

## An Interface and Framework Design for Interactive Aesthetic Sonification

Kirsty Beilharz and Sam Ferguson

Faculty of Arts and Social Sciences  
Faculty of Design, Architecture and Building  
University of Technology, Sydney

[Kirsty.Beilharz@uts.edu.au](mailto:Kirsty.Beilharz@uts.edu.au), [Samuel.Ferguson@uts.edu.au](mailto:Samuel.Ferguson@uts.edu.au)

### ABSTRACT

This paper describes the interface design of our AeSon (Aesthetic Sonification) Toolkit motivated by user-centred customisation of the aesthetic representation and scope of the data. The interface design is developed from 3 premises that distinguish our approach from more ubiquitous sonification methodologies. Firstly, we prioritise interaction both from the perspective of changing scale, scope and presentation of the data and the user's ability to reconfigure spatial panning, modality, pitch distribution, critical thresholds and granularity of data examined. The user, for the majority of parameters, determines their own listening experience for real-time data sonification, even to the extent that the interface can be used for live data-driven performance, as well as traditional information analysis and examination. Secondly, we have explored the theories of Tufte, Fry and other visualization and information design experts to find ways in which principles that are successful in the field of information visualization may be translated to the domain of sonification. Thirdly, we prioritise aesthetic variables and controls in the interface, derived from musical practice, aesthetics in information design and responses to experimental user evaluations to inform the design of the sounds and display. In addition to using notions of meter, beat, key or modality and emphasis drawn from music, we draw on our experiments that evaluated the effects of spatial separation in multivariate data presentations.

### 1 INTRODUCTION

AeSon Toolkit is a Max/MSP framework for aesthetic sonification. It includes objects for importing data, for formatting and synchronizing real-time data, for transforming data, for mapping data to sound and musical parameters and for synthesizing sound. AeSon is available for free download<sup>1</sup>, and requires Max/MSP and the freely available IRCAM FTM library<sup>2</sup>. The latest version of the sonification toolkit utilizes Processing<sup>3</sup> for the interactive user interface to enhance the visual appearance of the controls and aesthetics of the interface itself. The following sections explain its operational process particularly in regard to representations data and the rationale for prioritizing interaction, aesthetic quality and aesthetic versatility.

#### 1.1 Why another sonification toolkit?

A number of sonification toolkits and frameworks already exist. In Section 2 we provide a cursory overview of a non-exhaustive selection, for the purpose of identifying key features, goals and

purposes. This brief summary reveals that many operate from an auditory graphing perspective, i.e. primarily intended for representation and analysis of data in a literal, functional way.

Many existing sonification toolkits work from a numerical and statistical background. Our approach is comparatively interdisciplinary, as we thought it would be interesting to look at principles from the fields of information visualisation and information design to examine which organisational and representational ideas may be transferred to sonification (in Section 3) as well as reviewing research directly concerned with sonification aesthetics. We also demonstrate (in Section 4) that some approaches drawn from music and applied in the data mapping process can add useful markers of periodicity, accentuation of significant points and familiar pitch relations to provide aesthetic, melodic pitch contours. We contend that user customisation of the display according to preference, the data characteristics, listening context, adjusted for complexity and to enable different ways of the hearing the same data realised may be key to enhancing the general aesthetic quality and accessibility of data sonification and engagement with the information represented.

Sonification exists for the purpose of conveying data to the listener, but is also commonly used for the creation of aesthetic works that are presented in the same context as purely musical works, or are used in an installation setting. (e.g. interactive performative sonification [1]). These works often do not aim to communicate a dataset, or at least do not use information transfer as a measure of their success. They also usually lack a connection between the user and the data being represented; the users may be uninterested in the data's meaning or purpose as it is of no relevance, or is so complex as to be unintelligible. Their focus is on the aesthetic or musical content of the sonification. AeSon adopts this argument that aesthetic and effective design can enhance and strengthen the data communication and quality of the analytical experience.

Finally, in Section 5 we describe the AeSon Toolkit, in terms of the underlying framework built in Max/MSP and the way in which it facilitates aesthetic sonification outcomes. We also describe the user interface, which is a 'bolt-on' addition to the framework, and is designed to exploit the interactivity and customisability it facilitates.

### 2 SONIFICATION FRAMEWORKS

Some of the earliest frameworks, the Listen, Muse and MUSART frameworks overseen by Lodha [2-4] are research sonification platforms that developed many of the basic patterns for sonification frameworks. 'Listen' mapped data to auditory parameters for synthesis using MIDI – pitch, duration, volume, and pan. 'Muse' extended this model and added many more musical elements as mapping options – it used CSound for synthesis, and its purpose was to produce sonifications of a more musical nature. Finally MUSART extends this idea using

<sup>1</sup> <http://code.google.com/p/aesontoolkit>

<sup>2</sup> <http://ftm.ircam.fr>

<sup>3</sup> <http://processing.org>

‘audio transfer functions’, and adds still more mapping possibilities. They are notable for their early experimentations – and a lot of the ideas they put forward have been incorporated into later frameworks, such as the Sonification Sandbox.

Another early framework was Madhyastha’s *Personify* [5] presented at the first ICAD Meeting in 1992. It used simple GUIs that were configured using ‘widget files’ with ‘sound control files’ to configure a data mapping method.

The Sonification Sandbox<sup>4</sup> [6, 7] is motivated to provide a multi-platform, multi-purpose sonification toolkit. It allows an auditory graph to be produced with the minimum of effort enabling the user to map datasets to auditory dimensions such as, pitch, level and pan with multiple data-streams represented using different timbres. It works with tabular data, and produces both visual graphs and auditory graphs, in a similar way to Excel. Vickers has discussed the Sonification Sandbox, and has pointed out some of the problems inherent in the reliance upon the MIDI protocol [8], and indeed Davison and Walker do as well [6] – however they argue that the sandbox is not meant to produce every possible sonification design, but rather is positioned as a simple general-purpose tool to build auditory graphs. It remains a well-maintained piece of software, and many of the options for sonification it provides have been supported by the author’s research (such as [9]).

