



Areas of Attention

This early thinking had a profound impact on the ideas that led to later sound design investigations, many of which were later identified as "sonification". The first use of the term sonification can be traced to the writings of Carla Scaletti in 1994 when she proposed a formal working definition for her investigation of auditory data representation, as:

"A mapping of numerically represented relations in some domain under study to relations in an acoustic domain for the purposes of interpreting, understanding, or communicating relations in the domain under study" [3]

That same year Gregory Kramer and Bruce Walker stated the most familiar definition used:

"Sonification is the use of non-speech audio to convey information or perceptualize data" [4]

By 2008 Thomas Hermann in his first taxonomy where he endeavored to refine this definition:

Sonification Definition: A technique that uses data as input, and generates sound signals (eventually in response to optional additional excitation or triggering) may be called sonification, if and only if:

(C1) The sound reflects objective properties or relations in the input data.

(C2) The transformation is systematic. This means that there is a precise definition provided of how the data (and optional interactions) cause the sound to change.

(C3) The sonification is reproducible: given the same data and identical interactions (or triggers) the resulting sound has to be structurally identical.

---

[3] Scaletti, C 1994, Software synthesis algorithms for auditory data of an extensible software framework for sonification research and auditory display. *In Proceedings of 13th International Conference on Auditory Display, Montreal, Canada June 26-29, pg 224*

[4] Kramer, Gregory, ed. (1994). Auditory Display: Sonification, Audification, and Auditory Interfaces. Santa Fe Institute Studies in the Sciences of Complexity. Proceedings Volume XVIII. Reading, MA: Addison-Wesley. ISBN 978-0-201-62603-2.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

(C4) The system can intentionally be used with different data, and also be used in repetition with the same [5]

A more complete articulation of "Common Auditory Display Classes as Common Functions in Systems" was put forward by Nees and Walker in 2009: [6]

- Alarms - Brief, simple, sounds that capture attention
    - *Alerting, -warning*
- Auditory icons - Environmental sounds; ecologically relevant sounds
    - *Object, status, and process indicators; auditory menus*
- Earcons - Brief, abstract motifs with rule-based iterations
    - *Object, status, and process indicators; auditory menus*
- Spearcons -Brief, accelerated speech
    - *Object, status, and process indicators; auditory menus*
- Auditory graphs - Data mapped to frequency
    - *Data exploration aids*
- Audification - Periodic data sampled within audible range driving frequency
    - *Data exploration aids*
- Model-based sonifications - Various (extensive)
    - *Data exploration aids*

[5] Hermann, Thomas, Taxonomy and Definitions for Sonification and Auditory Display ICAD proceedings 2008

[6] Nees, M.A., & Walker, B.N. (in press). Auditory interfaces and sonification. In Stephanidis, C. (Ed.), The Universal Access Handbook. New York: CRC Press, pp. 509. (with the addition of pointers to our related research)

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

- 3-D audio displays -Virtual spatial audio (for VR via HRTFs)
    - *Spatial-orienting cues; navigation aids*
- Soundscapes - Various, often naturalistic
    - Ongoing status indicators; monitoring aids
- Audio in arts and entertainment - Various and most widespread version of sonification
    - Sonification as art; aids for enhanced and accessible experiences of exhibitions, games, etc.

The sonification research group at the Art Sci Lab has been working from these classification that were set out in 2014 after the groups preceding 3 years of research:

2. ATEC ArtSciLab Research Template

We continued work with these previous definitions and created some of our own guidelines and ideas as part of our long-range plan for the Data Stethoscope project. We first developed four general classes of sonification types.

2.1 Four classes:

- **Soundscape** – environment of events that represent multiple parameters that are changing in time; an accurate representation of the state of a multi-parameter complex data set over time

- **Sequence** – traversing through a data set, the movement is contrived as a journey across the data; NOT from the data
- **Audification** - direct unanalyzed mapping of analog signals and time varying data to sound (non-speech) or speech
- **Extrinsic** - (i.e. exformation) **–** time series ongoing which does not represent the data, but is modified acoustically, temporally, by the data.[7]

Each of these 4 classes or sonification types can trigger perceptual alerts, alarms or be perceived as sonic icons: symbolic punctuation of sonic occurrences that are perceived as discrete events. Both the Soundscape and Extrinsic classes can be thought of as subclasses of a larger acoustic ecology, whereas Sequences and Audification are encapsulations of gesture that can be parts of larger soundscapes but are not soundscapes in themselves. This opens up a complex web of interconnected audio centric user interface contexts which are outlined here but will require the rest of this paper to clarify individually. Some are more self-evident than others.

In our day to day world we are aware of the nagging turn signal clicking while driving and we have already changed lanes or distracted and looking away. We hear the water container begin to fill to the top and stop pouring. The actual use of audio in real world application are few and far between. The Geiger counter, the hang glider pilot's Vario or the

[7] Unpublished From: Lead Author Scot Gresham Lancaster with Andrew Blanton, Frank Dufour, Roger Malina UTDallas ArtSci Lab in an unpublished internal research document

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

submarine crew member sonar are some of the only tangible instances of sonification at scale.

However, there is a complex set of possibilities that the conversion of data to audio represents. The following outline can be read as a more elaborate version of the Table of Contents but was also provided to give the reader an overview of the concepts driving the models that we have built. Section 2.2 can be thought of as the executive outline of what is explained in the rest of the paper that follows.

2.2 The Five Types of Sonification Categories in Each Class

- **2.2.1 Alerts and Alarms**
  - All types of "Listening to Data" can trigger both Alerts and Alarms
  - Alerts provide warnings
  - Alarms are a call to action (getting out of bed instead of hitting snooze, for example)
- **2.2.2 Audification**
  - Direct mapping of data or raw signals into sound
    - Pre-data listening where signals are monitored directly in the audio frequency range.
  - Analog and time varying signal in the audio domain
    - Vinyl records directly convert audio electromechanically from records

- Early researchers converted spectral analog computing signals to waveforms that were listen to as audio spectrum formant playback with speech like results.[8]
- Real world sensors streamed to a digital to audio convertor
  - Electromagnetic to auditory
  - Physical manifestation
  - Or simply grinding a rock across a grate

2.2.3 1st order sonification

- Abstracting the data flow by only one degree of abstraction. For example, the value of data mapped directly to frequency of the sound signal.
- Earcons[9]
  - Characterized as brief, abstract motifs often with rule-based iterations
  - The most extensive research is from Laboratoire Musique et Informatique de Marseille called Unités Sémiotiques Temporelles (UST) or 10

[8] F. S. Cooper, "Some Instrumental Aids to Research on Speech," Report on the Fourth Annual Round Table Meeting on Linguistics and Language Teaching, Georgetown University Press, pp. 46-53,1953.

[9] Viewed in July 2019 http://sonification.de/handbook/index.php/chapters/chapter14/

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

- TSU's are described as sound objects and defined as semiotic objects. They can be thought of as generalized sonic gestures found in reality and musical composition both. (example: note sliding up indicates going upwards, etc.)
- When realized as phrases of musical material shapes these earcons may be displaced in scale, both in time and frequency and still retain a recognized shape and/or quality.
- Spearcons are a subset of Earcons and are characterized as brief, accelerated speech. They are a very specific application of Text to Speech technology for the purpose of creating audio menus.

- Scaling time and values of data streams
  - Remapping to Pitch (chromatic musical notes)
  - Remapping to a musical mode or scale
  - Each new event forced to a timing grid (rhythm)
  - Spectral remapping (tone/timbre)
  - Dynamic scaling of amplitude
  - Combinational configurations of all these are possible

2.2.4 2nd order sonification

[10] Viewed on July 2019 http://www.labo-mim.org/site/index.php?2013/03/29/225-temporal-semiotic-units-tsus-a-very-short-introduction-

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

Introduction to the concepts of 2nd order two degrees of abstraction

or data points made into sound in larger models

(Often referred to as Model based sonifications)

- Gestural - Melodies and Sequences of remapped data
    - direct "audification" like in 1st order sonification along with "reading" the data directly from gestural interaction or sonification of movement
    - Interacting with software instruments incorporating physical models synthesis types[11]
- Environmental Soundscape
    - Natural habitat emulation with sounds of natural phenomenon as semiotic markers with meaning (frogs, crickets, wind, rain, wolf howls etc.)
    - Noise on the factory floor (unstabilized state machine)
- Nodes in complex networks as generative machines
    - Temporal and gestural change as generator
- Exformation driven (Tor Nørretrander)[12]
    - Perturbation of "known" audio sources

[11] Perry R. Cook, Real Sound Synthesis for Interactive Applications - CRC Press Book. ... Applications. 1st Edition.

[12] Tor Nørretranders (1998). The User Illusion: Cutting Consciousness Down to Size. Viking. ISBN 0-670-87579-1.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

  - (familiar music or audio get varying amounts of reverb, for example)
- Familiar music and sound perturbed by temporally changing data
  - (establishing a familiarity by repetition then perturbing values in the now familiar new sonic context)
  - Ecological- soundscapes and "perturbed" prerecorded known musical pieces
- Visually driven reaction to screen based interactive models

- **2.2.5 Machine Listening**

  - Machine Listening is a field of computational auditory scene analysis systems that involves analysis by machine of signals in the audio spectrum.
  - All the techniques of sonification above can be reoriented to allow machines the opportunity to "listen" in their unique way and to contribute yet another approach to gleaning understanding and information from raw data sources.

- **2.2.6 DATA LISTENING: Techniques and Classifications**
  - Three techniques of Data Listening
    - Pattern Creation
    - Monitoring
    - Search and Discovery
  - Three classifications of Data Listening
    - Analog Conversion

- Data Synthesis
- Signal Processing

3. Presentations of Modules Using the Taxonomy

In this section each of the classification types will be examined in more detail and then there will be a description and examples of various realizations that have been used in our research. Much credit must be given to our colleagues and forbearers in the ICAD community. ICAD (International Conference for Auditory Display) 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. This is the segment of the market for audio visual display which is estimated at $26.5 billion.[13] Also, there has been a surprising lack of established conventions compared to visualization (e.g. road signs) and the fact that there is variability of interpretation tied to different cognitive traits that can be effected by culture and context. Our ongoing discoveries over the last few years of informal and heuristic testing have led to

[13]Viewed July 29th https://www.marketwatch.com/press-release/global-audio-visual-display-market-to-reach-usd-265-billion-by-2025-2018-10-25

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

the understanding of the necessity to re-test after new users are oriented to the expectations of each of these audio realizations. This has led us to integrate an approach of Learning, Exploring and then Testing the result of each particular technique. If the technique under discussion has been accompanied by user testing, the results of that testing will be added.

This includes the benefits of cognitive science as a means of understanding efficiencies this includes Auditory Cognition[14] and Music Cognition.[15] The effect of both these are part of the interaction with User interface design.

3.1 Alerts and Alarms

Alerts are familiar to all of us. In terms of this “fuzzy” vague taxonomy it is the most obvious example of a category that is present in all classifications of sonification. So, this first classification is actually a broader classification that reaches across all the techniques and types that we have identified. In the right context, each distinct type of technique articulated in the outline above can create a sound that will be heard as an alert or an alarm. The nature of auditory perception is that sudden changes in our acoustic setting

[14] Kramer, Gregory. Auditory Display: Sonification, Audification, and Auditory Interfaces. Reading, Mass: Addison-Wesley, 1994. Print.

