# Sound Shapes and Spatial Texture: Frequency-Space Morphology

**James Stuart**
Edith Cowan University
Australia
sgj@soundfieldstudio.com

## ABSTRACT

The use of Wave Terrain Synthesis as a control mechanism is a governing system that allows the performer to create a complex and coordinated change across an existing complex parametric system. This research has focused largely on the application of Wave Terrain Synthesis for the control of Timbral Spatialisation. Various mappings of the Wave Terrain mechanism are discussed, to highlight some various ways in which frequency-space morphology may be approached with such a model. With the means of smoothly interpolating between various terrain and trajectory states allow the performer to control the evolving nature of sound shapes and spatial texture generated by the model.

## 1. CONTEXTUALISATION

Robert Normandeau describes a technique of ‘spectral diffusion’ and ‘timbre spatialisation’ he explored whilst composing a series of acousmatic compositions: *StrinG-Dberg* (2001-3), *Éden* (2003), *Palindrome* (2006–09), and *Kuppel* (2006–09) [1]. Normandeau describes the process of separating sounds into different regions of spectra using four bandpass filters, and assigning them respectively to a series of speakers across a listening space. It was across this period whilst working on these compositions that Normandeau experimented with different approaches toward timbre spatialisation, altering the centre frequency and width of each filter in *StringDBerg* and *Éden*. In *Palindrome* Normandeau focused on controlling the extent of timbre spatialisation by mixing between the original unprocessed source and a spectrally diffused version.

Stefani and Lauke refer not only to the idea of timbre, but the notion of spatial decomposition and re-composition. Spectral Analysis and Resynthesis techniques have traditionally allowed for this kind of deconstruction/reconstruction process, and relate synonymously with the notion of spectromorphology. We can most certainly suggest that the central processes of ‘spectral diffusion’ and ‘timbre spatialisation’ are the conceptual notions of *spectromorphology* and *spatiomorphology* [2]. Both notions raise questions with regard to the potential scope of spatial gesture and movement possible by ‘spectral diffusion’ and the role and relative importance of *timbre* in clarifying the purpose of such a technique in relation to other approaches of spectral spatialisation. Timbre spatialisation suggests that timbre’s made of different frequencies are spread ‘across’ space, through a process of fragmentation that Normandeau describes. If the movement of spectra evolves in the time domain, the spreading of frequencies create a morphology of sound shapes, and this is evocative of the writing of Iannis Xenakis in his considerations of form, structure, and space, when he writes about the song of cicadas in a summer field, political crowds of dozens or hundreds of thousands of people, and the geometric transformation of shapes [3].

Twenty years earlier, we find a slightly different approach to the distribution of timbre across the listener space. Since 1993, Cort Lippe has written extensively on real-time strategies for controlling timbre in the frequency domain [4]. In 1994 Lippe and Zack Settel began spatialising spectra in two channels using the IRCAM Signal Processing Workstation (ISPW) [5], and in 1999 demonstrated how low-dimensional audio-rate signals could be used as *Spectral Processing Functions* (SPF’s) for generating evolving spectral diffusion [6]. In 2002 Topper, Burtner, and Serafin in a research paper describe a process of *Spatio-Operational Spectral* (SOS) Synthesis controlling localization independently for frequency bands across 8 speaker channels [7]. In 2004 Cort Lippe and Ryan Torchia extended the spectral diffuser software for a quadrophonic speaker arrangement, and allowing for the circular distribution of spectra based on azimuth and distance cues [8]. Christopher Keyes and Daniel Barriero have also published research concerning some 8-channel spectral diffusers [9, 10], and David Kim-Boyle proposed a simpler governing control methodology for spectral spatialisation using the Boids algorithm and integrating the model with vector-based amplitude panning (VBAP) [11, 12]. Whilst all of these research papers present control strategies to govern hundred’s, or potentially thousand’s of real-time parameters, conclusions from this research generally indicate a need for alternative and more effective solutions for governing the spatial distribution of spectra. This research focuses on another governing control system for spectral spatialisation, the use of Wave Terrain Synthesis as a control mechanism, but not only with the intent on controlling the nature of the way different frequency bands are localized, but how timbre is distributed and governed across respective speaker channels.

