1-1-2015

Spectromorphology and spatiomorphology of sound shapes: Audio-rate AEP and DBAP panning of spectra

Stuart James

Edith Cowan University, s.james@ecu.edu.au

James, S. (2015). Spectromorphology and spatiomorphology of sound shapes: Audio-rate AEP and DBAP panning of spectra. In *Proceedings of the 41st International Computer Music Conference 2015* (pp. 278-285). Available here

Spectromorphology and Spatiomorphology of Sound Shapes: Audio-Rate AEP and DBAP Panning of Spectra

Stuart James
Edith Cowan University
s.james@ecu.edu.au

ABSTRACT

Explorations of a new mapping strategy for spectral spatialisation demonstrate a concise and flexible control of both spatiomorphology and spectromorphology. With the creation of customized software by the author for audio-rate histograms, spectral processing function smoothing, spectral centroid width modulation, audio-rate distance-based amplitude panning, audio-rate ambisonic equivalent panning, a growing library of audio trajectory functions, and an assortment of spectral transformation functions, this article aims to explain the rationale of this process.

1. INTRODUCTION

This research has led toward the development of several spectral spatialisers that explore and utilize topographies and geometries in *Wave Terrain Synthesis* as a means of controlling the distribution of sound spectra; this has also involved an exploration of *morphologies* of such distributions. This article focuses on one of these spatialisers. In its most flexible form this spatialiser diffuses different spectral components of a source sound across a listener space (shown in Figure 1a) using potentially any arbitrary speaker arrangement in both 2D and 3D arrangements. These *particles* or points of spectra in space, as pictured in Figure 1, are each differentiated in frequency by their associated color. Whilst these points are reconstructed across the listener space, the *sound shape* can be distorted geometrically across space, resulting in a *spatiomorphology* as shown in Fig 1b.

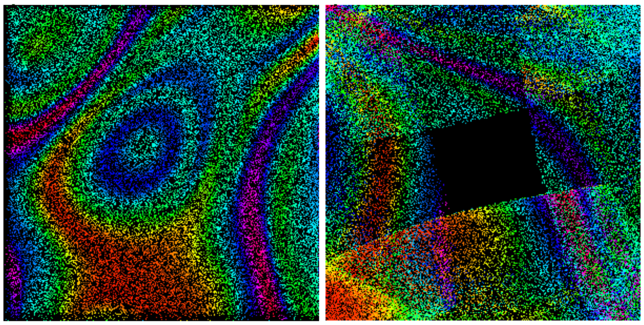

Figure 1 The image on the left shows a spatial distribution of spectra using software by the author. Spectral bands are differentiated by color. The image on the right shows a subsequent spatiomorphology of that distribution of spectra.

A *spectromorphology* is also possible with this spatialiser resulting in a perceived change in *timbre*. [1] Each point in this distribution is generated at the audio sampling rate. If one were to use a sampling rate of 44.1KHz this manages 44,100 points of spectra per second. The distinction made by the author of *timbre spatialisation*, as opposed to *spectral spatialisation*, is the exploration of both *spatiomorphology* and *spectromorphology* in combination.

Synthetic spatialization offers the computer musician the ability to render virtual moving sound sources across a discrete speaker array. Such spatialisation techniques generally fall into two categories: the first involves the simulation of localized and moving sound sources, what is often referred to as point-source panning aimed at rendering an auditory illusion of a sound moving across a listener area. This work is largely influenced by the work of John Chowning who published a paper on this, and who experimented with the coordinate control of virtual moving sound sources on computer over a quadraphonic system. [2][3] Another group of techniques are aimed more specifically at generating immersion. These are sounds that have a sense of spatial extent or what Denis Smalley describes as *circumspectral* space. [1]

The first category includes perception-based amplitude panning techniques like Vector-based amplitude panning (VBAP) [4] and Distance-based amplitude panning (DBAP) [5], Ambisonic panning, and Binaural panning. All techniques listed with exception to Binaural are adaptable to different speaker configurations supporting 2D horizontal arrays, and 3D spherical dome speaker arrays. Binaural panning techniques are instead intended for stereo headphone reproduction; whilst they model a 3D listening experience, the algorithm uses head-related transfer functions (HRTF) to reconstruct the illusion of sounds moving around the listener in 3D.

Immersive techniques, on the other hand, include rapid panning modulation synthesis [6], spectral decorrelation, spectral spatialisation [17], spectral splitting [7], and spectral delays. As a further distinction, Schumacher and Bresson use the term ‘spatial sound synthesis’ to denote any sound synthesis process that is extended to the spatial domain. [8] A number of these techniques include spatial swarm-granulation [9], Sinusoidal Modulation Synthesis [10], timbre spatialisation [11], and Spatio-Operational Spectral (S.O.S.) Synthesis [12]. Generally these fall into the immersive category, but results can also be quite localized too.