[15] Dowling, W. J. (2010). Music perception. In C. Plack (Ed.), Oxford handbook of auditory science: Hearing (pp. 231-248). Oxford, UK: Oxford University Press.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

instinctually trigger alarm.[16] Throughout our research we have attempted to create instances where our intent of focusing attention to important isolated moments of change and disparity were made obvious. When successful this sort of auditory trigger translates immediately, possibly instinctually, a warning (alert) or a call to action (alarm). The example of the rooster crowing in the morning tells the rural sleeper that sunrise is up them. Later in the mechanical world the Alarm clock is triggered by time, warns the listener that it is time to prepare to take action. It is time to pay attention to the intention encoded in the act of setting the alarm. In the richer context of multi-dimensional triggering, the listener can take cues from many parts of the sound stream. Multiple sonic events that are associated with these complex scenarios can be realized as sophisticated model-based sonification; this is an example of using several types of sonification at once, hence our use of a 'vague' taxonomy. These models can have various trigger types that emerge from the collision of streams of sounds. These streams create a new state in a soundscape and alert the listener to the new context.

Credit to Samuel Bordreuil[17] for making this distinction between the Alert and the Alarm.

Alert	**Alarm**

[16] Edworthy, J, Adams, A, Warning design: A research prospective; Taylor and Francis, London, 1996, 219 page
[17] Communication au 8ème Symposium Locus Sonus, « Audio mobility », 16-18 Avril 2014, Aix en Provence : « In between Marching Band and Sound Walk » (pages 250-262 des proceedings).

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

"The water is rising higher than usual"	"Leave your house immediately, a tidal wave is coming in"
"There is a feeling of hunger"	"There is great pain"

So despite the promise of giving an example for each of these classifications in these detailed writeups, this category is so "fuzzy" that Alerts and Alarm contexts will be pointed out in the other more distinct cases of sonification techniques.

Our research colleague Frank Dufour has added the conceptual basis of our investigation should be grounded in a context of phenomenology.

> *"Phenomenology is the study of structures of consciousness as experienced from the first-person point of view. The central structure of an experience is its intentionality, its being directed toward something, as it is an experience of or about some object."* [18]

Understanding this quote in the context of sonification, a more refined understanding of the fuzziness, the vagueness and those authentic moments of intensified alert attention and genuine alarm comes into focus. Philosopher Stuart Jones posits that "Sonification can offer a form of reconciliation between ontology and phenomenology, and between

[18] Smith, David Woodruff, "Phenomenology", The Stanford Encyclopedia of Philosophy (Summer 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2018/entries/phenomenology/>

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

ourselves and the flux we are part of."[19] Often when speaking about sonification the point becomes clear that we can shut our eyes but not our ears. This points up the important role that audio can play in multi-sensory contexts and in understanding the immediacy of sound relative to sight. Intuitively, this would seem to support the idea that all alerts and alarms in computer human interaction would be audio based, but at scale, this is not the general case.

There can be a threshold that is widely variable between sounds that alert and sounds that are genuinely an alarm and requiring immediate action. This gives rise to the instance that is illustrated by the classic "Boy who cried wolf" parable. If a warning alert is too prevalent and the sound of the alarm itself is based on the same sound design principle, there is a danger that the actual alert will not be interpreted as less significant and just regarded as another alarm. It is important for sound designers to keep in mind this distinction between an alert and an alarm in all the areas of sonification classification. It is an overarching "meta" aspect of this entire area of research.

3.2 Audification

The first and perhaps simplest technique for creating sonifications is the direct mapping of data into sound. An example very real-world version of this auditory perceptual mapping is when a mechanic takes a long screwdriver and puts the blade on the engine block while it

[19] Jones, Stuart "Now? Towards a phenomenology of real time sonification" , AI and Society May 2012, Volume 27, Issue 2, pp 223–231

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

is running and putting their ear to the handle of the screwdriver. They would do this to hear the sound of the valves, bearings and cylinder firings of the running engine.

Many of us have experienced the physician listening to out heartbeat and breath by means of a stethoscope and our research has led us to understand that the training and efficacy and accuracy of the effective use of the stethoscope can be shown to be more efficient in the hands of musically trained physicians. This again points to on one the many vague interactions at the threshold of listening and musical listening. The stethoscope itself has proved to be a very effective use of the auditory channel to enhance.[20]

In the digital and microcomputer age, this means taking a sequence of numbers from a database, sensor or some other digital source and taking that raw data and dumping directly to a digital to audio convertor. To do this in the analog era before digital technology, the voltage off an instrument was put through an amplifier connected to a speaker.

3.2.1 Examples of uses of these techniques

[20] Mangione S, Nieman LZ. Cardiac auscultatory skills of internal medicine and family practice trainees. A comparison of diagnostic proficiency. Journal of the American Medical Association Sep 3, 1997

Electromagnetic to auditory: Example, Valentina Vuksic: "Tripping Through Runtime"[21])

Valentina uses magnetometers to "hear" computer runtime operations by directly converting the output of magnetism from laptops and hard drives.

Transmitted data stream to auditory: Sharath Chandra Ramakrishnan: Traffic Jogja[22] detailed navigational real time data to "hear" radio transmissions from aircraft transponders.

Physical manifestation is the technique that Richard Lehrman uses via piezo electric pickups he converts the mechanical vibration of objects into direct audio signals.[23] Perhaps his most charming piece is the "Travel On Gamelon" where the spokes of bicycles are tuned to the Pelog scale of Javanese music and as the riders bicycle, the transducer 'strums' those spokes and amplifies the sound over a battery powered speaker. This creates a gamelan style percussion orchestra and a fleet of these bicycles go by.

A more digitally oriented demonstration of this technique can be found in the example of Robert Alexander's work helping helio-physicists at NASA's Goddard Space Flight Center in Greenbelt, Maryland, to pick out subtle differences by listening to satellite data instead of looking at it. This article "The bird's ear view of space physics: Audification as a tool for the spectral analysis of time series data" details this work.[24]

---

[21] Viewed July 2019 http://vimeo.com/4105144

[22] Viewed July 2019 https://bit.ly/2jz8cpO

[23] Viewed July 2019 http://www.west.asu.edu/rlerman/

[24] Website seen June 2019 https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2014JA020025

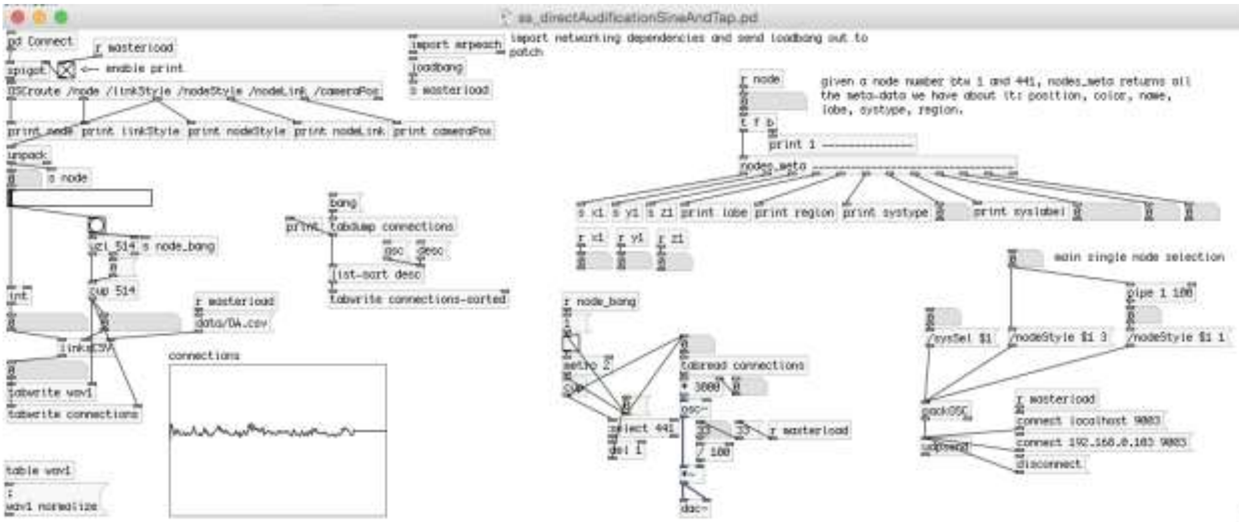
Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

3.2.2 Technical Realizations

In a more modern context, in an age of the "Internet of Things" (IOT), the opportunity presents itself to take the direct feed from the sensors of various online devices and "listen" to them in real time. This would require getting the numbers from the sensor in a buffer or what is referred to as a Round Robin Database (RRD). We have used the linux command rrdtool, for example, to create an ever-changing buffer fed from a live data stream. The output of this ever-modulating buffer can be played back as a digital signal. Often the scaling and speed at which this buffer is played back determines how the raw data is sounds. One trick is to make the buffer exactly 44,1000 with 16-bit values. In that way the ever-changing buffer can be sent directly to a digital to analog convertor at 1 second per scan resulting in a monophonic waveform that is the same audio quality as a Compact Disc.

The figure below is a snapshot of a working program in the computer music language pure data called ss_direct that was designed by Andrew Blanton. The connection graph in the lower left is a graph of the listing of connection strength between neuron groups in an fMRI data set. By loading in different sets of connections the user can listen for differentiations between similar sets of information. Every query to the database changes the information in the table called "connections" and the user immediately hears the change in the data being observed.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA



This is just one example of many realizations that we have created.

3.3 1st order sonification

The next level of techniques for "Listening to Data" is a very general and widely examined area. It consists of abstracting the data flow by only one degree. In other words, taking a data value and scaling it to become a unit in an audio generation algorithm. This can be thought of as a "1st order sonification" meaning that there is a direct correlation between a data value and the audio parameter being changed by that value. In the writings of sonification scholar Thomas Hermann these are referred to as "parameter mapping sonifications". [25] The most widely used version of this concept is taking a value of data and mapping it directly to frequency or more specifically pitch, for example. Frequency is an exact count of waveform cycles per second, whereas pitch is an approximate musical term.

[25] URL viewed June 2019 https://mct-master.github.io/sonification/2019/04/05/Thomas-Hermann.html

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

The frequency of a sound can be in a close range of the pitch and still be perceived a musical note if a little out of tune.

What follows is a detailed description of the processes that a musically trained "sonifier" might begin with. In this case rather than raw conversion of data to frequency, it is converted to musical notes.

Many of the sonifications one can find can be characterized as a type of musical performance as a compositional determinant. The works are often composed using familiar sounds that musical in nature. For example, a simple piano playing notes that are mapped to a time series data flow. This is not as simple as it may seem at first. The initial version of this may be a sine wave following the contour of a line on a graph. As the line increases in value the sine wave goes up in frequency corresponding to the increase value. This simple act of making this sine wave follow a set of time series data values is laden with choices and cognitive cueing that are not at all obvious. Let's say we have a series of floating point numbers that represent a constantly sampled changing data flow. (1.3, 7.24, 6.3, 0.5, 2.78 etc.) To decide to play these values as a sine wave we need to make some decisions that have a marked effect on our listening experience. If we choose to be completely literal then the sine wave would play at frequencies too low for human perception since the lowest auditory frequency is 20 Hz and the highest frequency in this literal translation is 7.24 Hz, inaudible. Also, at what rate do we choose to go from one element in this array to another. Only 5 elements were shown and then etc., so this array could possibly be any size or it could be a live stream from a sensor. So, this would affect the "scan" rate that we would use to go from element to element.