## 2. WAVE TERRAIN SYNTHESIS

Table lookup procedures for sound synthesis developed not long after the birth of the microcomputer in the mid-1970’s, and it was in 1978 that Rich Gold first coined the term *Wave Terrain Synthesis*. There have been many advantages to the table lookup procedure - it is considered to be computationally efficient as processing requirements are consistently low, and the procedure allows for storing and reloading tables of data. These data sets might consist of either arithmetically generated values, statistical information, and measurement data, provided the data range is adequately sufficient. For example we may use "samples" of real world data, be it audio, video, or any other collection. Extending the dimensionality of the table lookup procedure allows for describing the complex non-linear behavior of electronic components and other complex phenomenology such as the localized perception of sound in a binaural setting using HRTF's.

The idea of traversing data sets is well established in the realms of sound synthesis and computer music. W*avetable lookup* traditionally refers to the repeated scanning of a list of numbers which describe a single cycle of an audio waveform in memory; this was originally used for the generation of digital oscillators. Increases in memory capacity since allowed for the use of larger wavetables which has consequently seen flourish in other techniques in sound synthesis including audio sampling, *Waveshaping Synthesis* (Le Brun, 1979), audio analysis via *FFT* or *wavelet* methodologies and the subsequent resynthesis of this data, and *Convolution Reverb*. The extending of these wavetables by two or more dimensions has also seen other develops in sound synthesis including *Wave Terrain Synthesis* (Gold, 1979), *Graphical Synthesis*, *Head-Related Transfer Functions* (HRTF's) for the binaural treatment of existing sounds, and *Multi-Impulse Response* techniques. In 1995 the author suggests that Wave Terrain Synthesis may be a possible method of controlling sound spatialisation.

## 3. MAPPING STRATEGIES

Timbral Spatialisation is a technique that requires the control of potentially thousands of separate frequency bins, and for larger speaker arrays the number of relevant parameters increases. Such a system benefits from a control system that operates at audio-rate, is time-synchronised, and allows the control of multiple parameter sets that interact across a multi-channel speaker array. This sense of interaction between speaker channels is important if the results are to generate a morphology of "sound shapes" – for example a scene rotation would require the parameters of one speaker to be transferable from speaker to speaker, and henceforth. More complex sound shapes involve some considerably more non-linear transitions. The rationale of a Wave Terrain Synthesis control mechanism is such that a trajectory traversing a terrain may be time-division multiplexed or phase shifted in order to extract several interrelated contour maps from the one terrain surface, and these then might be used for controlling the state of a parameter that is connected across the entire system.

Evaluation of each system depended largely on the performers ability to control the predictable outcome of timbre, sound shapes, spatial gesture, and various psychoacoustic concerns: source location, source width, spatial coherence, and immersiveness, whilst leaving it feasible for the performers to inject a certain amount of the unpredictable.

Several mapping strategies were tested. The first strategy commonly uses a circular trajectory curve, shown in Figure 1, where the angle of the read pointer traversing the circle corresponds with azimuth cues, the distance from the origin or the center of the circle corresponds with distance cues, and the value of the terrain curve, or the height $z$, corresponds with the associated frequency bin. Whilst the height parameter $z$, may be represented visually in various ways, we can say that low values directly correlate with low frequency bins and similarly high values with high frequency bins. As the lookup procedure is performed at audio-rate, allocation of azimuth to various frequency bins is feasible, however there is the potential for the same frequency bin to exist in multiple locations. Some outcomes consequently require a "folding" down of data, such as a flat terrain curve shown in Figure 2 where the lookup process would generate many duplicates of the same frequency at various points of azimuth implying that we have one single frequency band diffused in many hundreds of positions. Apart from the unfeasibility of this with a single FFT for each respective speaker output, the perceived effect should be that one frequency is diffused across all output speaker channels available. Folding down this data to the appropriate speaker configuration used is necessary. There can also be azimuth and distance conflicts for frequency bins when trajectories cross over within a single FFT frame such as the Archimedes Spiral shown in Figure 3. However, this may also be resolved by unwrapping the angles of azimuth generated.