2. CONTROLLING SPECTRAL SPATIALISATION

Torchia and Lippe [15] and Kim-Boyle [16][17] and some others have expressed the need for researching other mapping strategies for *spectral spatialisation*. Most existing methods utilize control rate and semi-autonomous approaches to spectral spatialisation, one of the most successful and eloquent solutions being Kim-Boyle's use of the Boids algorithm for generating spatial coordinate sets to be rendered using VBAP panning over a quadraphonic speaker array. Lippe has eloquently described the problem:

> The biggest challenge for composers working in… [the frequency] …domain continues to be exploration of ways to meaningfully manipulate and interpret FFT data in a musical context. In a real-time interactive environment, expectation of some sort of correlation between a musician's performance and the sounding output of a computer system is high, and successful strategies for manipulation and interpretation of FFT data via performer input are paramount. Two mapping problems are completely intertwined: mapping of performer data, and mapping of FFT data. [21]

Hunt and Wanderley define several mapping strategies that have developed over the years: one-to-one, one-to-many and many-to-one, and several of these approaches can be used in combination for a variety of many-to-many mappings. [18]

Van Nort and Wanderley state that a mapping can be *explicit* or *implicit*. The former refers to a situation in which the mapping is known and can be expressed analytically, while the latter is based on internal adaptation of a system, and can be seen as a "black box" model. The explicit approach is beneficial in that having knowledge about the way the mapping occurs allows one to tune, alter, and expand it over time and for different musical contexts.

Van Nort and Wanderley have presented various mapping strategies as functions between geometric spaces of control and sound synthesis parameter sets. [19] Properties such as local/global definition, editability, continuity, linearity, smoothness, etc. take on more or less importance depending on a given musical control context and can even embed interesting musical gestures in the control system if designed properly. [20]

The use of *Wave Terrain Synthesis* as a control mechanism allows the performer to create a complex and coordinated change across an existing multi-parameter system. Whilst the *terrain* surface and *trajectory* geometry represent manifolds, the way in which these structures combine present a wide range of morphologies, what is in essence a parametric space that is itself malleable. The scientific concept of phase space is an important notion here as it represents all parametric possibilities of a system. [13] Previous research in *Wave Terrain Synthesis* by the author evaluated the flexibility of the system for synthesizing and transforming a wide range of curves. [14]

The instrument developed uses an explicit many-to-many mapping to bridge *Wave Terrain Synthesis* to *spectral spatialisation*. Each section of the instrument can be further defined, for instance the *Wave Terrain Synthesis* itself is an explicit many-to-one mapping. In the model shown in Figure 2, the *Wave Terrain Synthesis* lookup process is used to define the frequency or spectral bin to be spatialised, and the spatialisation of the spectra are determined by the real-time spatial coordinate of the trajectory. The result is an eloquent system: the topography of the terrain describes a timbre distribution or *spectromorphology*, whilst the trajectory defines a sound shape or *spatiomorphology*. In order to achieve this, some additional software needed to be developed by the author.

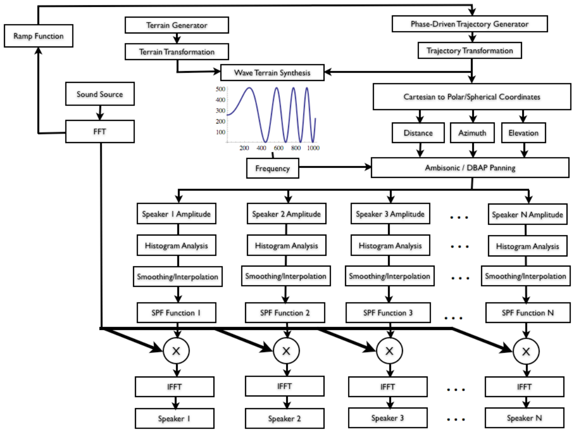

Figure 2 A block diagram describing the instrument.

It is also the case that a performer is not intentionally responsible for the diffusion of every independent frequency bin, but rather is responsible for the global distribution of all frequency bins. An analogy in audio engineering is the notion of a group fader, where the performer or engineer is not responsible for riding the level of multiple individual faders on the mixer, but rather a single fader that controls the proportional level of multiple audio channels. In this way parameter management reduces the necessary burden to one parameter change, hence making the system easier to manage, and simple enough cognitively for the performer or engineer to focus on the sounds themselves, and the relevant musicality of the performance. There are many styles and sub-genres of digital music performance that do not have the communicative aspects of human gestures as their primary concern, and yet real-time control and organization of sound materials is often still of paramount importance. [22]

3. NAVIGATING SPECTROMORPHOLOGY

In 2012 Eric Lyon describes a method where digital images provide the computer with detailed control information for diffusion. [23] Whilst his article describes a very different mapping process than what is investigated here, it was enough for the author to re-evaluate previously explored mapping strategies used for *timbre spatialisation*.