What if the numbers in the series suddenly spike up out of the distribution? (1.3, 7.24, 6.3, 0.5, 2.78, 1287.45, 542.7, 10.3, 6.7, etc.). One needs to know the minimum and maximum range of the data stream in question and then determine the scaling and offset to apply to the whole range so that it is within auditory range. This also requires knowing something about what the user is interested in examining. In one instance, the element at 1287.45 might represent a spike that would punctuate the data set as an interesting indicator of an event of interest and the lower values would be a simple "noise floor" that represents the turbulence in the data flow. In another instance the spike could be an anomaly or a standard and uninteresting blip that is a regular part of this sort of data flow. The point is that the choices given the sonification designer to scale, offset and listen in a data range are all at play. There is more to this, but the final aspect to point out is that the perception of frequency is logarithmic in nature. Our perception of the octave and the natural phenomenon of resonance and harmonics follow a logarithmic curve, just scaling with a fixed multiplier might result is a perceived frequency that does not belay that harmonic relationship that might be present in a data set. This harmonicity is an aspect of natural phenomenon which is the basis of much analysis in the context of Fourier transforms and other techniques in digital signal processing.

It should be pointed out that this is just the beginning of the process of one aspect of this approach of turning a series of numbers into musical notes to be played on conventional

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

sounding instruments or in the case of John Eacott's work transcribed directly to be played by instrumentalists.[26] We have given this approach to data listening the nomenclature of "musification" and with this approach comes a myriad of complexities and problems. These include the cultural and societal differences in "taste and discretion" between any user that might be listening to the result of this process. We cannot assume that each listener has the same musical sophistication, training or background to interpret the results of scaling that will be applied to a given dataset to create this new musical result

Cognitive Science can be used as a means of understanding efficiencies in both Auditory Cognition and Music Cognition. As early as 1999, the role of cognition in the successful scaling of sonification as a practice had been identified. Kramer et al. stated that "Perception and cognition research should focus on dynamic sound perception, auditory scene analysis, multimodal interaction, and the role of memory and attention in extracting information from sound. "[27] A part of the research regarding sonification has been focused on the role of cognition in the act of listening to data as an "embodied" experience. This approach attempts to transcend the cultural and sociological bias of sonification by setting the listening experience in a context. Carla Scaletti makes the strong point the **sonification ≠ music.**[28] She points out "Although data sonification shares techniques and materials with

[26] Eacott, John (2012). Flood Tide: sonification as musical performance—an audience perspective. AI and Society 27 (2):189-195.
[27] Sonification Report: Status Of the Field and Research Agenda
Gregory Kramer-Bruce Walker-Robin Bargar - International Community For Auditory Display - 1999
[28] Carla Scaletti (2018) Sonification ≠ Music , The Oxford Handbook of Algorithmic Music

data-driven music, it is in the interests of the practitioners of both music composition and data sonification to maintain a distinction between the two fields."

Researchers taking that to heart have made the content of any cognitive testing be framed in a setting that requires the listeners to have some introductory training and understanding that what they are listening to is not to be considered as "music" per se, but rather as a metaphorical sonic representation of unseen data. This is the way in which this very direct form of 1 to 1 correlation of data to audio is made more contextualized.[29]

However, music can often interfere with this intent of isolating the experience of sonification from the experience of music. Much research and work has been done in music cognition itself which is confusing in a setting where music is not supposed to be a part of the consideration of data that has been converted into musical sequences and phrases. Therefore, it is important to realize the distinction between pure musical expression and music like results of sonification practices. This distinction is often not readily available to the listener, either.

The combination of both pure auditory and musical cognition has to be taken into consideration with each decision in the conversion of data into some form of auditory

Edited by Roger T. Dean and Alex McLean

[29] Roddy, Stephen & Furlong, Dermot. (2015). Sonification Listening: An Empirical Embodied Approach. Proceedings of The 21st International Conference on Auditory Display (ICAD 2015) July 8–10, 2015, Graz, Austria

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

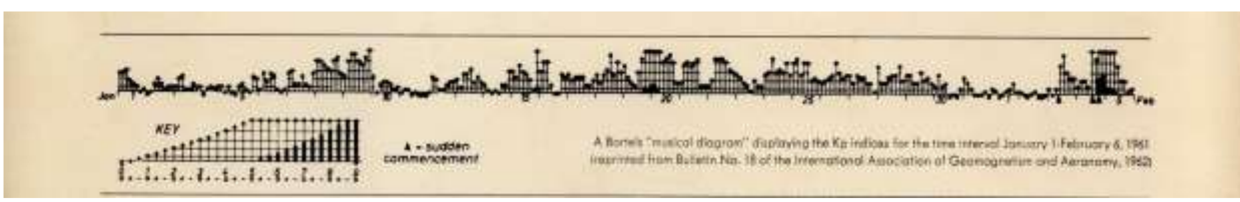
stream. The effect of both these have vast implications for success in designing interaction with user interfaces.

3.3.1 Examples of uses of these techniques

A very early example of this approach to this style of sonification approach is a very early computer piece Charles Dodge - Earth's Magnetic Field, from the Columbia-Princeton Electronic Music Center in 1970. “Dodge mapped magnetic field data to musical sounds. Over the course of a year, 2920 readings were taken of the magnetic field. This was then mapped to a four octave span, or 45 notes (the average span of an instrument). Between different points within this data, interpolations were made to create the other aspects of the music—tempo, dynamics, and register. “[30]

Here’s part of the readings that he used:

Magnetic field data



John Eacott’s tide base piece was mentioned earlier, but when the moon passes between the sun and the Earth on August 21, it didn't make a sound—but Exploratorium composer

[30] URL viewed June 2019 https://info.umkc.edu/specialcollections/archives/1056

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

Wayne Grim and the world-famous Kronos Quartet turned the total solar eclipse into a 3-hour piece of music.

Searching and searching the literature for an instance beyond a music composition characterized as "the sound of the universe" or some equally preposterous claim, there is simply no instance where real practitioners of science are using these tools in a genuine way to advance their research. In an interview with sonification expert Mark Balora he admitted that the strength of this work in his estimation is as a means of promoting interest in an area of science.[31] Alexander Supper does an amazing job of breaking down the interesting paradoxes of sonification in the context of research.[32] She identifies the overall use of this sort of technique of sonification as being like the "wonder shows" of the late 19th century.

However, if there is some interest in investigating the use of this 1st order style of sonification for use with your data. Here is a pointer to one of several websites that does an amazing job of taking your data and turning it into a musical style midi sequence file.[33] "6 basic steps to create music from your data!" is the header. Clearly, they are not thinking along the lines of Carla Scaletti that "sonification ≠ music". So here in this sub-category of a

[31] URL of podcast viewed June 2019 https://creativedisturbance.org/podcast/mark-ballora-sonification-research/

[32] Alexandra Supper (2016) Lobbying for the ear, listening with the whole body: the (anti-)visual culture of sonification, Sound Studies, 2:1, 69-80, DOI: 10.1080/20551940.2016.1214446

[33] URL viewed June 2019 http://musicalgorithms.org/4.1/app/#/howto

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

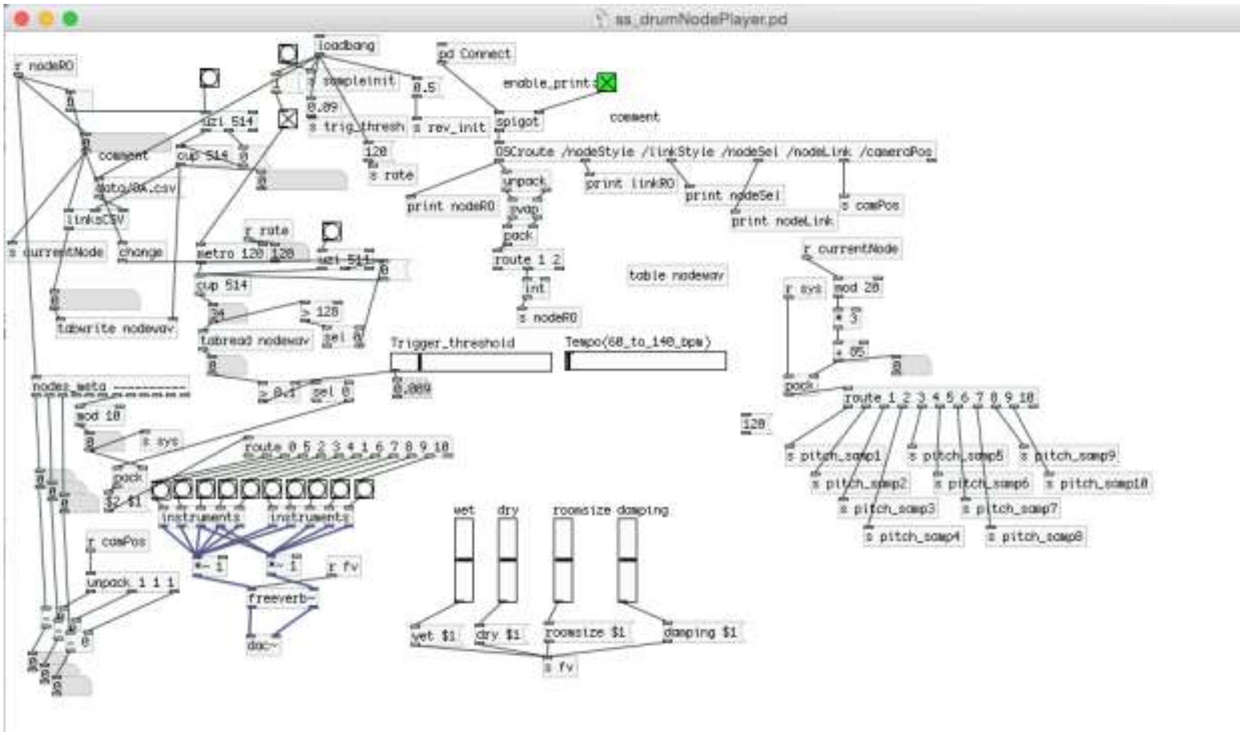
breakdown of sonification into a feasible taxonomy is a clear example of the absolute fuzziness of this area of research.

3.3.2 Technical Realizations

Here we explain a patch developed in our research that will act as a bridge between the more conventional 1st order technique of sonification and a type that is often referred to as Earcons.[34] Earcons are characterized as brief, abstract motifs often with rule-based iterations. In this case a sound that is different for each brain subsystem. Each system in the brain map is assigned a specific “drum” sound. Visual System = Bass Drum, Default System = Tomtom, Auditory System = closed hi hat etc. On node rollover, plays a distinct rhythm for each node. The distinct rhythm is driven by the strength of the connection of each subsequent node, so each node would have a distinct rhythm that contains information about the interconnectivity of the node the user is rolling over. It was called ss_drumNodePlayer.pd in the system for that reason. Here is a snapshot of the patch:

[34] URL viewed June 2019 http://sonification.de/handbook/index.php/chapters/chapter14/

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA



Working with Dr. Gagan Wig of the Center for Longevity Studies as part of a 1-year DARPA funded research on new techniques for examining neuroscience data we developed many modules. Using these modules required becoming familiar with the rhythms the data generated as the user searched the visual model of the brain for nodes in the model that made similar sorts of sounds and drum patterns. Although you could never find an exact match it did enable the neuroscience researchers the opportunity to find area where connection generated similar sets of drum sound and rhythms. The distinction between these similar systems would be nearly impossible to find using visual techniques but with patience and training the distinctions could be detected.