This mapping approach requires that the relevant spatialisation for specific frequency bins be decoded. Values are multiplexed in terms of their corresponding azimuth in the sound field, and then these lookup tables generated are interpolated to find a generalized SPF for each respective speaker channel.

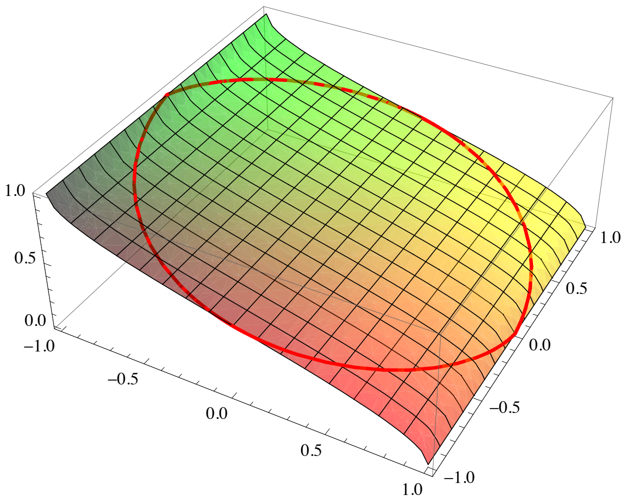

**Figure 1**. A trigonometric terrain curve traversed by a circular trajectory curve

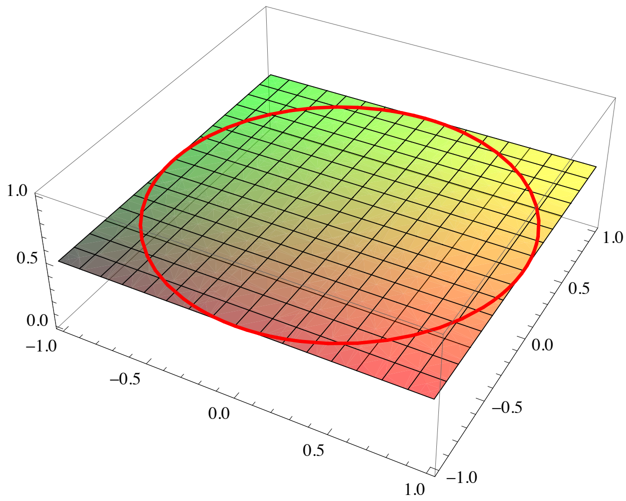

**Figure 2**. A flat terrain curve traversed by a circular trajectory curve

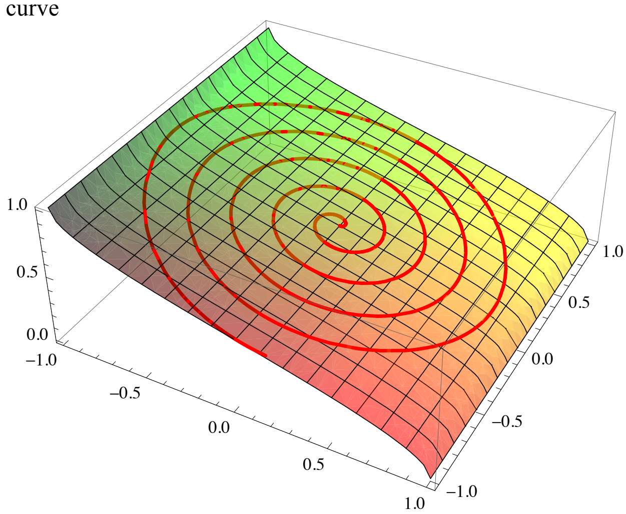

**Figure 3**. A trigonometric terrain curve traversed by the Archimedes Spiral

The second model uses the contour *z* generated by Wave Terrain Synthesis as the SPF. This SPF is phase shifted several times and subsequently windowed for different speaker outputs. As opposed to the first mapping strategy, this approach has some different benefits, for example when the terrain is flat, such as in Figure 2, the SPF is also flat, hence all frequencies are passed at equal amplitude through all speakers. When the terrain is tilted, and curvature is applied as found in Figure 1, there is a perceived effect of high frequency moving to one side, and low frequency to the other. Whilst the exact relationship between the terrain function and what is heard correlates superficially, they are in fact abstracted from one another in the sense that the "direction" of the trajectory as it passes across the terrain function determines the order of subsequent frequency bins, and hence any phase shift applied is also subject to this clockwise or anti-clockwise movement. The relationship between the terrain and the resulting spatialisation is further abstracted for trajectory functions that are more chaotic and stochastic in nature, however the audio-rate topographical movement contributes in a significant way in influencing the evolution of sound shapes generated across the soundfield.