SonEnvir<sup>5</sup> is built on the SuperCollider and PureData synthesizer/programming environments, and specifically targets the demands of fields with complex, multi-dimensional data for analysis. One of its drawbacks is the significant amount of knowledge required to use its implementation in SuperCollider. It is quite technically ambitious, and also seems to incorporate spatialisation methods for the IEM Cube spatial playback system. To modify the template requires significant programming understanding, and it seems that such sonification is often undertaken in a collaborative research context, such as in the ‘Science by Ear’ Workshop [10]. It benefits from a wide user-base, and is interesting in the diversity of data that it has been applied to [11].

Worrall et al’s *SoniPy*<sup>6</sup>, as the name suggests, uses the Python programming language as its foundation [12, 13]. He characterises the development purpose as a need to balance the data processing and sonification capabilities within one piece of software. His premise is that others will embrace the development of new modules and add-ins using this modular programming platform. However, he acknowledges the quality of the modules available can vary significantly, and the installation of each of these tools can increase the complexity of the process.

*Monalisa*<sup>7</sup> by Nagano and Jo [14], is a sonification toolkit that uses image filtering methods into data sonification of image and data files. While a serious sonification tool, its purpose might also be construed as exploratory and investigative or curious and fun, more than for information analysis. Generally, it operates by converting image data into audio data and applying image or audio filters resulting in audio output that is fairly unpredictable – the reverse, audio to image, is also possible. While its outcomes are quite diverse and feature a range of aesthetic results, its intention is not to provide a tool for analysing and rendering abstract datasets – it is specifically interested in the crossover between sound and image. SonART<sup>8</sup> is a similar image-based framework by Yeo et al. [15]. The

OSC network protocol facilitates a variety of real-time synthesis and distributed synthesis options; currently implemented in a Macintosh application. Their stated aims focus on network transmission of data, real-time sound generation, distributed synthesis and modular design. In recent incarnations the emphasis seems to have shifted towards image-based interfaces, using image data and layering effects to generate data to control sonification. The control data is taken from the layered images or data linked to image pixels or areas, and the framework generally outputs OSC data, leaving audio creation to the user’s choice of OSC client.

Pauletto and Hunt’s Interactive Sonification Toolkit is an environment in which datasets are scaled, sonified in a variety of ways and navigated with the mouse (with future plans for other interface controllers) [16]. Their toolkit principally deals with non-real-time datasets of pre-gathered information (although an on-the-fly theory is posited) and the aim lies in applying various sonification algorithmic processes to the same dataset to produce various auditory outcomes. It is cross-platform and uses Pure Data (PD) as its programming foundation. They propose a number of different ways of listening to sounds, ranging from distinct data dimensions mapped onto distinct auditory dimensions, to generating a complex timbral effect whose single-sound complexity reflects data structure.

Barras discusses his framework *Personify* in the context of a number of sonifications built using the TaDa design template, which was the topic of his Ph.D thesis [17] *Auditory Information Design*. The tool is divided into two parts. The requirements phase sets up the data according to the TaDa method, as well as configuring a default representation. Histograms are computed and mapping extremes can be configured, as well as data types. The representation phase has a number of methods for representation and a number of output audio devices.

Dombois’ *SoniFyer*<sup>9</sup> [18] is one of the most recent sonification frameworks presented. Two data domains are discussed, that of EEG data, and of seismic data. The software provides parameter mapping sonification using FM synthesis, as well as audification. Visual maps are supplied to assist the user in selecting sonification channels – in one interface a user may select an EEG signal from a map of the head, and in another they may select a seismic signal from a map of the world. Dombois points out the importance of expert knowledge in the design and interpretation of the sonifications, and so the software includes a type of “virtual forum” window for research group members to discuss the data and its interpretation. Finally there is a strong emphasis on community discussion and distribution of the sonifications produced through an accompanying website.

## 2.1 Framework Design Consideration Summary

The discussed frameworks can be summarized using several design attributes. In Table 1 we have briefly outlined each framework’s specific level of applicability to different data types, the user interface method, and audio rendering system.

A number of different approaches to the design of the interface for sonification become apparent through the process of surveying these frameworks. At one extreme there are the code-based frameworks that maintain a huge degree of flexibility and customizability, but require significant programming knowledge to be used adequately. An experienced practitioner can not only produce a highly original and engaging

<sup>4</sup> <http://sonify.psych.gatech.edu/research/auditorygraphs>

<sup>5</sup> <http://sonenvir.at/>

<sup>6</sup> <http://www.avatar.com.au/sonipy/>

<sup>7</sup> <http://www.monalisa-au.org>

<sup>8</sup> <http://www-ccrma.stanford.edu/~woony/software/sonart/>

<sup>9</sup> <http://www.sonifyer.org>

sonification using these types of system, but can also produce an interactive interface to the sonification, allowing real-time exploration of the dataset. The tradeoff is the small audience that is able to use these systems effectively, and the large amount of time necessary to learn their use. At the other extreme there are the graphically driven, easy to use systems, which allow simple auditory graphs to be produced quickly from straightforward data sets. Another approach was a multi-layered interface design – the software contains a number of well-organised classes for sonification, but also allows an easy to use GUI to be ‘bolted’ on, hopefully providing flexibility and ease of use. Finally, *SoniFyer* shows an example of software that, while still using a traditional GUI, provides a great quantity of options and flexibility, in a similar way to modern graphics and video authoring programs.