Much research has been done isolating musical cognition of short melodic information that an "earcon" would represent specifically in a dynamic context where subtle changes of

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

pitch classes create marked perceptual differences.[35] The non-linearity of response to even one note changes in a 7+ note sequence make the conceptual adherence to an idea of mapping "iconic" representation to melodic fragments problematic. A small change in a visual curve does not offset the coherence or gestalt perception of the visual object. However, with a short musical phrase, even the smallest change in constituent elements creates large and unpredictable reactions in listeners. The dimension of rhythmic scaling and structure and the subtlety and meaning of even less than 10 notes of musical gesture gains an insurmountable amount of complexity when brought to scale.[36] Paradoxically, when the intervallic order is strictly maintained and the scale of the rhythmic structure remains congruent, musical material shapes may be displaced in scale, both in time and frequency and still retain a recognized shape and/or quality.[37]

This opens the possibility for a more generalized and "semiotic" approach to creating recognizable auditory flags or "icons". An interesting theory of Temporal Semiotic Units (USTs) has been a cornerstone of the research at Laboratoire Musique et Informatique de Marseille.[38] UST's are described as sound objects and defined as semiotic objects. They can be thought of as generalized sonic gestures found in reality and musical composition both.

[35] Bailes, F. (2010). Dynamic melody recognition: Distinctiveness and the role of musical expertise. Memory & Cognition 38(5), 641–650.
[36] Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008) Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. Journal of Cognitive Neuroscience, 20(2):226–239.
[37] Shepard, R. N. (1982). Geometrical approximations to the structure of musical pitch. Psychological Review, 89, 305–333.
[38] Frey, A., Hautbois, X., Bootz, P., & Tijus, C. (2014). An experimental validation of Temporal Semiotic Units and Parameterized Time Motifs. *Musicae Scientiae*, *18*(1), 98–123. https://doi.org/10.1177/1029864913516973

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

As an example: Notes sliding up indicates going upwards, etc. these can be characterized as a proposal since no definitive cognitive studies are currently available. However, it does point to an intriguing approach for the dissemination of “earcons” throughout the user interface systems available. Earcons can be thought of as sonic emojis. A readily available instance type of audio icon is the sound of the “woosh” when Apple Mac users send an email or more fundamentally the sound of a turn signal in the automotive acoustic environment. The question of how to bring new versions of this approach to audio informatics is still not answered, but the proposals and thinking can be traced to the early papers of William Gavers and his proposal of a “Sonic Finder”.[39]

One troubling and unresolved aspect of this potentially bountiful area of research is the assumption of “universality”. In other words, “where” and “how’ does general and the actual segmentation or differentiation between these sound icons happen? This requires much more research and attention. Our interest in this area of inquiry is directly related to this aspect of universality since in each case we are attempting to not make “Art” pieces but make scalable and distributable, easily used types of sonifications. This first order style of sonification lends itself to easy realization by computer music practitioners. If there is a template or schematic for making data reactive sound generators that are modeled on these Temporal Semiotic Units.

[39] William W. Gaver (1989) The SonicFinder: An Interface That Uses Auditory Icons, Human–Computer Interaction, 4:1, 67-94, DOI: 10.1207/s15327051hci0401_3

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

The design of several of the units we constructed were informed of this thinking and followed certain easy to realize and yet structured control of data. Displacement mapping is accomplished by taking raw data values and slightly moving them in both frequency and time to map to familiar musical structures.

“Frequency displacement” this means mapping the raw frequency derived from a data set to the nearest pitch and then moving that pitch to the nearest scale tone of the scale the designer has chosen. If the frequency derived is 277.3567 Hz that is very close to 277.182617 Hz which is a C# in an equal tempered western music scale. However, if the sonification designer is trying to map the results to an equal tempered C major scale 277.3567 Hz would need to be displaced to 261.625580 Hz. The same would be true for a value of 277.2564 Hz etc. It is clear that this reads as very complex, but it demonstrates that even displacement of data requires a depth of musical understanding that results in some tangible recognition even to users without thorough musical training. Making little recognizable motives out shapes in data require this sort of displacement.

Surprisingly, this very complex example of pitch displacement is the absolute simplest version of displacement of data to music in 1st order techniques. Here is the mapping procedure for turning sets of data into “simple music” or “simple sonification” (if you agree with Carla Scaletti) derived from several sources.

1. Pitch Input
    a. Choosing what data will be sequenced for frequency

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

    b. Example: which row of CSV file will be assigned to pitch

2. Pitch Mapping
    a. This is where your data values are scaled to be read as a musical note
    b. Example: 1-88 represents the 88 keys on a piano with a value of 0 representing silence (no pitch).
3. Duration Input
    a. Choosing what data will be sequenced for frequency
    b. Example: which row of CSV file will be assigned to pitch
        i. *This is where each number of your data set is given a number representing a length of time to play. A higher number plays the respective note for a longer amount of time.*
4. Duration Mapping
    a. This is where the duration values get mapped a range of durations, each being associated with the pitch from Pitch Mapping.
        i. *Technically referred to as rhythmic quantization*
5. Scale Options
    a. This is where the Pitch Mapping output is adjusted to fit a designated musical scale.
        i. *Technically referred to as Major, Minor, Modal, Synthetic etc.*
    b. Example: all shapes will be mapped to the “white notes” on the piano
        i. *Technically referred to as pandiatonic*

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

Returning to “Earcons”. If one were to displace short bursts of two streams of a dozen or so data points, one for scaled pitch and one for rhythmic structure using this scheme, a type of earcon for the data stream would be created. The Temporal Semiotics cuts in when the resulting musical gesture can be identified with a class of the defined TSU’s that have been identified. Here is a list of the TSU’s that have been identified as Braking, Compressing, Stretching out, Endless trajectory, Fading away, Falling, Floating, In suspension, Moving forward, Obsessive, Propulsion, Spinning, Stationary, Suspending–questioning, Wanting to start, Waves, Chaotic, Divergent, Heaviness. These poetic descriptors give an idea of identifiers that can be used to construct an understanding of similarities between derived “earcons”.

Of course, this is just one of dozens of strategies that can be taken. Using unquantized, musically chromatic, for example will possibly provide more detail about the phrase in question. This will, however, place the example in a less musically palatable space and possibly affect the ease of adoption.

Another aspect that needs attention is timbre or more scientifically spectral remapping. Are these notes played as piano, sine tones, tuba, heavy metal guitar? Timbre is often used as a derived semiotic element in sonifications. Example: Tide level is indicated by string sounds, wind speed by flute sounds, temperature by piano, etc. The composer/researcher

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

who uses this technique extensively is Martin Quinn, for those with an interest in researching this further.[40]

There is an additional layer of complexity that introducing timbre into the mix creates since data can drive subtle aspects of the timbre itself. Data values can be assigned to aspects of brightness, attack, decay etc. Pointing again to the needs of even the 1st order sonification realization to approach the work being aware of the decades of work in auditory and musical cognition and testing that has been done.

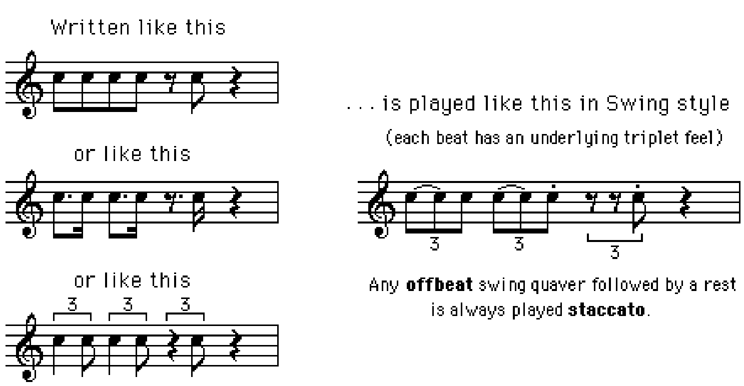
Using scaled data to effect temporal scaling and remapping has been touched on but is a very complex and essentially unexamined space relative to other areas of research into sonification. The research into this area has been associated with more advanced gait compensation for physical therapy and limited comparisons of "melodic" or pitch fragment oriented sonifications and rhythmic sonification.[41] In these cases the metaphor for interaction is deeply tied to the visual with both auditory and visual cueing and often the conclusion is that the visual is more dominant. A conclusion that neuroscience supports. Our research into strictly auditory mapping of parameters to a beat context arrived at some startling results. If you map a bass drum as a time reference on a regular beat and then scale the data from minimum to maximum to hit somewhere in the measure from bass drum hit to bass drum hit, almost no one, including trained musicians, can recreate an

[40] URL viewed July 2019 http://designrhythmics.com/
[41] Dyer, J. F., et al. "Advantages of Melodic over Rhythmic Movement Sonification in Bimanual Motor Skill Learning." Experimental Brain Research, vol. 235, no. 10, 2017, pp. 3129–3140., doi:10.1007/s00221-017-5047-8.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

approximation of the data values played in this way. So, assuming a 4-beat measure with the bass drum hitting on the first beat, if a set of values from 1 to 7 were played a trained musician would need to do weeks of ear training to recite back the values if they were changing every measure. This is with the values absolutely quantized to exact values. The problems are even more complex because of a non-linearity of "feel" that enters with non-quantized values. It ends up that we are extremely sensitive to beats that are just slightly off, so with this sort of rhythmic scheme as you approach the accurate beat, the listener is disturbed by the events that are close but not quite on the beat, almost instinctively. So, the model of linearity that is so familiar in visual representation does not work in the audio temporal domain. The cognitive processing of rhythm is complex and carries influences from the prosody of speech and cultural expectations of rhythmic structure that the listener is accustomed to.

Here is an example of how music is often notated to express this concept of rhythmic "feel" (at least in a jazz context):



This only hints at the nuance that milliseconds of delay bring to the perception of regular rhythmic figures. To compound the complication of this, the nuance of the combination of pitch and duration data driven displacement, and this nuanced rhythmic distribution it results in what is often referred to as melody. This is stepping quite some conceptual distance from Carla Scaletti's differentiation that sonification ≠ music, but again exemplifies how all these differentiations and sonic details combine to continue to "fuzz" the boundaries between any categories in a taxonomy of sonification practices.

Here is a musical example that exemplifies the cultural power of these rhythmic germs may hold for some readers.

Just to give it away with a hint, the lyrics would be "Shave and a Haircut … two bits". The rhythm of this phrase alone communicates its iconic need for completion, especially if two parties are familiar with this trope and one only plays the rhythm of the first measure.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

So, the concept of earcons is useful in this rhythmic context as identifiable repeated rhythmic structures and can be used in a semiotic style of construction that allows for rhythmic signatures to be identified often from raw material. To avoid the disquiet created by near beat events all events that are driven by data values must be quantized to a “tactus” or regular rhythmic dimension.