Some of the advantages of this second approach are that it is the most computationally efficient, generally has the "smoothest" response in terms of the SPF's generated, and is the most responsive in terms of audio-rate morphology. The system is also generally intuitive to control. For linear terrains and trajectories, such as in Figure 4 the user is able to generate a series of bandpass filters as seen in Figure 5, that sum perceptively to create a linear response curve across the frequency range.

Whilst this method is effective in creating evolving timbral shapes, the perceived change does not always correlate exactly with the wave terrain pictured in the interface window. This is particularly the case at the extremity of the spectrum where the highest frequency bands cross over to adjacent low frequency bands of the following FFT frame as can be seen in Figure 5.

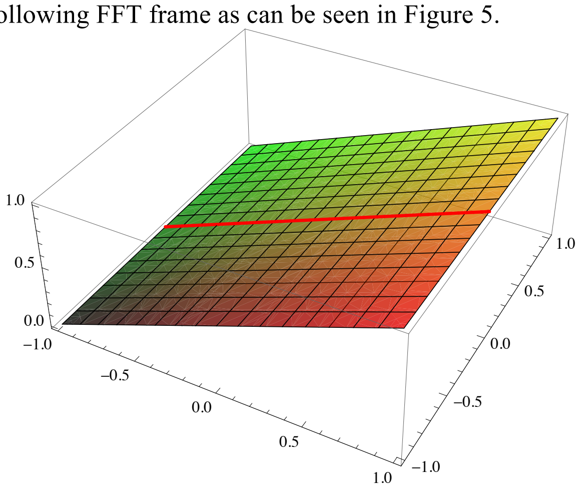

**Figure 4**. A linear terrain curve traversed by a linear trajectory and the resulting temporal curve

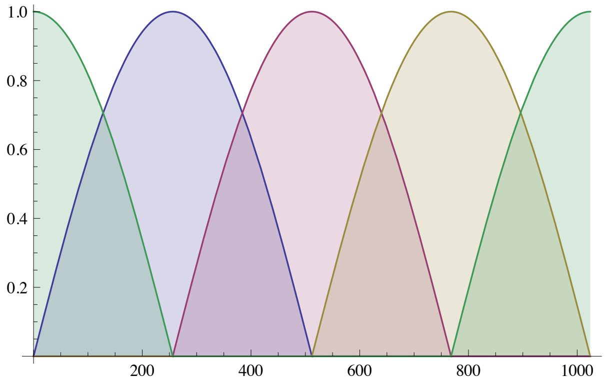

**Figure 5**. Color coded SPF functions assigned to 4 different loudspeakers

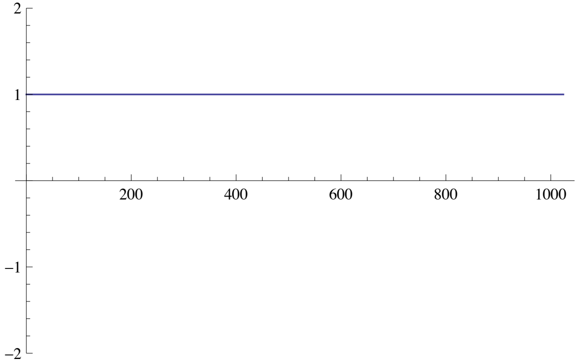

**Figure 6**. According to equal power panning and the phantom image, separated frequency bands in Figure 5 perceptively sum creating a linear response curve[1]

For more nonlinear terrain and trajectory pairs the resulting frequency response can also be considerably more nonlinear. Figure 7 shows a complex series of SPF's generated, and the resulting combined frequency response. In this way the nonlinearity and morphology of Wave Terrain Synthesis can be useful for generating a range of timbres and sound shapes.

