Competing Attractions, Orbital Decay and the Music of the Spheres: Force–based Relational Dynamics for Organizing Space and Timbre in Performance Using Physical Models

Richard Graham
Sensory Computation/Experimental Narrative Environments Lab
Stevens Institute of Technology, USA
rgraham1@stevens.edu
http://rickygraham.com

Brian Bridges
School of Creative Arts and Technologies
Ulster University, UK
bd.bridges@ulster.ac.uk
http://brianbridges.net

Proceedings of Korean Electro-Acoustic Music Society's 2016 Annual Conference (KEAMSAC2016)
Seoul, Korea, 15-16 October 2016

This paper describes the mapping of embodied metaphors found within physical systems to the spatial organization of voices and timbral processes. The intention of such an approach is to enhance the clarity and richness of connections between performance gestures and sonic structures. Previous system iterations have presented mappings informed by ecological–embodied metaphors of dynamic forces as a means to bridge cross-domain performance events across physical and figurative planes. The first iteration sought to reify gravitationally based tonal pitch space models by mapping melodic syntax computations (such as attraction, tension, and inertia analogues) to in–kind parameters of a flocking algorithm as a method of dynamic audio spatialization. Given the embodied physical bases implied by musical models proposed by Lerdahl and Smalley, we present a system that further explores the ecological bases of musical abstraction through the lens of force–based mapping models for spatial audio and timbral processing. The present iteration of the system utilizes a physics engine in a game development environment as a base for a practice–led exploration of mappings encompassing a wider variety of force–relational dynamics (derived from instrumental note–events) as applied to the evolution of spatial and timbral gestures. A particular focus is the treatment of energy-motion trajectories within the system's mapping. While spatialization and diffusion is an obvious output modality for such a mapping, practice–led explorations of these embodied concepts, as facilitated by this system, may also inform a relational model of timbral connections.

The impetus for our live performance system design is largely based on the extension of an electric guitar practice through multichannel audio processing per instrumental register. Our chosen processes focus on the extraction of instrumental performance data, which are utilized to forge gestural narratives between recognizable physical or figurative performance events and timbral and spatial audio signal processing. This paper outlines the motivational basis, theoretical underpinnings and current innovations of this project.

Metaphors and Mapping Strategies for Data Rich Performance Environments

The initial motivation for this project was largely based on the desire to extend the sonic characteristics of a conventional electric guitar beyond its conventional design. Previous iterations of our performance system have drawn influence from a series of fields, including embodied cognition and human-computer interaction. These systems incorporated metaphorical mappings based on familiar physical gestures to provide more intuitive access points for the listener and to allow the performer to manage complex sonic materials during a real-time instrumental performance. Emergent mapping strategies have included the a melodic model for spatialization (Graham and Bridges 2014a, 2014b), and more recently, a force-based model based on temporal and frequency–based continuity and coherence to drive spatial and timbral morphologies (Graham and Bridges 2015). This paper presents a series of recent system developments, including a physics and collision detection environment, driven by real-time physical and figurative performance gestures. These recent developments provide a useful framework for a composer/improviser to explore the notion of order to disorder (or a loose-tight continuum) in the composition of musical materials.

Emerging Live Performance System Designs for Multichannel Guitar

The basis of our new system is the Nu Series Modular Active Pickup by Cycfi Research in the Philippines. This low-power system boasts a hacker-friendly design, low impedance coils, and a low-noise preamplifier for each module (DeGuzman 2016). These innovations in multichannel audio pickup design are ideal for the next stage of development for our system, particularly given the wide frequency response of each pickup module and the improved crosstalk performance through permalloy ring shields and discrete preamp design. The guitar passes the multiple individuated audio signals using a 19-pin LEMO connector cable, which also carries power to the pickup system. Overall, the system carries nine channels of audio: eight individual audio channels per register and one channel carrying a sum of the existing mono magnetic pickups positioned at the neck and bridge of the guitar. As with our previous iterations, the audio feeds may then be sent to any analog or digital audio processing device using our available breakout board designs (Graham and Harding 2015).