This technique consists of transferring data values to time delays between events. It points up a huge conceptual problem that is a part of all of these 1st order sonification techniques. This is the idea that simply doing this in the same way that a value is applied to a graph will represent the continuity of the value as a stream, but the process of listening for information is so fundamentally different than the visual equivalent. Listening is a dynamic and recursive process that happens on many time scales. Very short time scales are remembered for rhythmic pulse, slightly longer time frames are remembered for similarities of equal but past time frames (musical measures). This requires repetition to detect differences and are noted as “feel” in both cases. The etymology of this clearly points to the root of subjectivity and therefore represents the quandary of an attempt to glean objective understanding from auditory sources derived from data.

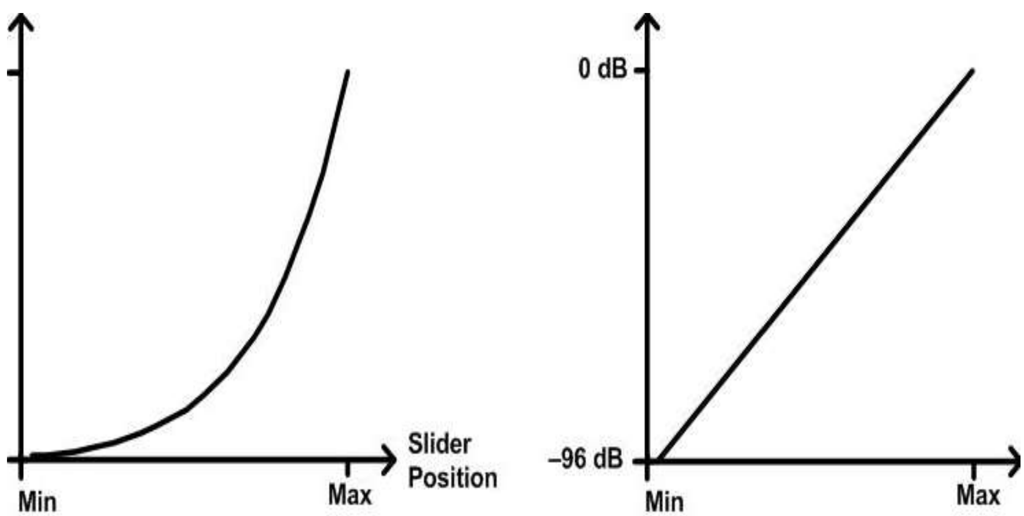
By scaling the parameters down to the bare essentials Sharath Chandra Ram has been able to make a methodology with demonstrable results. The “Pan Pulse Module” of Dynamic fixed pitch pulsing and location or panning to represent a scale. For example, given the range 0 to 100, then 0 would be represented by a sound pulsing very quickly in the left ear of a stereo field, 50 would be in the direct center with no pulsing and 100 would be the

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

same fastest rate of pulses but panned to the right ear only. 25 would be pulsing at half the rate and panned ¼ to the left and 75 would be half the fastest rate and panned to ¾ to the right, etc. We have referred to this as “Panned Pulse” and it is adaptive distance/deviation from a defined center, and the resultant Sonification is a mental illusion. Once the story of the representation is told, listeners are quick to be able to “hear” values.

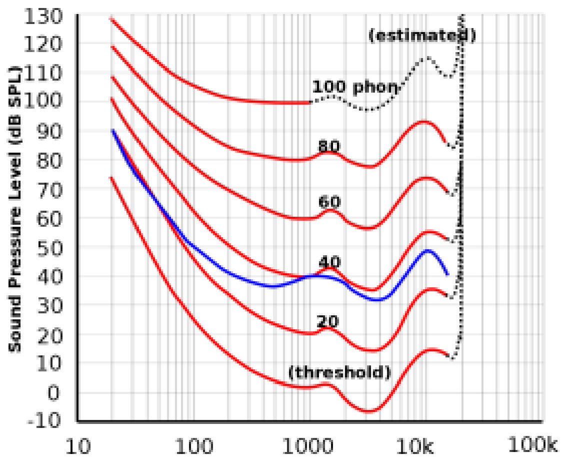
Dynamic scaling of amplitude using first order sonification techniques requires some understanding of amplitude perception. We perceive volume on a logarithmic scale which means when using amplitude of a signal to convey information via data, if using a linear scaling of data, it will result in skewed perceptions. The listener will hear no change of a large field of linear distributed numbers and then the change will be perceived as sudden and drastic at the very end of what should be experienced as an even line.

The figure below gives a visual representation of this phenomenon.



Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

The amplitude perception problem does not stop there. Depending on the frequency content of the audio there is an additional perceptual distortion or "equal loudness curve that purely objective amplitude audio representations that needs to also be designed into an amplitude based sonification.



Freq vs amplitude "equal volume curve" averaged for human perception (Wikimedia image)

So data scaled to represent a linear progression needs to be remapped to an exponential amplitude curve to compensate for logarithmic perception for it to be perceived as linear.

The reader may be beginning to understand the deep complexities that have hampered the obvious transition from a totally visually based set of techniques for abstractly conveying information to a more multimodal approach. The core of this problem is that sound is less objective than vision in terms of perception. (ref needed) The auditory field in humans has evolved as the alert and alarm system, the communication channel and then the part of perception that rubs up on the deep area of pure subjectivity, music, chant and ... speech.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

There is a strange proscription in the accepted definition of sonification. Just look at Wikipedia, at least in 2019, and just a fragment of the first sentence snags it. “Sonification is the use of non-speech audio” Can data scientists making visual charts of data to convey information function without the ability to label those charts? This rather arbitrary and, frankly ridiculous rule, has led our research to a substantive and under researched area of sonification that we have been calling “wordification”.

There is an additional and fully demonstrable aspect of this wordification audio dilemma. The existence of prosody or “the patterns of stress and intonation in a language.” go back to the problems of pitch/frequency and rhythm. This is evident even with single dimensional transfers of data into sound. We are faced with the fact that the perception of raw data in a single dimension transfer to audio invokes the listener to subconsciously attempt to find meaning via language, spoken word, the perception of which is driven by the experience of prosody. So, the ultimate utility of “wordification” in terms of sonification is expanded by this fuzzy connection between music and prosody. It is a startling realization to realize that the designation of “non-verbal” to the accepted constraints of the current definition of sonification may actually be why a true multimodal interface of visual and auditory in user interfaces has been so elusive.

The “smart speaker” has organically alleviated this problem. The “bazaar” not the “cathedral” has provided a pointer to a reasonable expansion of the use of audio spoken word into information sharing.[42] Of course this has been true since the advent of radio, but a more fundamental integration is needed to move this field forward. Emotive descriptions embedded with the narrative that is driving the sound the user is about to experience is essential to the

[42] Raymond, Eric Steven. "The Cathedral and the Bazaar". Retrieved 18 April 2012.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

future of scalable sonifications. Our research has led us to a basic understanding of how uninitiated users of sonification absolutely require an understanding of the narrative of any technique of sonification that is being presented to them. This can be in the form of a visual chart, a video explanation or more simply, a word based, well produced, audio explanation of the mechanism that the sonification they are about to experience is based on. 2nd order sonifications are often referred to as "model based" and require and even more elaborate explanation of the model itself and the connection it has to the data that the model is sounding. Even in this rather direct 1st order context the listener needs to understand the linkage between the data value and the sound that is being produced. The Cognitive Efficiency (CE) metric has been compared and the use of words in a cockpit or driving situation has been fully researched and quantified.[43]

The discussion of "wordification" or the use of spoken words in a sonification framework leads to an examination of the one area of sonification that has been included in previous taxonomies. Spearcons[44] are characterized as brief, accelerated speech. They are a very specific application of Text to Speech technology for the purpose of creating audio menus. " Spearcons are created automatically by converting the text of a menu item (e.g., "Export File") to speech via TTS and then speeding up the resulting audio clip (without changing pitch), even to the point where it is no longer comprehensible as speech."[45] This is just one strategy of many that would use speech as a component in various sonification strategies. An analogy can be made between the text legends on many graphs and charts. Labelling the X and Y axis of a chart gives the user a means of understanding of that chart and without those text labels that chart would be

[43] Yang S., Ferris T.K., Measuring Cognitive efficiency of novel speedometer displays (2016) Proceedings of the Human Factors and Ergonomics Society, , pp. 1934-1938.
[44] Walker, Bruce N.; Lindsay, Jeffrey; Nance, Amanda; Nakano, Yoko; Palladino, Dianne K.; Dingler, Tilman; Jeon, Myounghoon, Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. Human Factors, 2013, Vol.55 (1), p.157-183
[45] Ibid page 160

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

meaningless. In that same way using spoken word labels to identify the meaning of a sonification sound graph would obviously have the same user experience benefit. So, beyond Spearcons we are advocating for the general use of speech based audio labelling and the general use of what we are calling "wordification" of data labels and when appropriate values as part of some sonification strategies.

Combinations or temporal configurations of all these parameter mapping sonifications will spawn aural reaction that can be more clearly understood by an in depth understanding of previous research in music and acoustic perception. Systematic understanding of the range of reactions and expectations. [46] [47] This is a painstaking process that slows development time, but we feel is needed to make sure that our assumptions in design are supported by more than our own empirical experience. Additionally, there is the aspect of types of listening and the variability of results between individuals and their potential hearing deficits or loss.

3.4 2nd order sonification

Introduction

The concepts of 2nd order sonifications are varied but they share a common basis. They are all transferring the data being examined into some larger model that represents either an environment or some sort of sound producing object that requires the input of a user to

[46] Richard Ashley & Renee Timmers (editors), The Routledge Companion to Music Cognition." Printed January 1, 2019

[47] Baldwin, Carryl L. Auditory Cognition and Human Performance: Research and Applications Boca Raton, FL: Taylor & Francis, 2012.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

create the sound. So, unlike the previous section where the data is "played" by algorithms, in these various cases the data perturbs or is interacted with in a model that has nothing to do with the representation of the data itself. In parameterized displacement, that were the basis of the 1st order sonifications, one could think of the playback of those sequences, earcons and time streams as sonic charts, much like visual charts. In 2nd order sonifications a model is created, and the data perturbs the model in specific ways in each case.

Generative Processes

The most ubiquitous type of sonification in this 2nd order class are the processes that use some form of generative audio creation algorithms. Generative audio is generally created in 4 different ways:

Linguistic/structural generative audio

Uses generative grammars, that have rule-based suggestions for the completion of phrases or structures based on other music pieces or sound production systems. David Cope has done the most work in this area with his "Experiments in Music Intelligence."[48] There are rules in this instance, but they are not strict and allow variability. This allows a vast amount of variability based on the rules given.

48

Cope, David (1996). Experiments in Musical Intelligence. Madison, WI: A-R Editions.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

Interactive/behavioral generative audio

This technique is silent without interaction, but then uses human gesture to generate audio and/or music. The "Gestural Models" below are a good set of examples of this sort of approach in our work

Creative/procedural generative audio

In this instance a composer or sound designer uses data to create short Earcon size fragments of audio or music and then an interactive mechanism that is either algorithmic, interactive or a combination of both is put in place to create the structure of combinations of these fragments.

Biological/emergent generative audio

This may be the most often realized version of 2nd order types of sonification. The variations are infinite and are rarely repeated. An audification example from the readers life experience would be the non-repeating notes created by wind blowing through wind chimes.