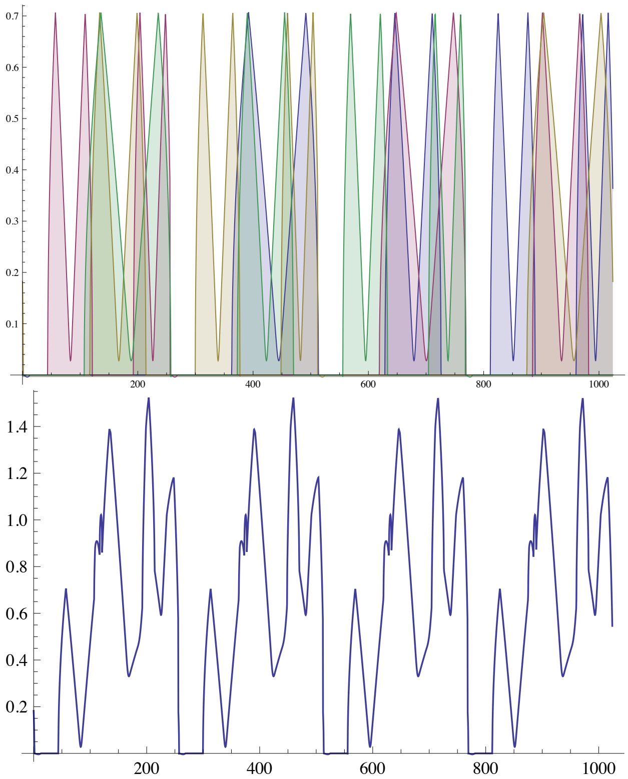

**Figure 7**. A nonlinear series of SPF's and their sum showing the perceived loudness of separate frequency bands

The third mapping strategy was aimed at bridging between the first and second methodologies. Like the first strategy, this method accounts for differences in color as representative of different frequency bins. By time-division multiplexing of the trajectory into N pieces, where N represents the number of output speaker channels, a histogram is calculated on the relative amounts of different colors analysed. This allowed for a more intuitive relationship between the visual color displayed, and the resulting spectrum generated. The method allows the terrain or trajectory to be rotated, resulting in a coordinated frame rotation around the speaker array.

## 4. NAVIGATING A TRAJECTORY SPACE

One of the benefits of Wave Terrain Synthesis is the wide-variety of audio-rate trajectory curves, but issues still exist in regards to how they are semantically defined, and secondly how they may be navigated in one connected parameter space. In other words, is it possible to navigate trajectories in such a way that they all exist along one multi-dimensional continuum? The author suggests they may be defined by two primary factors: firstly their level of *periodicity*, and secondly by what topologists refer to as their *homotopy equivalence* and *homotopy class*. Navigating this as a continuum means that the way different frequency bins are redistributed can be subtlety or drastically altered morphologically in one connected spatial gesture.

Besides interpolating between 2 points on a line, 4 points on a 2D plane, and 8 points in a 3D volume, this research was looking for a more compact way of navigating many trajectory curves across a 2D plane. One alternative for interpolating parameters using a 2D interface is the *nodes* object in MaxMSP. This object manages intensity levels for a number of different nodes that vary according to where a point lands in relation to a bounding circle around each spatial node as seen in Figure 8. Unfortunately this means that for nodes that intersect, we have several signals summed at varying weights, meaning that we cannot maintain equal intensity interpolation across this multichannel and arbitrarily designated array.