Since the authors' *Wave Terrain Synthesis* instrument uses images and video for terrain maps, it seemed logical to use the topographics (shown as differences in colour) to describe frequency. This aligns the approach with some other graphical sound synthesis methods such as Oramics [24], and Iannis Xenakis's Upic [25] system involve the drawing of sound in time. This concept has been extended into the digital domain with systems including Christopher Penrose's HyperUpic [26], UI Software's Metasynth [27], Izotope's Iris [28], and the Virtual ANS [29] iOS application for iPhone and iPad. Lyon also makes mention of automated mapping of video to space including the work of Shawn Greenlee and Augustus Leudar, a PhD student at SARC[1], who has worked with video to space in a quadraphonic context, using ambisonic tools for spatialization. Kim-Boyle has also been involved in building a perceptually informed interface for control, opting for a visually intuitive method of amplitude scaling spectra by the brightness of pixels in a Jitter matrix. [17]

The point of finding a different methodology to implement *timbre spatialisation* was motivated by the need of establishing a stronger correlation between the terrain structure and the resulting sound diffusion, or a means of translating the topography of the terrain structure from its visual representation to a more literal audible analogy. The fundamental motivation was also to be able to "paint" timbre across space. Graphical synthesis systems like the UPIC and Oramics were an influencing factor in the initial conception of this mapping.

3.1 Generating the Timbre Distribution

This process is essentially three-fold: firstly values are read from the terrain map using coordinates determined by the trajectory, the coordinates of the trajectory calculate an n-channel weighted-distribution dependent on n number of loudspeakers, and finally these weighted values are stored according to their spectral bin in n histograms that calculate the *Spectral Processing Function* (SPF) for each speaker.

Early implementations of this by the author utilized some of the processing functions in the *Jitter* library for *MaxMSP*. Objects such as jit.histogram and jit.bsort. At the time the model was also restricted to basic amplitude panning. This approach was highly problematic as each histogram computed in Jitter resulted in audible segmentation artifacts as they are much slower than the window size of the FFT. Consequently it became apparent that the author needed to develop an audio-rate histogram synchronized with the FFT. Furthermore, in order to compute each spectral band with the necessary weightings for speaker channels for a range of different speaker configurations, it was decided by the author to develop audio-rate implementations of Distance-based amplitude panning (DBAP) [5] and Ambisonic Equivalent panning (AEP) [30], shown in Figure 3, that would allow the system to be transportable and compatible with the scope of different spatial music installations around the world.

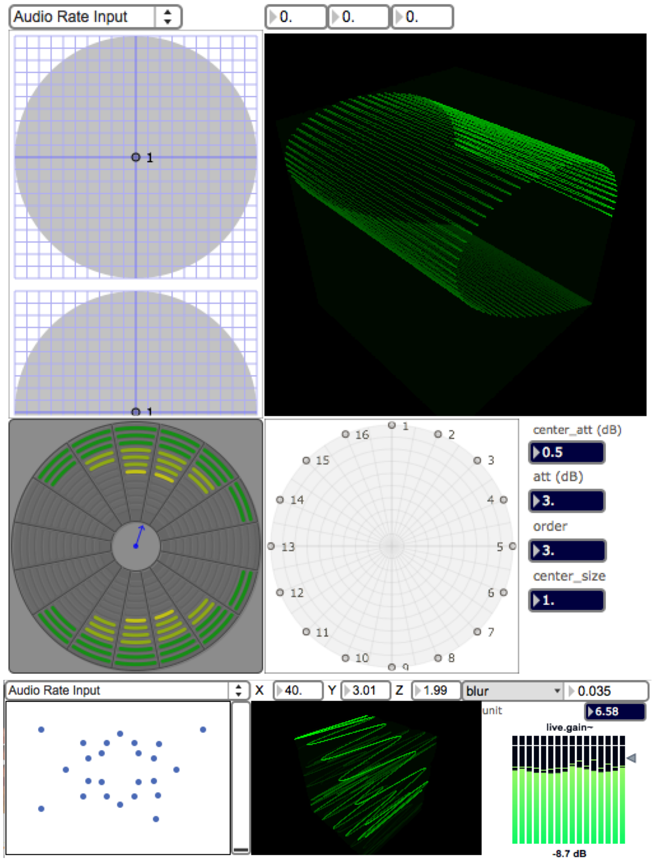

Figure 3 A presentation view of the audio-rate implementations of AEP and DBAP. These patches also include GUI objects of the HOA and ICST ambisonics externals, as it was important to ensure compatibility.

The importance of the histogram in statistical analysis is that it can accumulate a history of one parameter over a given time interval. The use of the histogram solved an issue whereby the information generated from the terrain was inherently bound by the order in which these events were read. By resorting to a histogram analysis of the terrain topography it means that no assumptions need to be made about the "order" in which this data is read from the terrain. This means that the trajectory could traverse in the opposite direction, and the results should be comparably the same.

These processes ultimately assign a frequency band with a spatial coordinate. The plots in Figure 4 show the terrain topography and the resulting spectral distribution. Notice how different trajectories reveal different sections of the terrain surface.