The audio rendering methods are also quite diverse – some of the systems simply output MIDI or OSC, and do not attempt to control the synthesis directly, but parametrically. Many newer systems ‘hook into’ synthesis methods provided in their associated environment (eg. SuperCollider, Python, PureData).

Each of the frameworks had a different approach to data import methods, but many opted for a form of text file input. Often the data import methods provided by the software or programming environment used are employed, meaning that data is best cleaned and prepared in another program before being presented to the sonification system. As an aside, it is possible sonification frameworks may benefit from a standardized data format – they all seem to need to grapple with data types, cleaning, filtering or transposition methods.

Table 1: Several attributes of each the discussed frameworks. These attributes are not investigated exhaustively, and are based on reading the papers and documentation supplied for each system.

Framework	Data	Rendering Method	Interface
<i>Listen, MUSE, MUSART</i>	Primarily Tabular Data	MIDI, CSound	GUI
<i>Porsonify</i>	Text Data File	Custom Libraries	Audio GUI and Config. Files
<i>Sonification Sandbox</i>	Primarily Tabular Data	MIDI	GUI (Java)
<i>SonEnvir</i>	Real-time Data and	Supercollider Synthesis	Code and Custom SC GUIs
<i>SoniPy</i>	Python data import	Many synthesis methods	Python Code, Custom Python GUI
<i>Monalisa</i>	Image Data, Audio Data	Core Audio Signal Processing	GUI (Cocoa), Audio or Image Unit Interfaces
<i>SonART</i>	Images or Image Associated Data	Sends OSC data to OSC client, audio rendering by user.	Image Layering GUI created (Cocoa)
<i>Personify</i>	Tabular Data	CSound	GUI
<i>Interactive Sonification Toolkit</i>	Tabular Data	PureData Synthesis and Audification	PureData GUI
<i>SoniFyer</i>	Text, EEG, Seismology	FM Synthesis, Audification	GUI (Cocoa), Image Based Navigation

## 2.2 AeSon Toolkit Design

The strategies taken in the AeSon Toolkit have a lot in common with the above surveyed frameworks. The interface to the toolkit attempts to allow both flexibility and simplicity by using a popular and intuitive audio programming language (Max/MSP), and providing modular sonification routines (in the categories of Data-handling, Mapping, Signal Manipulation, Synthesis and Sampling) to simplify the process of building new sonification methods. Concurrently, for those that still do not wish to engage with digital audio programming, there is a graphical user interface that allows interactive configuration and playback of data sonification – hopefully engaging both user groups.

Audio synthesis fundamentals, as with other frameworks, are provided by the overarching environment (Max/MSP), and there are packaged routines designed specifically for simplifying polyphonic FM synthesis or time corrected sample playback. There remain, however, ample possibilities for the user to replace these routines with their own version.

Data import and storage methods are provided by the FTM extensions to Max/MSP [19], which are a significant improvement on the rudimentary data structures provided by Max/MSP. Not only does FTM provide matrices and dictionary data types, but it also allows simple smoothing, interpolation, quantization, recording and playback and statistical functions.

## 3 INFORMATION VISUALIZATION AND INTERACTION DESIGN PRINCIPLES TRANSFERABLE TO SONIFICATION

### 3.1 Characteristics of Representation

Representations of data are tools for investigating and ultimately for understanding data resulting from measurements of a particular phenomenon. Representations can be used to explore data, when the user is unsure of the meaning of a data collection, or they can be used to confirm a hypothesis, when a user seeks verification of something they already know to be true. Sometimes these two purposes are described as exploratory data analysis and confirmatory data analysis [20]. Graphs (either visual or auditory) are often more useful for the exploratory phase, as they can highlight outliers, describe multiple trends, or draw similarities between variables intuitively. Once effects within the data are *explored*, statistical tests can be used to *confirm* their significance.

Representations are often used in various ways – Walker and Nees [21] describe the various possibilities using terms such as *trend analysis*, *point estimation*, *pattern detection*, and *point comparison*. They then point out that it is simply impossible to use many auditory graphs to achieve these generic tasks. There are a number of distinct data elements that a representation may attempt to convey to a user. Measures of central tendency or location (e.g. various types of means, the median, or the mode) describe the general location of the dataset, while measures of dispersion (e.g. standard deviation, variance) can describe its spread.

### 3.2 Information Visualization Approaches

Benjamin Fry's simplification of complex data visualization principles is manifest in his methodology (and tool) for *Computational Information Design* [22, 23] that converges the fields of information visualization, data mining and graphic design normally employed in the complex task of representing

data visually. One of the first challenges presented beyond merely mapping data to representation is how to handle large volumes of data and how to glean a "big-picture" oversight or meaningful understanding. Fry's description of the process flow is not domain-specific and applies equally well to auditory display as visualization: *acquire - parse - filter - mine - represent - refine - interact* (p.13) [22], a process he explains has reflexive and iterative rather than linear stages. This iterative process is possible with AeSon Toolkit. However it is the choice of representation method and forms for conveying data that govern psychological perception and where aesthetics can vastly influence the efficacy of the design, as well as its pleasingness. Approaches that can rapidly, if qualitatively, convey the overview meaning and significant gist of the information representation are much more ergonomic than methods that require detailed analysis and short-term memory recall, e.g. representing quantities with number values to be read vs. with size-scaled visual dots. This view is underscored by Maeda's *The Laws of Simplicity* [24], the goal is non-active information assimilation – a pre-attentive display. Examples of how quantity and relativity information can be represented visually include object's number, shape, alignment (orientation) relative to the group, proximity, shape, size/scale, likeness vs. dissimilarity of appearance. Also colour, motion and spatial position can impart pre-attentive signification.