Gestural Models

In many instances sonification interfaces are made to be "played" or interacted with physically by the user. This can take many forms, but it is a familiar type of HCI (human computer interaction) experience. Users push buttons, choose from menus and/or mouse

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

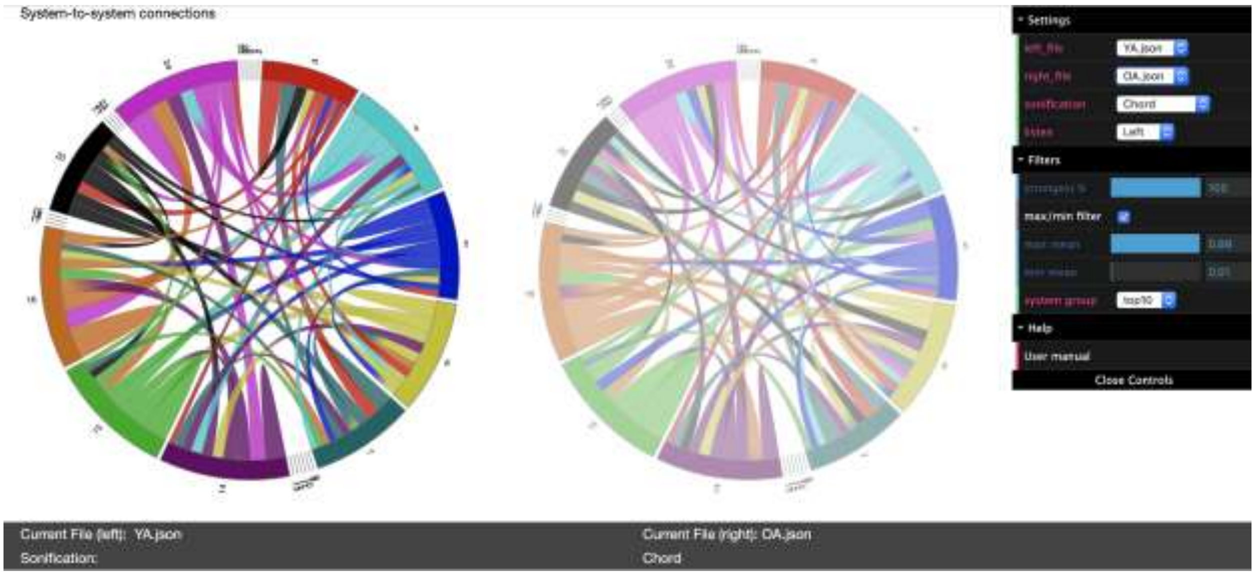
click on specific images on the screen which initiate sounds that are linked to the state of the model the user has just adjusted it to.

Often this takes the form of a direct "audification" like earlier work with "reading" the data directly from gestural interaction. The user is presented with a version of the data that is rendered as a 2d or 3d map and they navigate through the static model to instantiate sounds that are directly correlated to the data that is being examined. Taking the needle on the phonograph metaphor as a one-dimension version of this approach, imagine moving in 2 or 3 dimensions through the data set. The resulting sound would be based on your gestural input in that context.

There is a simpler context that we have realized by mapping data to a specific type of graph and giving the user an opportunity to interact with that graph, to make adjustments to the parameters of the graph and not only to see, but also to hear the results of that interaction.

Here is an example of a user interface of that type that was designed as part of the Data Stethoscope project at the ArtSciLab ATEC UT Dallas. It is called stethoX and is an interactive sonification of standard ‘Circos’ tool.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA



The user makes settings in the menus on the right and when the mouse is rolled over a dividing arc of each Circos diagram a sound is made. What follows is a more complete "user manual" for this particular interface.

User manual

- *To operate the application. Simply hover your mouse over any of the nodes on the circumference. Doing so will trigger the sonification and hide all the connections except for the system on which the mouse is hovered on.*
- *Use 'z' key to switch data source for sonification from left to right. Use 'h' key to show/hide tool palette.*
- *It is recommended to refresh the webpage before using it if a sonification has been running for more than a few minutes.*

Settings

- *Left file: Select the data set for the left diagram. There are 4 data sets to choose from currently.*
- *Right file: Select the data set for the right diagram.*
- *Sonification: Chose the type of sonification. Currently we have 5 types: Chord, DynKlank, Exformation, Écureuil, Melody.*
- *Listen: Select which diagram you want to use for sonification. Sonification will be performed on the set chosen here, irrespective of which diagram you perform a mouse hover on.*
- *Filters: These are you used to filter the data being displayed and sonified. The user can chose to display all the systems, only the top 10 system and the top 17 systems. In addition they can chose to*

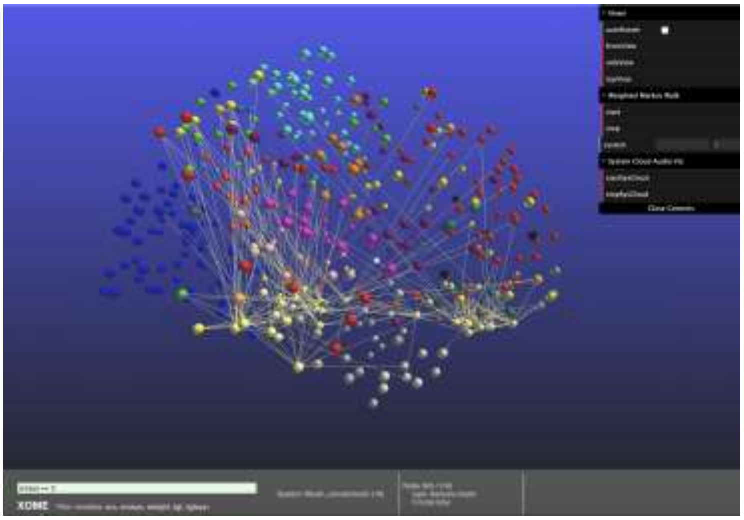
Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

have thresholds for the maximum and minimum mean values being used. It might take some time for these changes to reflect.

This is just one instance, but it is a clear demonstration of the flexibility and nuance of this sort of approach when compared to the brute force of the audio parameter graphs of the 1st order type of sonifications. It is interesting to note that this particular realization is that it gave the user the option to use this model and fixed data set with 5 different sonification algorithms. We found it to be a very effective way to discover diverse sonic perspectives on the data set and is another aspect of the 'vague' taxonomy structure.

The real challenge of this type of interactive gesture reactive interface is communicating the relationship of the interface itself to the model being demonstrated. In some cases, this can be solved by making the visual model in the style of gaming reality. The clearest example of this from our work is the creation of a 3d model of the various neuroscience defined neuron clusters in a model of the brain as a physical layout. Tim Perkis realized this design using the JavaScript library three-js and made such a model pictured below.

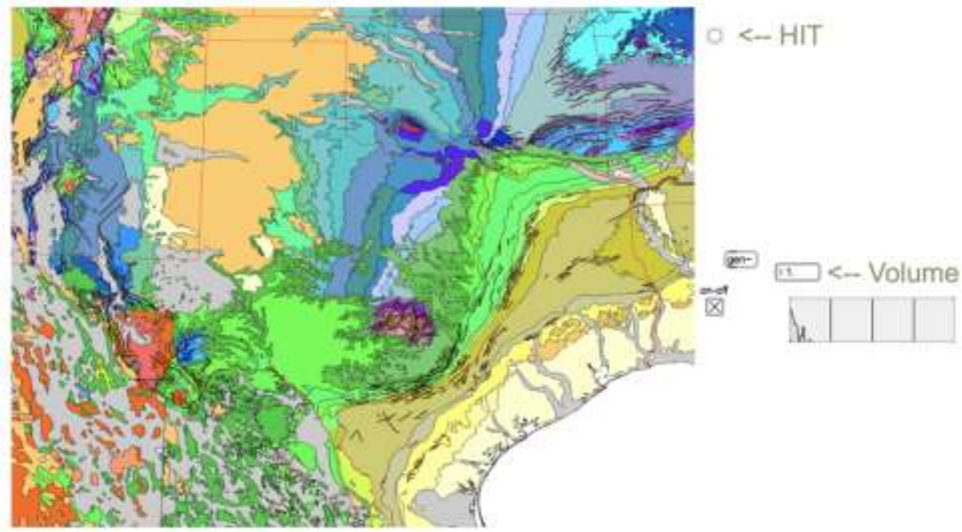
Attribution — NonCommercial — ShareAlike
CC BY-NC-SA



The user becomes oriented to the visual model and then uses the now familiar mouse point and click and/or roll over to trigger sound events. In this instance the sound events are reinforced with accompanying visual reactions that lead the user to point and click at other parts of the model. Additionally, there were automatic scripts which replaced the user interaction with a random walk through the brain systems. This will be discussed in more detail later, but the user was given the opportunity to explore possibilities of combinations and then able to manually interact based on those experiences.

An area of interest that has become part of sonification research is model based sonification. Thomas Hermann has referred to this as Interacting with software instruments physical models of acoustic instruments that are changed by the input of data, so that as the instrument is interacted with the sound of the model changes based on the data. (ref needed) The Texas Gong is a good example of this sort of interaction.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA



Thanks to UT Dallas geology professor Carlos Aiken we built the concept of this project was to create an interactive “gong” from a map of the types of LIDAR geographic materials found on a physiographic 2d map of the state of Texas that he provided to us. The user is encouraged to use the mouse and click on the various regions and excite a “gong” sound with each new mouse click. This was achieved by using a 2d mapping of the techniques of waveguide synthesis to create an interactive physical synthesis model of a dynamically mutable gong. This was a realization of the idea of using the manipulation of waveguide synthesis with data to communicate information via interactive audio.[49]

Environmental Soundscape

We identified a specific new area of sonification technique that we refer to as “environmental soundscape”. It is an interesting and under researched area of sonification

[49] Gresham-Lancaster, Scot “Waveguide Synthesis for Sonification Distributed Sensor Arrays ” (AI & Society ISSN 0951-5666 Vol.24 Number 2 Springer 2011), 289-292

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

and can be thought of as a subset of the larger concept of Acoustic Ecology. Here is a link to a comprehensive resource of materials dealing with this area of research[50]

This algorithmic technique allows us to connect sound and/or music via data to the environment and thereby by using data through sound as it participates in creating an acoustic environment. Scot Gresham-Lancaster and Peter Sinclair wrote a paper that addresses these core concepts in more detail.[51]

What this gives us is an ability to “listen” to complex sets of data in parallel. By creating independent short loops or samples out of sections of a data set and then distributing them in binaural 3d space where the listener gets auditory cues regarding the distance of each separated data set from the equivalent of the camera position in the virtual 3d space. This would include not only the position in three-dimensional space of this virtual camera but metaphorically the position of the user’s head in that space. Looking north along the plane or possibly looking up and to the northwest from that same position. This added complexity adds realism and detail to the user’s experience.

To do this requires complex techniques for distributing the audio. In headphones for virtual reality level authenticity, it is required that head tracking and position compensation is also integrated. As the need for accurate representation of audio in virtual reality becomes part of the production cycle new tools have facilitated the research in this area.

---

[50] Viewed July 2019 https://www.leonardo.info/isast/spec.projects/acousticecologybib.html
[51] Gresham-Lancaster, Scot & Sinclair, Peter “Acoustic Environments and Sonification” Leonardo Music Journal 22, MIT Press 2012

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

FLEX is part of SPAT the IRCAM an audio spatialization software an example of 3d binaural mixing

We have experimented with 2 metaphorical contexts.