---

[1] *Sonic Arts Research Centre*, Queens University, Belfast

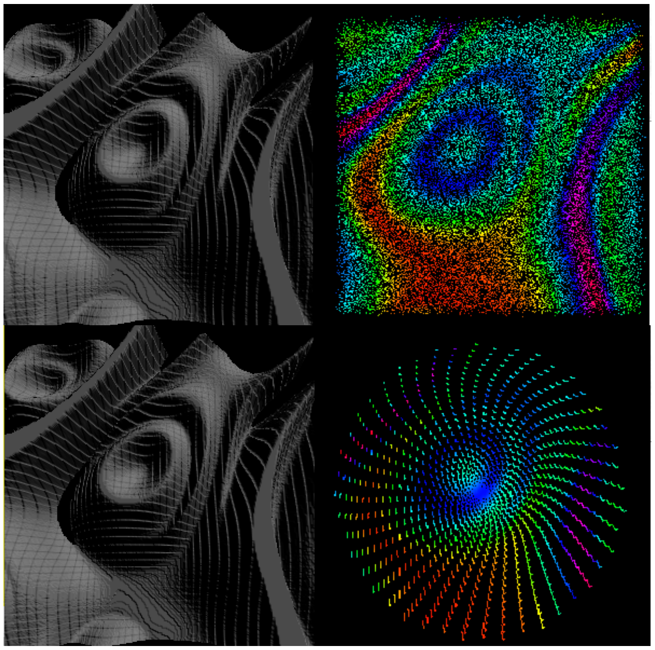

Figure 4 Top left is an image of the terrain surface, and top-right the resulting spectral distribution with a noisy lookup trajectory. Below is the same terrain, but a different sound shape that is revealed by a spiral trajectory.

3.2 Smoothing Histogram Data

Depending on the resolution of the data read from the terrain, this will influence how "rough" the histogram appears. Figure 5 shows an audio-frame of a histogram generated and visualized as a video frame and graphed below. This shows one of the inherent problems of statistical analysis in that data isn't smooth, but it clumps in certain regions of the data range.

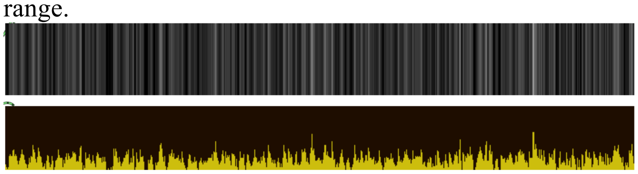

Figure 5 A typical histogram showing a clumpy distribution

Smoothing algorithms are essential, and whilst several were considered, these algorithms will have taken too long for computation. Since these had to be computed within a small number of FFT frames, it was decided to focus on two implementations: *spectral centroid* and *catmull-rom spline interpolation*. Both implementations are by the author. Spectral centroid has been generally the preferred, being the most efficient, and allows the user to control spectral centroid width (shown in Figure 6), spectral wraparound so the user can control whether high frequencies fold over into low frequency bands and vice versa. This spectral width control allows for a very musical way of shifting from noisy sounds to pitched sounds. The catmull-rom spline interpolator allows the user to control the interpolation width. Both of these processes were implemented in *gen~*.

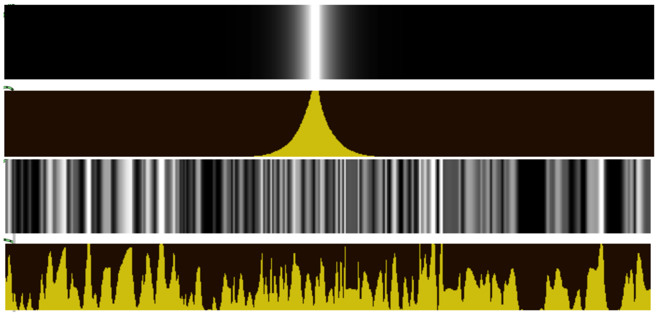

Figure 6 The use of the centroid width modulation applied to a single frequency band shown above, and catmull-rom spline interpolation applied to a noisy distribution below.

3.3 Morphology of the Spectral Topography

This research has explored a range of different terrain surfaces and their morphologies. For algorithmically generated terrains using *jit.expr* and *jit.bfg*, it is possible to modulate the surface in interesting ways given the surface describes a mathematical function. However other terrains have been explored such as those generated by video or those achieved through the visualization of other kinds of data. Video feedback presents some interesting possibilities.

Techniques of morphology have focused on affine transformations, such as scale, translation, and rotation. Spatial transformations and topographical deformations of the curves have been a point of interest too.

4. NAVIGATING SPATIOMORPHOLOGY

The exploration of sound shapes is ultimately defined by the geometries exhibited by a *trajectory*, a multi-dimensional audio-rate signal. Whilst this family of different trajectories could be classified categorically based on their computational derivation, that is whether they are derived by simple algebraic, trigonometric, iterative, or procedural processes, this did not feel appropriate given the application. Similarly, classifying them by their periodicity versus quasi-periodicity or chaotic and stochastic qualities was note appropriate either. If the trajectories were being used within the context of the FFT frame, they might be categorized by their synchronicity or asynchronicity with respect to the FFT frame. Another contributing factor relates to the space-filling qualities of the trajectory curve.