Comparing these form modifiers with sonic representation, the broad division of similarity and dissimilarity within any representational parameter is important. For example, like values can share similar timbre, pitch, duration, intensity (loudness/brightness) while outliers and dissimilar values exhibit contrasting characteristics. This emphasises the need for a subtle enough display that allows for variety or diversity within representational parameters and sufficient dimensions to convey the necessary dimensions of the dataset (especially timbre). Vande Moere [25] and others have argued that the choice of representational form is a juncture in the design process where the mode of representation can be aligned with the aesthetics and semantics of the data, e.g. colours and icons in graphics or timbre and character in sound can symbolise the tonality, gravity and message of the data. Information about fatalities in wars should be differently represented to the number of products sold, for instance. This a stage in the design process where aesthetic and semantic input can be imposed on the rendering.

Tufte [26] advocates minimising non-data ink, i.e. reducing superfluous and distracting elements in the representation that are not integral to clear understanding. Musically-speaking this is equivalent to minimising irrelevant auditory information such as harmonisation, beats and pulses that do not serve a useful purpose. Tufte's texts [26, 27] are his clear articulation of the importance of visual design for understanding data, noting that "graphical excellence consists of complex ideas communicated with clarity, precision, and efficiency," [26] a divergence from the expectation that somehow visual design serves to make data pretty or entertaining. This points to a potential conflict between aesthetic and informative representation – or is it simply a conflict between unnecessary complexity/complication and clarity. Lau and Vande Moere [28] discuss this possible conflict in a model of information aesthetics, characterised by 'the relationship between what a visualization facilitates and the means by which it achieves' the outcome. Childs [29] has investigated Tufte's theories with regard to sonification, and has developed a number of possible guidelines. For the AeSon Toolkit, we have adopted the second reading – namely that efficiency, simplicity, intuitive signification through pre-

attentive formal representation, and careful data filtering can each contribute to a less complicated rendering.

We still face a situation in which a complex information representation ostensibly may map abundant, abstract and unfamiliar information, and hence the next level of interactive sonification looks to ways in which we can assist the person analysing and interacting with the information through our representational interface. In information visualisation, Fry discusses making key information obvious in his example using dense genetic datasets, highlighting key areas, and applying rules for deciphering akin to pattern- and rule-recognition, e.g. adjacency and recurrence are useful features. He likens understanding new data to hearing a foreign language. Rare events are statistical anomalies or outliers or perhaps significant events, hence colour (in Fry's case) or a second layer of signification can be used to foreground and differentiate data characteristics detected in computational analysis. In our toolkit, we are able to do this by using timbral adjustment and other methods to foreground critical data-moments in relation to thresholds and constraints set by the user. This is an important aspect of the customisable interface and interactivity, allowing the user to define importance and re-examine the dataset according to different measures.

Fry {Fry, 2004 #539} (p.62) identifies clustering similar data, use of colour (or any differentiating feature) to improve clarity and mouse (or other interface interaction) to highlight location. Teasing this out and applied to a sonic context, colour could be considered literally: spectrum, timbral profile, tone colour or the general gist of illuminating relevance in a distinctive way, e.g. with a high-pass filter, amplifier or modulator of some kind. Highlighting location enhances the interface design, acknowledging the user interaction, navigation and locating or orienting the user. As we have designed our AeSon Toolkit for a variety of interface refinements, mostly intended for physical interaction (touchpad, drawing with a stylus on a graphics tablet, using multi-touch or fingers on a tangible surface or Apple iPhone-style screen) orienting the user and the ability to locate sonic objects in 2 dimensional space are critical features in the spatiality of our interface. With 3-dimensional controllers such as a Wii controller, we enable the user to move sonic objects in the three spatial dimensions and use the rotation, tilt, yaw sensors for motion control, scaling, scrolling and panning the data. This gestural interaction and the relationship between intuitive spatial manipulation of information is further intended to give the user pre-attentive and intuitive control of the interaction relying on proprioception and naturally familiar spatial motion to interact with the data.

We also see that low latency for interactive sonification and the ability to graph or map the same datasets in a variety of ways in order to interrogate it with a fresh "view/audit" are necessary. In our AeSon Toolkit, the user can scale the data at different granularities and refine the sonic dimensions applied to the axes and adjust tempo or pace of time-series sets, performing multi-dimensional scaling transformations. Contrast and differentiation can be used to transfigure the aesthetic incidence of the data as well as to clarify meaning. Hierarchy is contextual according to the dataset so our toolkit allows the user to spatially drag mappings in space, positioning the representation along the axes and adjusting the level of different datastreams.

In Fry's information visualization [22], size, weight and placement are criteria used for interpretation. Musically, there are a number of ways to represent the idea of size (i.e. contextual scale and importance), such as density of a cluster of sounds - such as Edgar Varèse's "masses of sounds" or Iannis Xenakis' "clouds"/statistical clusters of values whose density,

mass and distribution affect the perceived size and opacity of sound; dimension of the sound; intensity (loudness, spectrum); spatial proximity - perceived nearness appears larger, hence vary reverberation and virtual spatial audio characteristics to give sensations of nearness and remoteness and movement; dynamicism / agitation / animation or excitement effectively attracting attention; articulation.

In the spatial dimension, Fry proposes that spatial rules, connected elements and proportion rules govern the representation [22]. In the AeSon interface we employ a spatial, physical interface, and use natural acoustic rules of drawing acoustic sources closer, to the foreground, panning sounds to the left and right of the display-space. In auditory representations, connectedness can be metaphorically likened by temporal proximity or timbral similarity (proximity of palette) or connectedness in pitch or modality. Again, this is a grouping activity, where clustering refers to temporal convergence, connectedness refers to similarity but can be dispersed over time. Proportion rules require strict scaling and multiplying functions. Scale and zoom in a time-series dataset effectively describes the time-stretching or distribution of the data over time and ability to mine down to finer granularity, "magnifying" features of form. We address this using the spatial sliding axis principle that simply transforms how much of the sample dataset fills the inspection "window" (in auditory terms, what section of data is heard in the duration).