The Factory metaphor is an unstable state machine meaning the sound palette evokes the cyclical sound of machines on a factory floor. This approach is ideal for monitoring multiple streams of data in parallel. The user is listening out over an entire field of noise making sonic objects whose sounds are being produced by techniques of audification or one of the generative techniques outlined above. Each has the feature of drastically changing the sound they are producing at thresholds and value ranges that indicate alert action. If these virtual machines are placed on the “factory floor” in a distribution that represents the

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

functionality of the entire data field being examined, then the sound of the factory would be consistent until one of the machines triggered the changing sound of an alert. In a complex network of real time data, the sounds, alerts themselves, as a collective, would create a distinct new acoustic environment that, over time, could be learned and distinguished on a state by state basis. These ideas owe special thanks Tim Perkis via Jeffery Shaw at Interval Research in the late 1990's when we were experimenting with creating tangible audio objects of this nature.

Another type of virtual acoustic environment could be characterized as the "Natural habitat emulation" where the looped sounds of natural phenomenon as semiotic markers with meaning. For example, the sounds of frogs, crickets, wind, rain, surf, babbling, brooks, thunder, wolf howls etc. could be distributed throughout and navigable gaming environment and the amplitude and base frequency of the sample recording playback could be tied to the parameter being associated with that particular nature sound loop. As one moves through the field of cricket sounds some virtual crickets are lower in frequency and softer in volume, etc. Doing this with enough data points would create a distinct sound field that would be representative of an entire multidimensional set of dynamic data points. We have realized this in some of our research and performance work using the 3d model of the physiological neural network sections (vision, auditory, default etc.). In this way a user can "fly through" the model and hear the sound field of an entire neurological grouping of interconnections with the strengths of connection and the boundaries between systems identified with distinct general waves of sound that change as the user moves

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

through the virtual space. The instance of the crickets above was the description of a user experience that was manifested with this system. Interestingly we then stumbled on two modes of utilizing this approach which could be realized in any of the 3d acoustic environments. The first mode was of exploration and user interaction, allowing the user to move about the model freely using 3d navigation to move the point of view and listener orientation relative to the model. The second mode was developed after the model had been explored and was what we began to refer to as the “roller coaster” mode. In this context, the user was on a virtual track but would have the option to “look around” from a moving position that was following a predetermined and informative path. This mode allowed an experienced user the opportunity to help guide a new user to areas of interest with this virtual acoustic environment.

Both of these virtual environment-based techniques of listening to massive amounts of data in a parallel fashion were a logical approach when we started working with many nodes in complex networks. Using the factory metaphor first we made tight algorithmic sound generators as generative machines in this binaural location based setting. It points to an area of research that needs much more attention and critically requires extensive multiuser cognitive tests to really understand where the limits and promise of this approach lay.

Finally, we were led to investigate temporal and gestural change as a generator of sound clusters. These could be useful as a means of becoming familiar with the temporal shape of events and lead to understanding when processes are transforming from a state of

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

running well towards a potential problem. This is still a very vague area of research but holds a lot of promise for the future. Again, the range of listener perception is problematic at just about all levels of this work. The person doing the sonification has their own has their own set of years of experience, taste and hearing ability. This, like a fingerprint, will be different for each individual. Each collaborator/editor will have a similar and different set of hearing.

This unique attribute of the auditory realm led Roger Malina to really question this aspect of generalized listening as an aspect of the core complication of scale in the act of data listening. Combined with the arbitrary music listening skill sets across the full spectrum of listening abled users there is this complication of hearing loss and the varying ability to discern the location of a sound just left and right and even more complexly in three dimensions. To that end we spent a lot of effort to embed consistent and coherent listening tests into each of the modules we have been rolling out. A great find was the http://hearingtest.online site that Stéphane Pigeon graciously created for general use, and with his permission we have integrated into our introductory data listening modules. Of course, any serious listener should consult with a professional audiologist to get an accurate and concise determination of one's hearing abilities. One interesting take away from all the listening test research we did is that all audiology tests only test to 8 kHz which is an interesting aspect of this testing since all the finesse and spatial acuity of fine audio depends on the upper audio octave from 10 kHz to 20 kHz+ As audio professionals it is listening with refinement that drives much of our decision processes and yet audiology

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

professionals have determined that for functionality of day to day tasks this level of refinement is not even worth testing for.

This aspect of the creation of sonification examples points up a recurring theme that trained composers and sound designers do not always immediately intuit. The sense of what is obvious to the refined ear of a musician and/or sound designer with years of experience can be an obstacle to creating data listening methodologies that have a median listenability or obvious comprehensibility. It is hard to emphasize the power and importance of this point.

Exformatic Approaches

Philosopher and Science writer Tor Nørretrander has coined the term Exformation[52] and clarified its meaning with a very simple tale. Arguably the shortest written exchange ever was that between Victor Hugo and his publisher, following the publication of Les Misérables, with Hugo enquiring with a simple "?" about the book's success, and his publisher responding: "!".

This simple example goes to the heart of the concept of exformation. The definition may be a little more understandable in the context of this story. “All the shared body of knowledge, which is not explicitly used when people communicate, but without which

[52] Tor Nørretranders (1998). The User Illusion: Cutting Consciousness Down to Size. Viking. ISBN 0-670-87579-1.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

communication would be impossible." So, going back to the "shave and haircut" example from the previous section, if it is part of your vernacular understanding then if the first part is played "dum didi dum dum" in your mind going something like "dun dun". The completion of the phrase is automatic. There is some crossover as well to the expectations of that you may experience with each of the Temporal Semiotic Units, discussed before. Their completion is embedded in the experience that they create. For example, "Falling" could be a mix of sound that are dropping in frequency continuously. Our experience of reality leads us to get the embedded meaning here.

There is a gross assumption again of universality of the material chosen and there is a boundary layer of clarity between sound objects that communicate directly from the shared human cognitive experience, i.e. falling frequencies have been found to be understood across many cultures. Whereas, something like the "shave and haircut" analogy requires the shared experience of this fragment of the American cultural experience that this small phrase is a part of. It is more complex than that because something as simple as this rhythmic icon (dum didi dum dum …. dun dun) can be learned and recognized after very few repetitions.

Taking this a step further requires a leap of imagination and the confidence that the cultural context of any given song or extended sound sample will be shared with the sonification designer and the intended user of the sonification information to be conveyed.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

It is the culturally biased perturbation of “known” audio sources that is the basis of this technique. For example, familiar music or audio could get varying amounts of reverb with the level of reverb being driven by the data stream. The same audio stream could be spectrally filtered by the curve of the information stream. One instance of this was embedded as one of the choices in the Circos sonification example discussed earlier. In that instance the song “I want to be Sedated” by the punk band the Ramones, was filtered, like a graphic equalizer, by the fMRI data derived from participants brain scans. One could recognize that the song was being played, but as the data being examined changed the “tone” of the familiar song changed, more high frequencies, then more bass and/or mid-range. All easily discernible if one was familiar with the song. Actually, it was even more general than that. If the song was not familiar but what a song should sound like over a loudspeaker was, then the listener could still hear the subtle changes the data flow was having on the sound of the song as it played.

While exformation remains a difficult concept to communicate in words, the experience of it is quite simple. What you are used to is being slightly changed in an unexpected way, the consistency of that perturbation becomes an experience of the data itself. This all requires the same care curation, scaling and adjustments that we discussed with 1st order sonifications and points to the musical craft that is needed to design this sort of interaction mechanism.

The most obvious framework to immediately get a feel for the efficacy of the music part of this “exformation” approach is to create a set of examples based on familiar musical

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

samples or as a sound environment that is perturbed by temporally changing data flows o. There are several types of acoustic processing that can provide the perceptual cues that demonstrate this technique. The higher the value of the data the more reverb or the higher frequencies are cutoff etc., works as a way of cueing the listener that the familiar sound they are listening to is being modified. Like with the 1st order sonification, there is a need for training so that the listener is aware that the semiotic flag of this variability is embedded in the sound signal and to listen for it as representation. For example, the more reverb on the podcast you are listening to the higher the oxygen level.

A variant of this approach is establishing a familiarity by repetition then slightly changing the notes of what has now become a repeated sequence. This set of material, in our tests in a 'minimalist' style like Phill Glass or Steve Reich soon became a familiar new sonic context. Once familiar this "bed" of sound can change in a variety of ways each with specific meanings associated with various data streams. The sufficiently trained listener can discern the variations the data itself is generating. In keeping with the “vagueness” quality of all these techniques, this variation is counting on creating familiarity in interested trained listeners and not necessarily using material that they are already familiar with.

The use of exformation also can extend to the ecological soundscapes that were described earlier as a factory or forest. In the exformatic context these soundfields are “perturbed" by the data flows being examined. This could also extend to prerecorded known musical pieces distributed in synchronized distributed binaural space. The exformatic influence on all the separate parts of these sonic environments of schemes can be varied combinations

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

and will, as before, be endowed with semiotic relationships between data and sound experience that will need to be learned.

The complexity of this approach necessarily goes to the metaphors that underpin the semiotic or even linguistic structures of these more advanced exformatic scenarios. Our research has led us to experimenting in these contexts in a limited way. It was enough to get a sense of the potential narrative strength of this approach relative to the data changed decision tree that results. What has not happened is the rigorous UX and cognitive/perceptual tests that we have used with our simpler modules.

The term metonymy or symbolic representation comes to closest to the connections that a successful version of this type of complex soundfield will allow. In the best scenario, the way that crown represents the king, so these sound environments represent the information buried in the data. This is made vaguer and more complex by the fact that the success of these narratives often relies on a visually integrated reactions users have. The users' experiences an integrated visual and auditory or multimodal scene. These are the most fully multimodal representations we have had the opportunity to work with. In many cases these scenarios are biologically virtual and metaphoric. For example, the neuroscience representations of the initial Data Stethoscope project are complex and requires some supportive dialog and explanation for a novice user to even understand what they are seeing, let alone the abstract sounds associated with those animations. It is only by perturbing those expectations that the information comes forward and hence the

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

exformatic factor. More research is needed to expose if this is a new type of non-verbal semiotics.

At this point it is informative to return to Alerts and Alarms to reiterate that that all the experiences described in the exformatic sonification approach are intended on a natural proclivity to react to sudden integrated auditory and visual stimulus and react with caution (alert) or action (alarm) to it. Additionally, there are thresholds of perception that are on the synesthetic perceptual range. This is not the pathological definition of synesthesia, where it can be thought of as an affliction, but normal integration of visual and auditory reactions. We have calibrated our experiments to just the median or "top of the bell curve" of these reactions. This is an attempt to estimate a setting that will have the broadest impact for the largest footprint of future users.

3.5 Machine Listening

Machine Listening is a field that involves applied computation by the machine to embody biological and cognitive aspects of hearing, to help capture and represent information inaudible or invisible to humans. This technique allows for abstraction of a wide variety of physical phenomena to be interpreted and represented as audio events. It involves the transformation of information between the time and frequency domains, using techniques now widely covered in the applied mathematics and physics of signal processing and Fourier transformations. Within the context of sonification, machine listening is the act of obtaining useful information from the real world, using techniques that allow for the abstraction of a wide variety of physical phenomena to be interpreted and represented as sound.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

As early as 1958 data scientists have been using audio tracking techniques to glean otherwise unavailable information from computer systems.[53] Machine listening fits into the category of Pre-Data Listening mentioned earlier in this taxonomy.