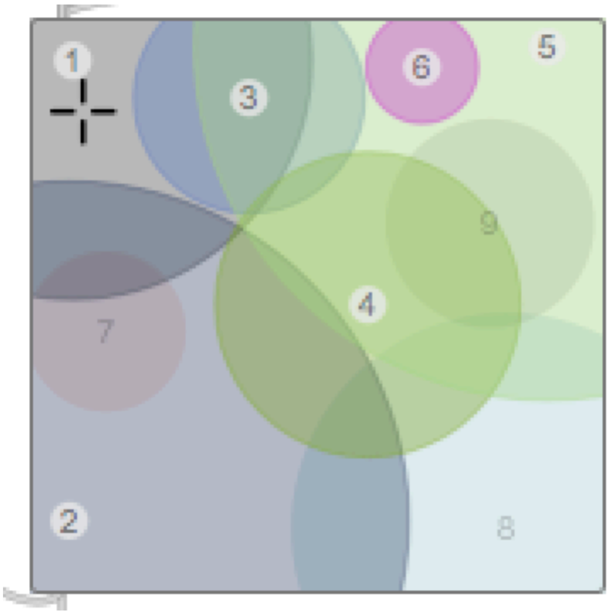

**Figure 8**. The *nodes* object in MaxMSP allowing for interpolating across nine nodes positioned across a 2D plane

A more flexible solution for interpolating different audio sources is the distance-based amplitude panning technique, 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. Each input source is arbitrarily positioned within this virtual navigable

---

[1] This is stated without taking into consideration other variable factors such as speaker frequency response, room acoustics, the Haas effect, air absorption, and the non-linearities of the Pinna, head and shoulders.

space. The advantage of DBAP over nodes is the ability for DBAP to adapt appropriate loudness curves where different sound sources might normally intersect as managed by Equation 1. This also ensures that loudness roll-off curves are extended for where sources do not intersect.

$$I = \sum_{i=1}^{N} v_i^{2} = 1 \qquad (1)$$

Where v refers to the amplitude of each source, and i refers to the source number, N being the maximum. This is a slight point of difference in the use of the equation for DBAP panning where i is normally used to denote the speaker number. Navigation across many trajectories are possible as shown in Figure 9.

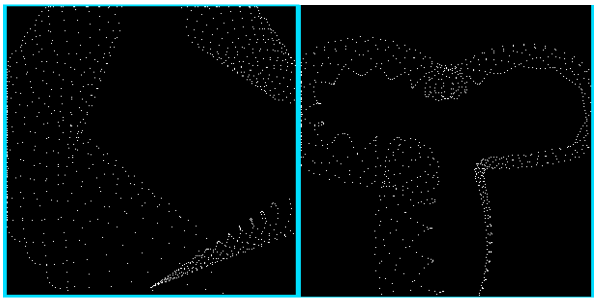

**Figure 9**. DBAP panning between a circle, square, rose curve, Archimedes Spiral, Henon attractor

The evolution of the trajectory function is central to the performers ability to control the morphology of sound shapes generated by Timbral Spatialisation. Controlling the relevant evolution of these curves gives the performer the ability to control the level of periodicity, aperiodicity, chaos, randomness, as well as the morphology between these various states.

Other morphologies are possible such as the geometric transformation of a trajectory curve generated. Besides standard methods of affine transformation such as translation, scale, and rotation, there has been interest in exploring trajectory curves that are distorted in some way through various manipulating means: smoothing functions *rampsmooth~*, foldover and wrapping with *pong~*, and bit reduction with *degrade~*. Other potential distortions can be created through the use of phase distortion algorithms such as *kink~*. Some of these can be seen in Figure 10. Feedback and crosstalk can also be effective ways of introducing distortion in the evolution of the trajectory as shown in Figure 11.

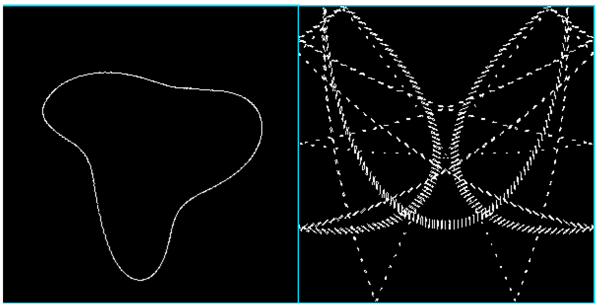

**Figure 10**. A trajectory curve with a smoothing function applied to it, and secondly a trajectory with foldover applied