John Maeda's "Laws" or design principles seem to apply to information sonification in the interest of achieving efficacy and clarity [24]: thoughtful reduction to achieve simplicity; organisation which makes a system of many appear fewer; avoid time-wastage; contextualise and utilise peripheral information to contribute to the context of the central thread; subtract the obvious [which is implicit] and elucidate the meaningful.

In *Envisioning Information* [27], Tufte talks about escaping flatland, i.e. for two-dimensional graphic design this means finding representational techniques that interleave layers of information. That is, within a succinct "gestalt" view or time-frame for assimilation, overlay various concurrent meanings and seek an information-rich display, to "sharpen the information resolution ... to increase (1) the number of dimensions that can be represented ... and (2) data density" (p.13). In sonification terms, these goals of multi-layering concurrent information streams and condensing the amount of information that can be accommodated within a space equates to maximising information reviewed in the time-scale. Multi-dimensionality also adds diversity and aesthetic interest, hence it is both quantitatively and qualitatively enriching.

On layering and separation, Tufte states [27]: "confusion and clutter are failures of design - not attributes of information", hence organisation can be achieved using (1) proper relationships among information layers, and (2) representation relevant in "proportion and harmony" (p.54) which harks back to aesthetic goals found in music. One way in which separation and delineation can be achieved is by foregrounding and emphasising important information, akin to differentiating the graph from its background in graphic design or eradicating competing patterns, pitches, interfering elements, confounding density of information. In musical and sonification contexts, we can use distinguishing pitch register, timbres and filtering to distinctively characterise competing elements, as well as temporal off-set (rhythmic separation, asynchronicity).

#### 4 AESTHETIC APPROACHES FROM MUSIC AND VISUALIZATION

Visual quality, according to Fry {Fry, 2004 #539}, is often omitted from the discussion of information visualization due to its immeasurable, un-quantifiable and subjective characteristic, yet this is precisely the crux of our discussion about the importance of aesthetics in contributing to the uptake, acceptability and ergonomics of listening to information sonification. Fry argues the importance of representational quality precisely because small aesthetic design decisions become dramatically amplified when applied to large and complex datasets (p.40) [22]. Tukey {Tukey, 1977 #523} implies also that smallest inaccuracies and distortions in data can influence the quality of the graph, hence he places greatest importance on "procedures for analyzing data, techniques for interpreting the results of such procedures, ways of planning the gathering of data to make analysis easier, more precise or more accurate". In addition to Grinstead's "data select - data manipulate - representation - image operations - visual interactions" [30], Fry [22] (p.45) and Tufte [26] posit that there needs to be further visual refinement, the refinement of aspects of the image for clearest communication. This final step, while intended to refine and clarify information representation, also allows scope for aesthetic refinement and fine-tuning that can, for example, allow user and designer intervention in the selection and grading of timbre. Hence timbral control is an area of refinement, scaling and fine-tuning we ascribe in our AeSon toolkit.

Paul Vickers posits that, "many sonifications have suffered from poor acoustic ecology which makes listening more difficult, thereby resulting in poorer data extraction and inference on the part of the listener" [31]. Composers using algorithmic processes have also explored sound at a more microcosmic or microscopic level (creating sonic materials, synthesising musical materials) not only organising at the level of the note object but creating the spectrum and characteristics of individual sounds, organising using metric and geometric methods as well as conventionally understood ones. This connects with one of Tufte's organisational principles: the reading of the micro and macro. We hear with Xenakis' use of the Serial Modulor approach [32] or the French Spectral School, a musical architecture that relates the structure of sonic grains and sub-note-sized organisms to grosser temporal durations and overall forms of pieces, spatial distribution and textural organisation. Tufte's [27] theory for information design is about containing detailed information that can be zoomed in on, or mined at a deeper level, but it is also about (beyond comprehensiveness) organisational structures: "detail cumulates into larger coherent structures" (p.37) and "simplicity of reading derives from the context of detailed and complex information *properly organised*. A most unconventional design strategy is revealed: to clarify, add detail." In sonification, we can translate this idea as embedding zoom-able detail or nesting (unpack-able related hierarchies of information) in the detailed qualities of notes.

As Vickers [33] reminds us: '*Edgard Varèse defined music as organised sound, and sonifications organise sound to reflect mimetically the thing being sonified. so many different sonification examples have been built. What is apparent from listening to them is the wide variation in the [musical/aesthetic] quality.*'

#### 4.1 Musically Aesthetic Time Domain Representations

Sonification places the representation in the time domain, and therefore it must grapple with the limitations of perception in the time domain. Visual perception and auditory perception are different due to this ephemerality and sound's dependency on memory. The texture of the auditory representation is generally determined by the number and length of notes presented. Sonification generally uses a specified number of data-points represented as sound events (usually notes) presented at a constant rate or tempo, although there are many exceptions. This constant rate also often determines the length of notes – they are designed to sound only until the next note starts. This determines that none of the notes will sound simultaneously, and also makes sure there will be no silence. Therefore, often in traditional sonifications single (or sometimes multiple) notes are presented at a consistent rate throughout the sonification. This has the effect of avoiding any dissonance between notes due to temporal overlap, meaning that an unintended effect of the choice of pitch mapping is avoided in all cases. An alternative method that provides more control over the textural complexity of the sonification is to choose a pitch mapping where all or most of the possible note combinations are consonant.