The use of machine listening techniques of sonification dates back to remote transmission of data as sound in radio telescopy and communication. An instance of early image sonification, is the Pattern Playback machine that converted pictures of acoustic patterns of speech in the form of a spectrogram back into sound. 49

A contemporary example of this is the Ceribell - rapid response EEG project[54] with a sonification designed by Chris Chafe of the CCRMA at Stanford. He turned 32 channels of EEG signals into speech-like undulating sound using algorithms that played them all in parallel.

In a recent machine listening application named Abuzz developed at Stanford School of Medicine by Dr Manu Prakash, the characteristic hum sound made by mosquitoes is used by medical researchers to determine if the mosquito is virus carrying or not. [55]

And finally, the most recent success in the use of machine listening based sonification techniques was the conversion of molecular structure of proteins into a musical structure, that upon modification of auditory attributes resulted in the creation of new protein structures that have never been seen before in nature. [56]

---

[53] Viewed July 2019 https://youtu.be/byyQtGb3dvA?t=1014

[54] Viewed July 2019 https://ceribell.com/

[55] Viewed July 2019 https://stanmed.stanford.edu/listening/innovations-helping-harness-sound-acoustics-healing.html

[56] Viewed July 2019 http://news.mit.edu/2019/translating-proteins-music-0626

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

Here is an outline of the current areas in which this machine listening approach to sonification is used, that has been described extensively in the machine listening grey paper that we recently published. [57]

- AI based Machine Listening for Pre-data Sonification:
  - Process monitoring (external sensor and signal listening, Ex: EEG, EMG, bio-feedback, industrial sensors, IoT also known as Internet of Things)
  - Kinematic listening - (Sport kinematics, movement therapy, video sonification)
  - Speech Applications - (Machine listening into Podcast archives, Ai based bibliography sonification and wordification)
  - Electro Magnetic Transmissions - (Satellite Image Sonification, Air Traffic Control, Wireless Networked Data, Radio Telescopy)

All the techniques of sonification above can be reoriented to allow machines the opportunity to “listen” in their unique way and to contribute yet another approach to gleaning understanding and information from raw data sources.

4. Classifications and Techniques

The first section of this paper has been structured around the historical approach to creating a taxonomy of sonification. Reorienting the perspective more on the task of model building and specifically which types of classifications and techniques were used has led to a different more focused approach to constructing prototypes for testing and general distribution. The section applies to the approaches to sonification intended for human users or machines that are listening in a way that is closely mimicking human listening, which is a small set of the larger machine listening paradigm.

[57] Viewed July 2019 https://bit.ly/2TVK3uR

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

4.1 Techniques of Data Listening:

The previous section details a set of approaches that motivated our research but also lead to discoveries of scale and implementation that need to be discussed in a slightly different context and a tangibly more realizable way.

In order to realize the workflow that we are suggesting there is a need to identify a slightly different set of techniques and classifications of sonification. There are many ways to take raw data and convert them to sounds of an almost infinite variety. Categorizing each of the techniques that can be realized with the distinct classification of audio types that can be created using computer music software techniques is essential.

To most effectively do this we have identified three general techniques of interacting with computer generated audio. Technique in this case is the guiding template for the actual algorithm combinations that are used to realize the final sets of audio signals that constitute the sonification output.

4.1.1 Pattern Creation

First is the technique of **pattern creation**. Data is often held in computers in arrays. The rows and columns of a spreadsheet are familiar, where rows of information are organized by distinct columns of repeated information. Name, Sex, Age, Weight, Height for example. These columns can be sorted, filtered, etc. leaving a set of distinct numbers on each row. In the case of the example each array associated with a row would be have four members, something like Bob, male, 37, 190, 5'10". In this case, Bob is not a value, but everything else can be construed as a value. This can be thought of as a 5-member pattern. These discrete number arrays lend themselves to being used scaled and stepped through with all the various computer music

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

software packages. They are all designed to deal with arrays of repeated numbers as parameterized fields of pitch, amplitude, filter etc.[58]

4.1.2 Monitoring

The second technique, **monitoring** (soundscape) is more architectural in nature and relies on the "sonifier" to create a soundscape from the data. Each element in the sonification has an identity and the listener is tasked with listening to the resulting in what is often referred to as a soundscape. There is a rich history of human observation of the world through this approach. This concept may be unfamiliar to some, but it is a big part of the way that we use audio to interact with the world. We are presented with many layers of sound in all sorts of acoustic environments. We are experts at ascertaining the details of each new environment and isolating the identity of the individual sound and placing them in the mental map of the the situation we are listening into This technique is an approach to this very real if seemingly abstract approach to listening to data. All modern computer musics have systems for this sort of mixing of elements into acoustamatic sound fields. All modern mixes can be thought of as being created in this context.

4.1.3 Search and Discovery

The third technique, **search and discovery**, is an extension of the second. Once the audio environment is created, it can be explored. The user/listener can use the techniques of navigation to move throughout the "acoustic ecology" and discover relationships and differences in this virtual space. As mentioned before there is a rich history of the exploration of set and setting purely through the auditory which is often called this and remind the reader of this

[58] Dubus G, Bresin R (2013) A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities. PLoS ONE 8(12): e82491. https://doi.org/10.1371/journal.pone.0082491

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

resource of many books and writings about this area of study. [59] This moves the act of sonification into the realm of gaming, in that it assumes that the acoustic space created will have discoverable artifacts that are the complex combination of navigation and orientation in a new audio “data/information world”. The music technology techniques for this sort of computer music realization are just becoming a part of the sound designers tool kit and are closely associated with the sound production for gaming, virtual and augmented reality etc.[60] These are the newest techniques of audio production, but at the time of this writing they are finally fully a part of the audio industries production capabilities.

These three general techniques can be realized with three separate classifications of computer sound synthesis. Analog conversion, Synthesis and Signal Processing

4.2 Classifications of Data Listening

Classification in this case is very specifically about the synthesis techniques used to create the audio for specific types of sonifications. This is broken down into three rather broad areas. Each of which are very distinct computer music processes that will be familiar to anyone acquainted with creating music using synthesizers and computers. They are:

1. Analog Conversion
2. Synthesis
3. Signal Processing

4.2.1 Analog conversion

[59] Viewed July 2019 https://www.leonardo.info/isast/spec.projects/acousticecologybib.html
[60] Viewed July 2019 https://sonicscoop.com/2018/02/05/audio-mixing-for-vr-the-beginners-guide-to-spatial-audio-3d-sound-and-ambisonics/

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

Analog conversion is the direct conversion of data into sounds. The metaphor of the record needle may help in understanding the fundamental principle of this approach. The data, in this case, can be thought of as the raw contour of the data. By dragging the phonograph needle over these contours, one gets the direct sounding of that groove. The speed and pressure of the needle itself give varying results, but the analogy (or analog) of the data is sonically represented. This is the mainstay of computer music production. The audio buffers, signal generators and sound synthesis routines are designed to take sets of parameters at both audio and control rates to play both discrete and continuous notes and sounds. Each sound event can also be embedded with discrete cognitive attributes (frequency and amplitude modulation, filtering, phase shifting etc.) that are fully integrated into modern computer music languages

4.2.2 Synthesis

Synthesis requires remapping of the data. In this case the numbers can represent direct mappings of audio qualities. One number list can be interpreted as pitch, another amplitude, another as amount of reverb etc. etc. While the most prevalent style of data sonification, this abstract remapping of quantities presents a current key problem with the conveying sonifications. How does the user understand that the pitch represents some notion of a data column? Why is pitch the height of the user and not the age, for example? This problem is what has prompted us to not only include speech markers to help label and communicate to the user/listener. Also, we proceed the user experience with a training module that orients the user/listener to the data remapping first as an abstract concept. Later the user/listener is shown the details of the remapping in the actual instance of the specific use case.

4.2.3 Signal Processing,

Signal Processing is a further level of abstraction in sonification creation. One assumes that material familiar to the user/listener (a song, podcast etc.) is being sent through a digital signal processing box. The parameters of data are applied to the settings of various signal processes.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

This could be filters, reverb, panning of the sound to various speakers, amplitude and/or frequency modulation, etc. The artifacts of signal processing are still audible as a distorted version of the original sound revealing a new interpretation.

4.3 Table of Sonification Techniques and Classifications

Given all these options a wide array of potential combinations is made available for testing and realization. If a matrix is created of the 3 techniques and 3 classifications, a large and distinct catalog of the 9 resulting general sonification types emerge. These distinct and vaguely interrelated sets of approaches constitute the basis of much of all the techniques in sonification that have been articulated previously.

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

3 x 3 Table of Sonification Techniques and Classifications

Techniques x Classifications	Pattern Creation	Monitoring (Soundscape)	Search and Discovery
Analog Conversion (Audification)	Data as gestalt - the discrete representation of the data. Individual data "packets" or icons of sound shaped by the scaled data - raw sound buffers, waves from time series data	Data as environment - Humans and Machines can listen to these environments and detect information	Data as presence - the combination of sound possibilities driven by interaction and navigation - Data Forest 9e2 performance of Reunion 2
Digital Synthesis (Sequencing)	Data as melody or harmonic movement - Musification of data units as notes Circos item 1, 2, 5 for example *Most common technique (bar chart: as melody or chord)*	Data as gesture - a sea of sequences that are experienced as a sound field, motion driven, loops of fixed data or data driven processes mapped on a 2d surface, Texas gong, example Use case: 1	Data as music construction - Fragments of data set are combined horizontally and vertically (eg Tunification or Musification) many examples throughout the ICAD literature
Signal Processing (Exformation)	Data as extrinsic: parameter change - freq., pitch, tone, reverb driven directly by data dimensional vectors**.** (Circos menu 3 & 4)	**Data as location:** familiar or biological signals and audio material distributed as a 3d field and displaced by location and reverb. sound mapping (human and machine listening) Data Stethoscope of the Connectome fly through)	**Data as anomaly:** enigmatic sonic configurations that make the familiar, unfamiliar. Playing music or familiar speech (podcast) through various parameterized digital signal processing (resonate filter, delay, reverb etc.) (Circos menu 3 & 4)

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA

5.Conclusion

At the time of this writing, we feel confident that with these strategies we can begin to realize the vision of the early pioneers of this field who envisioned a time when the human computer interface was not solely based primarily on visual cues. We can foresee a balance of visual and auditory cues that will convey information to the user in faster and more complete ways. We can even further envision purely auditory interfaces that take advantage of the perceptually unique attributes of the "always on" sense of hearing. Harnessing the power of data analytics and machines that are also listening for attributes of audio signals that are rich with subtle new powerful information signals. We have gone to some lengths to make the reader aware of the many vague and complex boundaries of sound and its relationship to sight. This begins to give a clue as to why some 25 years after the definition of this field of sonification there are very few if any significant new interface paradigms have found acceptance at a global scale.

The next steps are to rigorously test and confirm all of these approaches to data listening and work to design modules for specific use case scenarios that require the specific strengths of each one. Our confidence is based on an assurance that our methodology can extend to a global scale. This is the ultimate goal of all this work, to make these "ways of listening for information" ubiquitous and worldwide.

6.Acknowledgments

To be determined after ExPuCu process

Some initial names:

Gagan Wig, Tim Perkis, Andrew Blanton, Cassini Nazir, Shruti Ayloo, Vina Somareddy, Anvit Srivatsav, Michela Chan, Neil Savalia, Paul Fishwick, Mihai Nadin, Frank Dufour, Carlos Aiken,

Attribution — NonCommercial — ShareAlike
CC BY-NC-SA