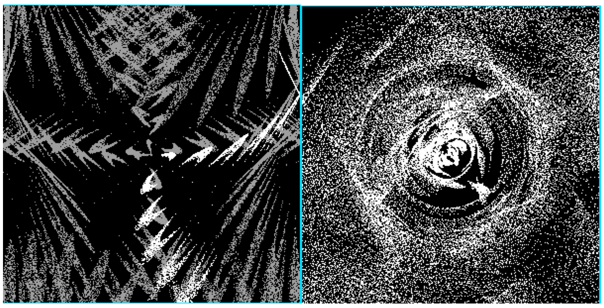

**Figure 11**. Trajectory curve generated with feedback and crosstalk applied

## 5. SPATIAL TEXTURE AND SPECTROMORPHOLOGY

The success of morphological systems depends on the smooth transition from its current state to the next. This model uses the jitter library in MaxMSP for managing the multidimensional data sets that are used for Timbral Spatialisation. As a consequence of using Jitter frames, a stepped-like artifact is evident at the frame-frame transition. This is pictured in Figure 12 below. This stepped-like artifact is not evident however in audio-rate transitions.

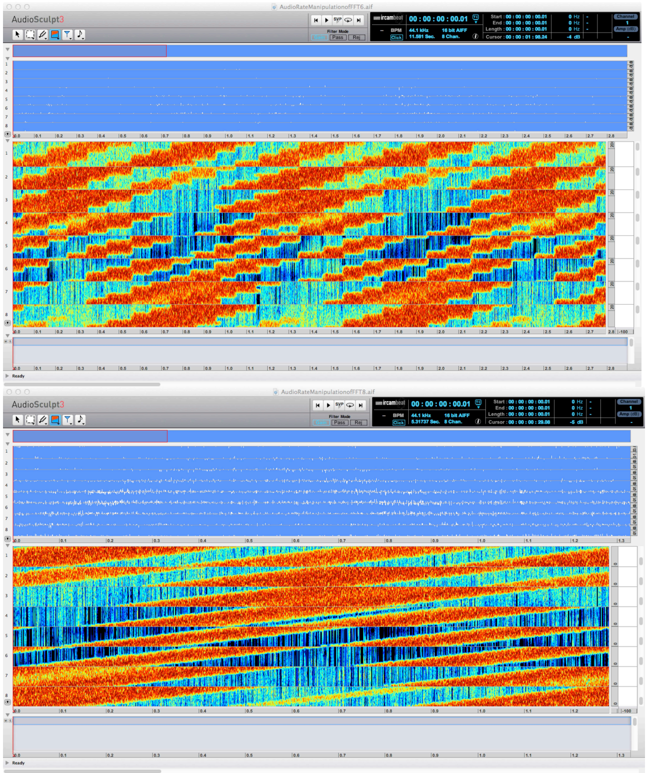

**Figure 12**. The stepped-like result of Timbral Spatialisation when a terrain is rotated, and the equivalent when the trajectory is rotated instead

It is possible to smooth over these stepped-like transitions by buffering a number of jitter frames at a time, and interpolating across these frames. The author created a circular buffer for dynamical terrain maps allowing the extraction of continuous terrain data at audio rate.

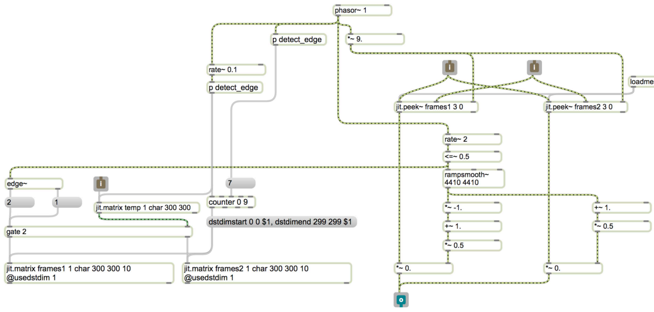

**Figure 13**. The video circular buffer generated in MaxMSP

The complex interaction of terrain and trajectory structures often create a spatial texture and morphology that is both coordinated yet in some cases erratic if based on the evolution of chaotic and random trajectories. One can see in Figure 14 the movement of different spectral bands across an 8 loudspeaker system, tracing the various nonlinear spatial movements generated.