Contingency plays another important role in time domain representations. The intervallic leaps made to reach each note affects the perception of that note. Also, depending on the effect of rhythmic accents, musically notes usually have different meanings in different parts of the bar. Temporal fusing is also a very important time domain effect. If two note's onsets are simultaneous (eg. within approximately 30 ms) it is likely that rather than two notes being perceived separately, a chord will be perceived, especially if the notes are harmonically consonant. The chord's components are difficult to separate, and the character of the chord is usually different to either component presented separately.

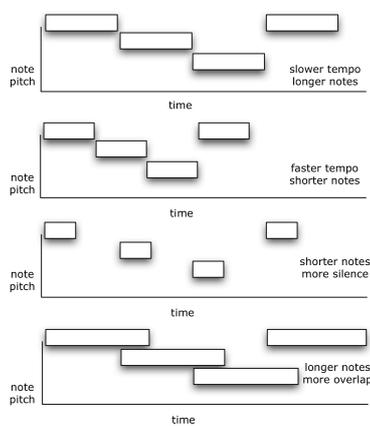


Figure 1: Temporally, the speed of presentation and length of the notes determine the number of notes sounding simultaneously, and the resulting harmonies.

#### 4.2 Pitch and Harmony

Pitch is the most commonly manipulated attribute of sound for data representation. It is also the most complicated attribute to control, in terms of fusions, dissonance and their implied meanings. Shepard has described interrelations between pitched sounds using a geometrical spiral shape [34]. The most obvious connection between notes is at the octave, where fusion typically occurs as all the harmonics of the two notes. Rather than attempt to directly control or correct for these factors we

seek to work around them through careful identification of purposes of representations and the application of appropriate scales and ranges. Furthermore, musical aesthetics typically requires the use of musical scales and harmonic movement through chordal states. Musical scales or sets of chordal notes are sometimes used in sonification, but the mapping is usually direct, and therefore it is assumed that the mapping (the frame of reference) cannot be changed during the sonification. By contrast, in musical situations, the lack of chordal diversity typically results in predictable, uninteresting music. Given a small set of chordal choices (perhaps 3 chords), and a selection of notes from those chords that fall within a 1-2 octave range, the general contour of the data can be represented adequately, while the sonification can be invested with chordal movement and direction. The use of a scale also provides a frame of reference implicitly, as each note in the scale has a unique relationship to each other. In a chromatic scale (and also an augmented scale), each note is equally important, and notes do not bear a unique position in the scale. But commonplace scales (such as minor or major scales) with non-regular interval leaps do provide a frame of reference that the listener can use to orient themselves with respect to the key of the music.

### 5 THE FRAMEWORK AND INTERFACE FOR AESON TOOLKIT

The AeSon Toolkit is made up of two parts: a set of abstractions in Max/MSP that work together to import data and map it to synthesis and sampling objects, and a graphical user interface that provides real-time interactive control over the chosen mappings and configuration as the sonification is being played.

#### 5.1 AeSon Framework

Figure 2 is a complete sonification patch, using a signal chain in which data is stored (Dataset.Storage), smoothed (Var.Cont.Smoothing) mapped to pitch (Mapping.Pitch), synthesized, enveloped with rhythmic accents (Mapping.Rhythm) and manipulated in level (Manipulate.Level), all according to the data value. The control and audio signals are passed down the left-hand side inlets and outlets in the typical Max/MSP manner, but the imported data (which is parsed and structured) is passed down a separate chain (green path in Figure 2), which allows each of the objects to function independently with respect to the imported dataset.

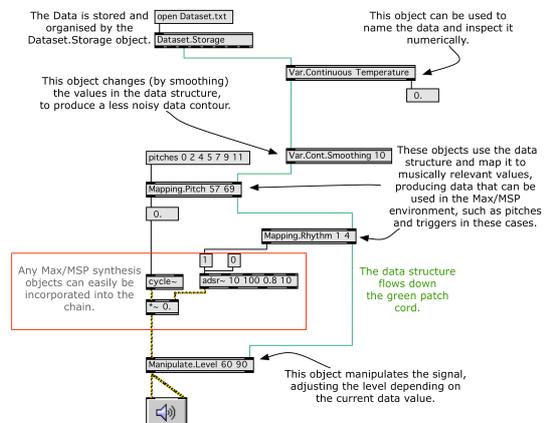


Figure 2. The Max/MSP signal chain shows the way in which data is stored, smoothed, and mapped into a sonification process, and manipulated in rhythmic accent and level according to the data value.

The various mapping and synthesis processes are undertaken in real-time – there is no ‘compilation’ of the sonification – and therefore each process can be altered in real-time. An advantage of this approach is that many interaction devices can be easily incorporated into a patch and therefore can be used to control the mapping and sonification processes. Also, real-time data and static data can be treated in almost exactly the same way, and used simultaneously in sonifications.

There are several categories of object in the framework, Data-Handling routines import data (including realtime data), control its presentation and perform various statistical and filtering tasks. Mapping routines convert data from the data structure into auditory parameters, such as pitch values, levels or rhythmic triggers. Sampling and Synthesis routines create audio signals from these auditory parameters. Manipulate objects perform digital signal processing on signals for spatial and filtering purposes. Each of these categories of objects comes with a help patch that describes the type of commands the abstractions accept, and give examples of their usage.

### 5.2 AeSon Interface

To provide an intuitive and general user interface for the AeSon Framework, we have designed the AeSon Interface to abstract many of the mapping concepts into a simple sonification interface. This is a ‘bolt-on’ graphical user interface that communicates with a sonification system that uses the AeSon framework. Data is imported and then each data ‘stream’ is represented as an audio stream in a spatially arranged interface.