Since the trajectory is evolving through time, one of the considerations of this mechanism is how this may influence *spatial texture*. Smalley states that spatial texture is concerned with how the spatial perspective is revealed through time, and that it is ultimately a question of *contiguity*. He suggests spatial texture is divided into: *non-contiguous space* (which is revealed when spectromorphologies are presented in different spatial locations such that two successive events are not considered near neighbours and *contiguous space* (which is subdivided into *spread settings* and *trajectories*). Smalley also states that the differences between contiguous and non-contiguous spaces are not definite and that a space can be non-contiguously erratic at a low level but contiguous at a higher level. [31]

4.1 Generating Sound Shapes

All of these curves are derived algorithmically, however there is no reason why they couldn't also draw from existing data sets. All of the curves render coordinates in three dimensions where possible.

The least space-filling curves tend to be periodic and FFT synchronous (i.e. those that follow the same repeated path). The geometry expressed by these shapes can be any periodic curve, whether continuous or discontinuous parametrically in the time domain. This includes geometric archetypes such as circles, squares and other polygons. It also includes slowly moving single dimensional pathways such as slow random walks. Even some chaotic functions under certain conditions behave in a very stable and repeating way.

Curves that exhibit higher degrees of space-filling characteristics, although this is largely contextual and dependent on variables, include spirals, the Rose Curve, hypotrochoids, iterative functions, and stochastic and random processes.

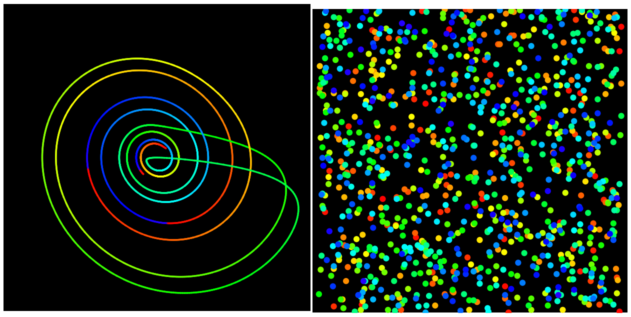

Figure 6 A typical chaotic attractor and a random distribution showing very different space-filling properties. One curve is more localized, whilst the other spreads across the entire space.

With this mapping strategy, if the contour of a trajectory is changing shape when moving from one FFT frame to the next, it will reveal different aspects of the terrain surface over time. However by freezing the coordinates of the trajectory it will also freeze the spectrum generated, and ultimately the sound shape. This is particularly applicable to noisy systems such as a Pink Noise Generator:

```
case 17:
//Pink Noise using Paul Kellet's economy method
for(i = 0; i < o1.length;i++)
{
    x = (float)Math.random();
    p1 = (float)0.99765 * p1 + x * (float)0.0990460;
    p2 = (float)0.96300 * p2 + x * (float)0.2965164;
    p3 = (float)0.57000 * p3 + x * (float)1.0526913;
    mathx = p1 + p2 + p3 + x * (float)0.1848;
    y = (float)Math.random();
    p4 = (float)0.99765 * p4 + y * (float)0.0990460;
    p5 = (float)0.96300 * p5 + y * (float)0.2965164;
    p6 = (float)0.57000 * p6 + y * (float)1.0526913;
    mathy = p4 + p5 + p6 + y * 0.1848;
    z = (float)Math.random();
    p7 = (float)0.99765 * p7 + z * (float)0.0990460;
    p8 = (float)0.96300 * p8 + z * (float)0.2965164;
    p9 = (float)0.57000 * p9 + z * (float)1.0526913;
    mathz = p7 + p8 + p9 + z * (float)0.1848;
    o1[i] = ((float)mathx-(float)26)*(float)0.25;
    o2[i] = ((float)mathy-(float)26)*(float)0.25;
    o3[i] = ((float)mathz-(float)26)*(float)0.25;
}
```

4.2 Interpolating Across a Library of Sound Shapes

A number of researchers have explored the navigation of a parametric "surface" or hyper-plane which represents all traversable regions of a parameter space. These are derived through the interpolation, extrapolation, and regression of control/sound data. [22] Several mapping toolkits, LoM (Library of Maps) [32] and MnM (Mapping is not Music) [33], have also been developed to facilitate higher dimensional navigation across multi-parametric sets, and providing access to a toolbox of few-to-many mappings. Some other implementations using interpolation and regression techniques include the manifold interface [13], Momeni and Wessel's (2003) models, the MetaSurface (Bencina, 2005). Most other interpolated mapping strategies approach this problem with slower control rates. This project, on the other hand, is primarily interested in audio-rate morphology, that is changes that occur to the trajectory curve that result in audible transformations in the spatial distribution of spectra.

In the case of the *timbre spatialisation* it was ideal to have the choice of multiple curves, and to navigate through this space, interpolating across the differences. The interest here is in smoothly moving from one geometric state to another. This can be achieved simply through interpolating the Cartesian points of one state to the next. In mathematics this would be considered a continuous bijection from one curve to a different curve. The ability to shift from one state to another, and explore these "in between" states is a highly expressive control parameter. By using a 2D slider interface, early versions of this controlled the relative amounts of a periodic, quasi-periodic, chaotic, and random trajectory sources, much like the principle of *vector synthesis*.