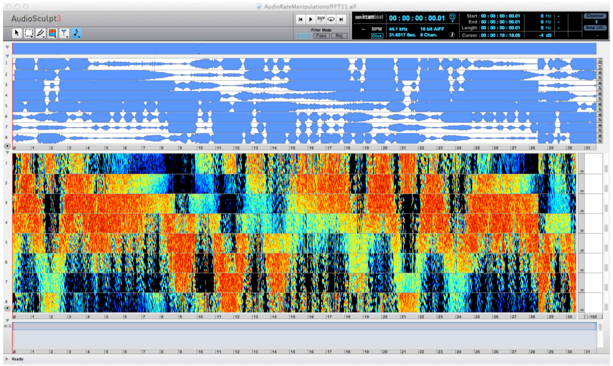

**Figure 14**. The coordinated spectral movement and spatial texture generated using a chaotic trajectory that is geometrically modulated over a linear terrain function

Intuiting frequency-space morphology largely depends on the visualisation provided for the performer. The Timbral Spatialiser software provides an interface for visualising the terrain and trajectory movements generated.

## 6. CONCLUSIONS

The intersection of both morphology in the terrain curve and independent modulation and morphology of the trajectory curve allow for a complex interaction of two different temporal structures, resulting in some unexpected sound shapes. The ease of control through the use of an existing and adaptable multidimensional lookup strategy such as Wave Terrain Synthesis allows for a wide variety of different topographies. The ability to navigate trajectory curves of different type using some of the principles of DBAP allows for the user to navigate periodic, quasi-periodic, and aperiodic trajectories in one connected space. Trajectories may also be categorized according to their homotopy equivalence and class.

Morphology of audio-rate trajectories allows for smooth transitions in Timbral Spatialisation from frame to frame, but changes in the terrain curve result in a step-like artifact when reading data from this multi-signal at audio-rate. Jitter frame interpolation was necessary in order to generate continuous audio-rate signals.

Sound shape generated, as well as the perceived changes in timbre are various, depending on the mapping strategy used. A more intuitive approach mapping color to frequency bins allow the performer to quickly determine how and where frequencies are allocated across a soundfield.

Further research is required in developing lower-level software for performing many of these processes, and applying physical gesture as a means of performing Timbral Spatialisation.

## 7. REFERENCES

[1] R. Normandeau. “Timbre spatialisation: The Medium is the Space.” Organised Sound. 2009.

[2] D. Smalley. “Spectro-morphology and Structuring Processes.” In EMMERSON, S. *The Language of Electroacoustic Music,* London, Macmillan Press Ltd, 1986, pp. 61-93.

[3] I. Xenakis. *Formalized music*. (Rev. ed). New York: Pendragon Press, 1992.

[4] C. Lippe. “FFT-based Resynthesis for the Real-Time Transformation of Timbre,” 10th Italian Colloquium on Computer Music, Milan, 1993, pp. 214-219.

[5] Z. Settel and C. Lippe. “Real-time Timbral Transformation: FFT-based Resynthesis.” *Contemporary Music Review*, Harwood Academic Publishers, 1994.

[6] C. Lippe and Z. Settel. “Low Dimensional Audio Rate Control of FFT-Based Processing”, *Proceedings of the 7th Biennial Symposium on the Arts and Technology*, Connecticut College, 1999.

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

[8] R. Torchia and C. Lippe. “Techniques for Multi-Channel Real-Time Spatial Distribution Using Frequency-Domain Processing.” *Proceedings of the 2004 Conference on New Interfaces for Musical Expression (NIME04)*, Hamamatsu, Japan, 2004.

[9] C. Keyes. “Three Approaches to the Dynamic Multichannel Spatialization of Stereo Signals.” *Proceedings of the 2004 International Computer Music Conference. San Francisco*: ICMA, 2004, pp. 325–9.

[10] D. Barreiro. “Considerations on the Handling of Space in Multichannel Electroacoustic Works”, *Organised Sound,* 15(03), 2010, pp. 290–296.

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

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

[13] T. Lossius, P. Baltazar and T. de la Hogue. “DBAP – Distance-Based Amplitude Panning” *Proceedings of the International Computer Music Conference*, Montreal, 2009