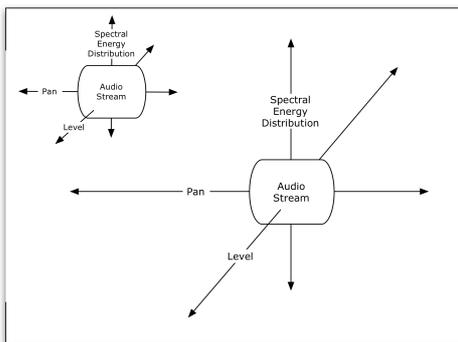


Figure 2: The spatial layout metaphors used for the control of all sonification stream objects on the canvas. Objects are positioned in auditory space, panned left and right or height transformed by adjusting spectral content. Moving the objects towards the user alters relative level.

Figure 2 shows that lateral motion pans an object in auditory space, vertical motion transforms the spectral content to lend an impression of vertical height and depth, while z-axis position (or distance) controls the level of the stream. This spatial treatment is intuitive, but is also based on the investigations of Song and Beilharz [35] and that of Ferguson and Cabrera [36]. It is consistently applied to all streams of data in the interface and thereby allows the user to separate multivariate data, prioritise interesting streams or increase loudness to assist in the analysis. In a creative performative context this feature can be used like a mixer. This interface was especially designed with physical and tangible contexts for interaction in mind e.g. touchpad, touch-screen, graphics tablet or Wii-mote. Canvas position can be utilised to distinguish multivariate data-streams through the reinforcement of the corresponding visual and auditory location in the interface.

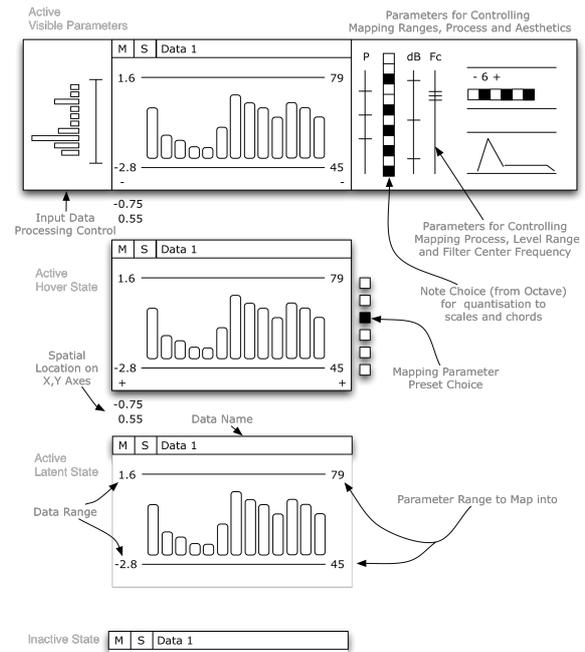


Figure 3: Object activation states. Objects can be customised in vertical scale (pitch range and distribution), horizontal spatial location, envelope characteristics and pitches/modality for mapping; the data-mapping range of pitch, loudness and length are selected by the user with the sliders.

Tufte and Fry place great importance on malleability of data and iteration, fine-tuning and re-contextualisation (namely interaction) during explorative phases and refining of the information representation during the intervention/refinement stage. Thus, we enable the user to hear the same data at many different configurations of concurrent streams in multivariate data, different tempo, keys, scaling, granularity, distribution over pitch and to foreground different elements to thoroughly explore the data from various perspectives. This live interaction may also enhance the engagement of the data examination.

Figure 3 shows the various activation states of sonification objects. Each stream object expands to reveal more layers of control in the hover and active menu states. The user can use familiar sliders and scaling axes to manipulate the scope of pitch, level and filter, in terms of the data it represents. In the active state, the user selects pitch values that constitute the mode/scale by choosing intervals within the octave for quantising data values, as well as pitch range/spacing, note length and loudness/level.

Further technical information on the AeSon Toolkit can be obtained by downloading the framework and using the accompanying documentation.

## 6 APPLICATIONS AND CONTINUING WORK

AeSon is designed to suit versatile contexts for interacting with data, such as with mobile handheld multi-touch screens (e.g. iPhone), tangible and tactile surfaces. Its application may range from analytical and information-seeking activities to performative live generation of music from data. Our continuing work includes the evolution of the interface and experimentation with multi-touch integration, for example with the newly released *beta* framework. We are currently investigating aesthetic feedback with a survey exploring qualitative impressions concerning durations, overlap, rhythmical emphasis, pitch distribution and correlation to visual

graphs to further inform the design of sound controls in the AeSon Toolkit.