This research uses DBAP to interpolate across many sources, a principle that extends the principle of equal intensity panning from a pair of speakers to a loudspeaker array of any size, and is adaptable to arbitrary input and output configurations [13]. The significant point of departure from traditional DBAP is rather than diffusing an input source across a multichannel speaker array, we are interpolating across a multi-channel input source for summing into a single output channel. The advantage of DBAP is the ability for the algorithm to adapt appropriate loudness curves where different sound sources might normally intersect. This also ensures that loudness roll-off curves are extended for where sources do not intersect.

The *nodes* object in *MaxMSP* can be integrated with the DBAP model for a more compact way of controlling and visualizing these relationships. This does not inherently use

DBAP's own interpolated weights for each node, but rather is used as a 2D graphical interface for control and visualization. The interpolation is still performed by DBAP.

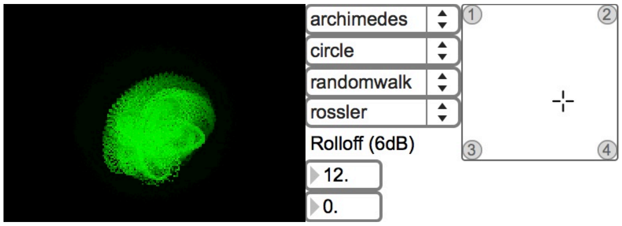

Figure 7 The trajectory interface using the nodes object as a 2D slider. The DBAP implementation is performed at control-rate in Javascript

The flexibility of the nodes 2D interface and DBAP interpolation is that such an approach is not limited to a fixed number of nodes, but rather it can be extended in non-linear ways to allow for many may varied states of expression. In this way this notion of inbetween states becomes more fruitful and rich the more nodes that exist offering a very powerful and intuitive way of navigating the many possible trajectory curves that exist.

4.3 Transformation and Evolution of Sound Shapes

This may arguably be the most interesting and powerful aspect of this instrument: the ability to transform spectral distributions at audio rates. All of these processes apply after the lookup process, but before the spatial coordinates have been used to derive the various speaker weightings for respective spectra.

With affine transformation of all of the coordinate points it is possible to achieve many of the scene-based spatial gestures Trevor Wishart documents including scene rotation, contraction and expansion, translation, swing, twist, and spiral. (Wishart, 1996)

Further transformations can be applied through time-domain signal processes such as low and high-pass filters, band-pass filters, bit reduction, sample-rate reduction, and feedback systems.

Individual spatial transformations may be applied to all of the bins independently too such that each spectral bin has its own trajectory and evolution. These may be randomized, periodic, or they may explore chaotic behaviors or swarm-based characteristics – cpu permitting!

Spectral Freeze can also be useful at this stage of the process. This is particularly the case for random and chaotic distributions where there is often no temporal synchronicity with the spectral process. Once the coordinates are frozen spatially, they can then be transformed using another process. There are several scenarios that are useful here. For instance, if one takes a spatial formation that is discontinuous, such as white noise, one is able to "freeze" a frame of this white noise, essentially freezing all of the 2D coordinates associated with all FFT frequency bins. These frozen coordinates can then each undergo some kind of modulation, such as oscillation, random 2D walks, or they can be given behaviors. Since these values are computed at audio-rate it does place some limitations on the nature of these behaviors too given the quantity of points used. In this way the Boids algorithm would be far too intensive. Particle systems are also of interest.

Here is a simple procedural section of Java code responsible for generating random circular orbits for all spectral bins.

```
        case 19:
        //Circles
        for(i = 0; i < o1.length;i++)
        {
            var2 = (int)in1[i];
            var4 = (float)((myInitial[var2]) * (Math.cos(myNew[var2] + (p1
* myInitial[var2]))));
            var5 = (float)((myInitial[var2]) * (Math.sin(myNew[var2] + (p1
* myInitial[var2]))));
            o1[i] = var4;
            o2[i] = var5;
            o3[i] = (float)0;
            myNew[i] = Math.abs(myNew[var2]+(p1*myInitial[var2]));
        }
        break;
```