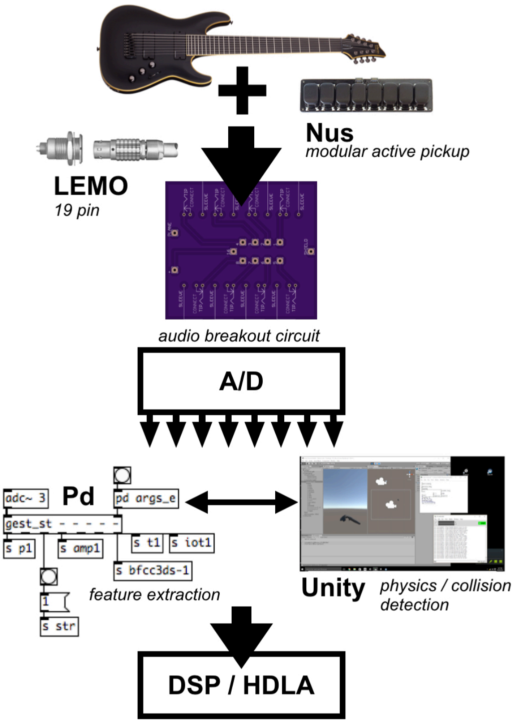


Figure 1. Current Multichannel Guitar System Design incorporating Nu Modular, LEMO, Pure Data, and Unity

This version of our system uses a basic circuit design by Graham (2016) designed for compatibility with analog synthesizer modules. Audio channels are then fed to a multichannel audio interface to capture pitch, amplitude, and spectral flux data per instrumental string register. Data may then be organized into higher-level tonal abstractions, scaled, and mapped to establish new gestural narratives between the performer, performance system, and resulting sonic structures. In previous iterations, the systems design primarily utilized the Pure Data (Pd) visual programming environment for spatialization and timbral processing. This iteration will seek to develop a bi-directional and multidimensional relationship between the feature extraction and effects processing patches programmed in Pd and a more sophisticated physics and collision detection system found within the Unity game development platform.

An Introduction to VESPERS at SCENE Lab

In 2016, Graham co-founded a research facility at Stevens Institute of Technology (USA) with Messrs. Cluett, Manzione, and O'Brien named the Sensory Computation/Experimental Narrative Environments (SCENE) Lab. The goal of SCENE Lab is to create hybrid software and hardware systems for the development and presentation of immersive virtual spaces. SCENE Lab houses the Virtual Experience, Sonic Projection, Extended Reality System (VESPERS) comprising of a 24.2 high-density loudspeaker array. The system can function as three individual eight-channel loudspeaker systems, including a first-order ambisonics (FOA) cube and a higher-order ambisonics (HOA) ring. This first-order 3D loudspeaker system permits musical experimentation using performative models that exploit height or elevation cues within virtual reality environments.

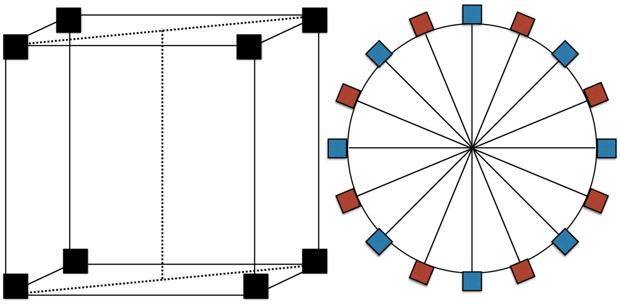


Figure 2. VESPERS: SCENE Lab's 24.2 high-density loudspeaker array including an FOA 3D 8-channel cube and HOA 2D 16-channel ring.

Previous iterations of our live performance system enabled the composer/improviser to superimpose or map abstract musical structures or image schemas onto a physical performance space. An obvious progression to accommodate more detailed musical models is to increase the resolution and dimensions of our loudspeaker array accordingly. VESPERS also integrates the HTC Vive (Valve 2016) virtual reality headset system, which provides highly sophisticated motion capture through two base stations emitting pulsed IR lasers. This motion capture system provides high precision tracking for user head movements and other bodily gestures through compatible hand controllers. The multidimensional, nonlinear tools provided by the VESPERS environment provide an ideal base for the development of more layered musical models for real-time instrumental performance systems.

Exploring a Tonal Model in 3D Space

Previous iterations of our performance system utilized melodic syntax computations to inform the steering behaviours of a flocking algorithm for spatialization.