## 7 REFERENCES

- [1] G. Weinberg and T. Thatcher, "Interactive sonification: Aesthetics, functionality and performance," *Leonardo*, vol. 16, pp. 9-12, 2006.
- [2] C. M. Wilson and S. K. Lodha, "Listen: A Data Sonification Toolkit," in *Int. Conf. on Auditory Display*, Palo Alto, CA, 1996.
- [3] S. K. Lodha, J. Beahan, T. Heppe, A. Joseph, and B. Zane-Ulman, "MUSE: A Musical Data Sonification Toolkit," in *Int. Conf. on Auditory Display*, Palo Alto, California, 1997.
- [4] A. J. Joseph and S. K. Lodha, "MUSART: Musical Audio Transfer Function Real-time Toolkit," in *Int. Conf. on Auditory Display*, Kyoto, Japan, 2002.
- [5] T. M. Madhyastha, and Reed, D. A. , "A Framework for Sonification Design," in *Auditory Display: Sonification, Audification and Auditory Interfaces* G. Kramer, Ed.: Addison-Wesley, 1992.
- [6] B. K. Davison and B. N. Walker, "Sonification Sandbox Reconstruction: Software Standard for Auditory Graphs," in *Int. Conf. on Auditory Display*, Montreal, Canada, 2007.
- [7] B. N. Walker and J. T. Cothran, "Sonification Sandbox: A graphical toolkit for auditory graphs," in *Proceedings of the 2003 Int. Conf. on Auditory Display*, Boston, MA, USA, 2003.
- [8] P. Vickers, "Whither and wherefore the auditory graph? Abstractions and aesthetics in auditory and sonified graphs," in *11th Int. Conf. on Auditory Display*, Limerick, Ireland, 2005.
- [9] B. N. Walker, "Magnitude estimation of conceptual data dimensions for use in sonification," *Journal of Experimental Psychology: Applied*, vol. 8, pp. 211-221, 2002.
- [10] A. d. Campo, C. Dayé, C. Frauenberger, K. Vogt, A. Wallisch, and G. Eckel, "Sonification as an interdisciplinary working process," in *Int. Conf. on Auditory Display*, London, UK, 2006.
- [11] A. de Campo, C. Frauenberger, and R. Höldrich, "Designing a Generalized Sonification Environment " in *Int. Conf. on Auditory Display*, Sydney, Australia, 2004.
- [12] D. Worrall, "Overcoming Software Inertia in Data Sonification Research Using The SoniPy Framework" in *Int. Conf. on Music Communication Science*, Sydney, NSW, Australia, 2007.
- [13] D. Worrall, M. Bylstra, S. Barrass, and R. Dean, "SoniPy: The design of an extendable software framework for sonification research and auditory display," in *Int. Conf. on Auditory Display*, Montreal, Canada, 2007.
- [14] K. Jo and N. Nagano, "MonaLisa: 'see the sound, hear the image'," in *8th Int. Conf. on New Interfaces for Musical Expression*, Genova, Italy, 2008.
- [15] W. S. Yeo, J. Berger, and R. S. Wilson, "A Flexible Framework for Real-Time Sonification with SonART" in *Int. Conf. on Auditory Display*, Espoo, Finland, 2001.
- [16] S. Pauletto and A. Hunt, "A toolkit for interactive sonification," in *Int. Conf. on Auditory Display*, Sydney, Australia, 2004.
- [17] S. Barrass, "Auditory Information Design." Ph.D Thesis: Australian National University, 1997.
- [18] F. Dombois, O. Brodewolf, O. Friedli, I. Rennert, and T. Koenig, "SoniFyer: A Concept, a Software, a Platform," in *14th Int. Conf. on Auditory Display* Paris, France, 2008.
- [19] N. Schnell, R. Borghesi, D. Schwarz, F. Bevilacqua, and R. Müller, "FTM — Complex data structures for Max," in *International Computer Music Conference*, Barcelona, 2005.
- [20] J. W. Tukey, *Exploratory Data Analysis*. Reading, Mass.: Addison-Wesley, 1977.
- [21] B. N. Walker and M. A. Nees, "An agenda for research and development of multimodal graphs," in *11th Int. Conf. on Auditory Display*, Limerick, Ireland, 2005.
- [22] B. Fry, "Computational Information Design." Ph. D Thesis: Massachusetts Institute of Technology, 2004.
- [23] B. Fry, *Visualizing Data*: O'Reilly Media, Inc., 2008.
- [24] J. Maeda, *The Laws of Simplicity*. Cambridge, Massachusetts: MIT Press, 2006.
- [25] A. Vande Moere, "Infosthetics: Where form follows data," <http://infosthetics.com/>, 2008.
- [26] E. R. Tufte, *The Visual Display of Quantitative Information*. Cheshire, Conn: Graphics Press, 1983.
- [27] E. R. Tufte, *Envisioning Information*. Cheshire, Conn: Graphics Press, 1992.
- [28] A. Lau and A. V. Moere, "Towards a Model of Information Aesthetics in Information Visualization " in *11th Int. Conf. on Information Visualization* Zürich, Switzerland: IEEE Computer Society, 2007, pp. 87-92.
- [29] E. Childs, "Musical Sonification Design," Ph.D Thesis, Hanover, New Hampshire: Dartmouth College, 2003.
- [30] G. Grinstein and M. O. Ward, "Introduction to Data Visualization," in *Information Visualization in Data Mining & Knowledge Discovery*, Ed. U. Fayyad, G. G. Grinstein, & A. Wierse: Morgan Kaufmann, 2002.
- [31] P. Vickers and B. Hogg, "Sonification Abstraite/Sonification Concrète: An 'Æsthetic Perspective Space' for Classifying Auditory Displays in the Ars Musica Domain," in *12th Int. Conf. on Auditory Display* London, UK, 2006, pp. 210-216.
- [32] K. Beilharz, "Designing Sounds and Spaces: Interdisciplinary Rules & Proportions in Generative Stochastic Music and Architecture," *Journal of Design Research*, vol. 3, <http://jdr.tudelft.nl>, 2004.
- [33] P. Vickers, "Ars Informatica - Ars Electronica: Improving Sonification Aesthetics," in *Understanding and Designing for Aesthetic Experience Workshop, Proc. HCI 2005 The 19th British HCI Group Annual Conference*, L. Ciolfi, M. Cooke, O. Bertelsen, and L. Bannon, Eds. Edinburgh, Scotland, 2005.
- [34] R. N. Shepard, "Geometrical approximations to the structure of musical pitch," *Psychological Review*, vol. 89, pp. 305-333, 1982.
- [35] H. J. Song and K. Beilharz, "Concurrent Auditory Stream Discrimination in Auditory Graphing," *Journal of Computers*, vol. 3, pp. 79-87, 2007.
- [36] S. Ferguson and D. Cabrera, "Vertical localization of sound from multiway loudspeakers," *Journal of The Audio Engineering Society*, vol. 53, 2005.