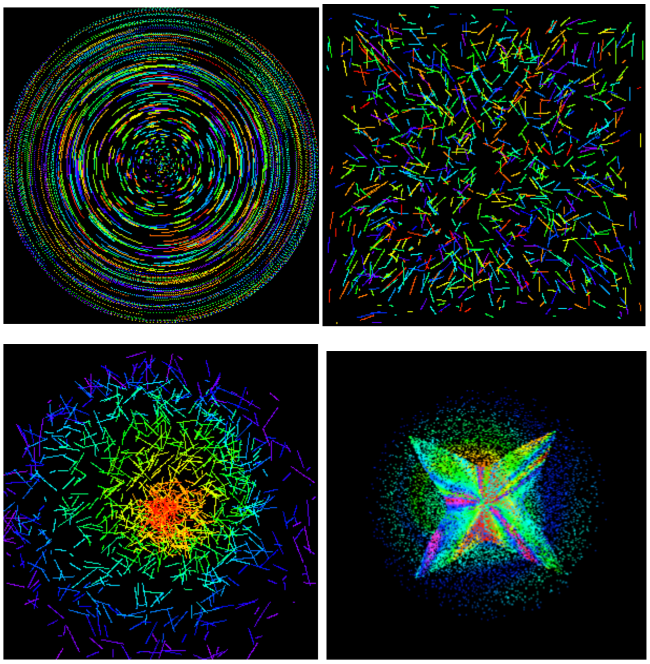

Figure 8 Top Left is a circular independent distribution of spectral points, Top Right is a randomized independent distribution of spectral points, Bottom left is a morphology between a random distribution applied to a fixed spiral formation, and bottom right is an affine transformation of the sound shape shown in Figure 4

5. ARTISTIC WORK

A performance opportunity arose at the end of October 2014 at the Pakenham Street Art Space, an opportunity where there was intended to be a 32-channel sound system installed. This immediately became the opportunity I had been waiting for to test these new spatialiser tools. Rather than prepare a new piece, I wanted to re-work and develop upon a previous composition. *The Overview Effect* was a composition by the author that utilized most of the software discussed in this article.

The speaker installation consisted of two inner circular arrays of eight speakers each, four additional distant speakers set across the room, and four sub-woofers. There were originally intended to be an additional eight channels of speakers hanging from the roof pointing down onto the audience, however this became a logistical complication with the venue. The originally proposed schematic is shown below.

The importance of this opportunity was due to the fact that the DBAP and Ambisonic software could be tested within a multichannel system with enough distance across space to really maximize the use of spatial gesture. It was also an opportunity to compare the existing Jamoma DBAP model with the audio-rate version I had coded, and furthermore compare the ICST, HOA, and Jamoma Ambisonic models with the audio-rate version I had coded too. The coordinates of the speakers were calculated with respect to the coordinate dimensions of the room. This gave a concrete series of points from which to program the DBAP panners and Ambisonic panners used.

The coordinates routed into the Jamoma DBAP implementation seamlessly, as well as the authors own implementation. These were subsequently hard-coded into the software used at the gig.

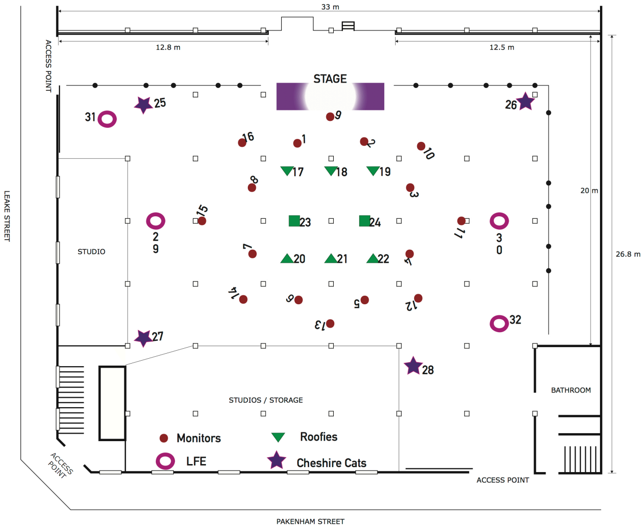

Figure 9 A birdseye view of the Pakenham Street Art Space with the speaker coordinates highlighted.

6. CONCLUSIONS

In the case of *Timbre Spatialisation*, a successful evaluation could be said to include the extent of the performers ability to shape timbre, generate a variety of sound shapes and spatial gestures, and the ability to control some relevant psychoacoustic parameters such as perceived source location, source width, spatial clarity, and immersiveness. An expressive control strategy should allow the performer the freedom and scope to be able to explore outcomes that are both predictable, as well as leaving the opportunity open for exploring those results that are considerably more chaotic, unpredictable, and esoteric. This research has not addressed the perceptual evaluation of these different implementations, but rather several possible mappings are presented with supporting theory and rationale, and the quantitative data generated has been subject to qualitative discussion.

Acknowledgments

Thanks to Dr Cat Hope for her support as my PhD research supervisor.

7. REFERENCES

[1] Smalley, D. “Space-Form and the Acousmatic Image.” *Organised Sound*, 12(1), 2007.

[2] Zvonar, R. (2005) “A History of Spatial Music: Antecedents from Renaissance Antiphony to Strings in the Wings.” http://cec.sonus.ca/econtact/7_4/zvonar_spatialmusic.html

[3] Chowning, John. (1971) “The Simulation of Moving Sound Sources.” *Journal of the Audio Engineering Society* 19: 2–6.

[4] Pulkki, V. (1997). “Virtual sound source positioning using vector base amplitude panning.” In *Journal of the Audio Engineering Society*, 45(6), 456–466.

[5] Lossius T., Baltazar, P. & de la Hogue T. (2009): “DBAP - Distance-Based Amplitude Panning”. *In Proceedings of the International Computer Music Conference*, Montreal, 17-21

[6] Schmele, T. and I. Gomez. (2012) “Exploring 3D Audio for Brain Sonification.” Proceedings of the 18th International Conference on Auditory Display, Atlanta and Schmele, T. 2011. Exploring 3D Audio as a New Musical Language. Masters Thesis. Music Technology Group, Department of Information and Communication Technologies, Universitat Pompeu Fabra

[7] Harrison, J.; Wilson, S. (2010). “Sound <–> Space: New approaches to multichannel music and audio”, *Organized Sound*, 15(3), pp. 183-184.

[8] Schumacher, M. & Bresson J. (2010). Spatial Sound Synthesis in Computer-Aided Composition. In *Organised Sound* , 15(3), 271–289.

[9] Wilson, Scott (2008). Spatial Swarm Granulation. In *Proceedings of the International Computer Music Conference*. Belfast, Ireland.

[10] Cabrera, A. and G. Kendall. (2013) “Multichannel Control of Spatial Extent Through Sinusoidal Partial Modulation (SPM)." *Proceedings of the Sound and Music Computing Conference 2013*. Stockholm, Sweden.

[11] Normandeau, R. (2009) “Timbre spatialisation: The Medium is the Space.” *Organised Sound*.

[12] Topper, D., M. Burtner, and S. Serafin. (2002) “Spatio-Operational Spectral (S.O.S.) Synthesis.” Proceedings of the 5th International Conference on Digital Audio Effects (DAFX-02), Hamburg, Germany.

[13] Choi, I., R. Bargar, and C. Goudeseune. (1995) “A Manifold Interface for a high dimensional control space.” *Proceedings of the 1995 International Computer Music Conference*, Banff Centre for the Arts, Canada

[14] James, S. (2005) “Developing a Flexible and Expressive Realtime Polyphonic Wave Terrain Synthesis Instrument based on a Visual and multidimensional methodology.” (Masters Diss., Edith Cowan University).

[15] Torchia R, and Lippe, C. 2003. “Techniques for Multi-Channel Real-Time Spatial Distribution Using Frequency-Domain Processing.” *Proceedings of the International Computer Music Conference*, Singapore: 41-44.

[16] Kim-Boyle, D. 2006. “Spectral and Granular Spatialization with Boids” in the *International Computer Music Conference Proceedings*, New Orleans, USA.

[17] Kim-Boyle, D. 2008. “Spectral Spatialization: An Overview.” *Proceedings of the International Computer Music Conference*, Belfast.

[18] Hunt, A. and M. Wanderley (2002). Mapping Performance Parameters to Synthesis Engines. *Organised Sound 7*(2), 97– 108.

[19] D. Van Nort, M. Wanderley, and P. Depalle. (2004) “On the choice of mappings based on geometric properties.” In *Proc. of 2004 Conference on New Interfaces for Musical Expression (NIME)* ., pages 87–91.

[20] D. Van Nort and M. Wanderley. (2006) “Exploring the Effect of Mapping Trajectories on Musical Performance.” In *Proc. 2006 Sound and Music Computing Conference* (SMC), pages 19–24.

[21] Lippe, C. “Real-time Interactive Strategies for Timbral Control in the Frequency Domain.” The 4th Sound and Music Computing Conference, 2007, Lefkada, Greece.

[22] Van Nort, D.& Wanderley, M. (2007). Control strategies for navigation of complex sonic spaces. In *NIME '07 Proceedings of the 7th international conference on new interfaces for musical expression*, 379–382.

[23] Lyon, E. (2012) “Image-Based Spatialization,” Proceedings of the International Computer Music Conference, Slovenia.

[24] Oram, D. “Oramics,” http://daphneoram.org/oramarchive/oramics/

[25] Marino, G., Serra, MH., and Raczinski, JM, (1993) “The UPIC System: Origins and Innovations,” Perspectives of New Music 31(1): 258-269.

[26] Penrose, C. “HyperUpic” (1991, no longer publicly available) http://www.music.princeton.edu/~penrose/soft/

[27] UISoftware. MetaSynth. http://www.uisoftware.com/MetaSynth

[28] Izotope, Iris https://www.izotope.com/en/products/effects-instruments/iris/

[29] Virtual ANS https://itunes.apple.com/au/app/virtual-ans/id711384847?mt=8

[30] Neukom, M. and J. Schacher. (2008) “Ambisonics Equivalent Panning”, *Proceedings of the International Computer Music Conference*, Belfast, Ireland.

[31] Smalley, D. 1997. “Spectromorphology: Explaining Sound Shapes.” Organised Sound: Vol. 2, no. 2. Cambridge: Cambridge University Press: 107-126

[32] D. Van Nort. And M. Wanderley. (2006) “The LoM Mapping Toolbox for Max/MSP/Jitter” *Proceedings of the 2006 International Computer Music Conference*, New Orleans, USA

[33] Bevilaqua, F., R. Muller, and N. Schnell (2005). MnM: a Max/MSP Mapping Toolbox. In *Proc. of 2005 Conference on New Interfaces for Musical Expression (NIME)*, pp. 85– 88.