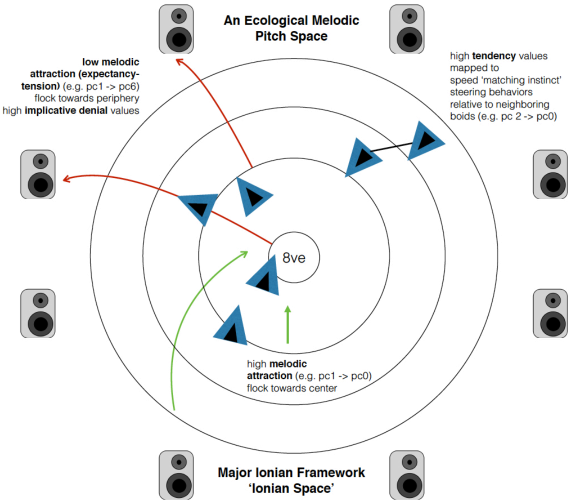


Figure 3. Dynamic Tonal-Spatial Mappings from Graham and Bridges (2015)

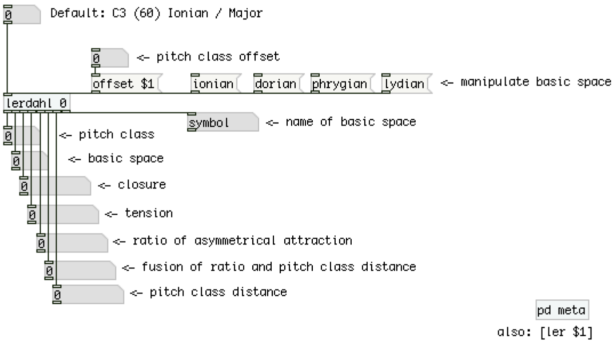


Figure 4. Introducing the cross-platform [lerdahl] external for Pure Data written by Graham (2016)

Graham has since developed an external for the Pure Data (Pd) visual programming language. Version 1 receives any MIDI input range and outputs values for pitch class, basic space, closure, tension, ratios of asymmetrical attraction and pitch class distance. The user may change the configuration of the basic space to accommodate any of the seven modes of the major scale. The user may also offset pitch class zero if they want zero to be something other than MIDI note C3 or C4. Version 2 will have list outputs and more useful construction arguments. These discrete values have proven to be a useful representation of real-time figurative gestures. Given the ability to support elevation cues in VESPERS, this paper presents mapping strategies extending our tonal-spatial model to exploit pitch height schemas within SCENE Lab's 3D FOA cube.

Unity and Pure Data: Physics informed Timbral and Spatial Processing

Prior to our current system's design, we focused on metaphors found in flocking, gravity, and tonality. These mapping strategies may be further developed in an interactive environment using the physics and collision detection components within the gaming development environment, Unity.

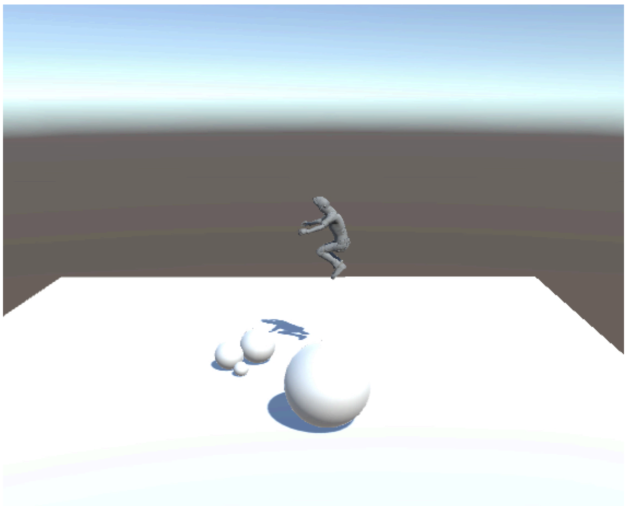


Figure 5. Third-person controller interacting with objects with Unity's physics engaged, with variable mass, material and drag. Gravity may also be manipulated in real-time.

Unity provides an ideal canvas for multi-dimensional and user-directed interaction. Spatialization may now be driven by physical gestures captured using the highly sophisticated Lighthouse motion tracking system (Valve 2016). Early tests using the Vive hand-controllers allowed any individual to morph the spatiotemporal structure with minimal latency. A performer may morph timbral and spatial parameters through regular or familiar instrumental hand gestures (Graham and Cluett 2016).

However, this paper will focus on the extraction, scaling, and mapping of discrete digital audio signals from our instrumental sound source as opposed to physical performance gestures explored in the aforementioned paper by Graham and Cluett. The goal of future work will be to combine both physical and figurative systems within a virtual scene to create an embodied and interactive musical experience.

Navigating A Virtual Performance Space

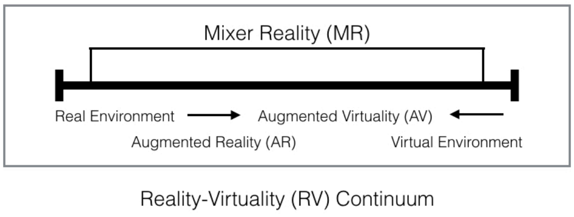


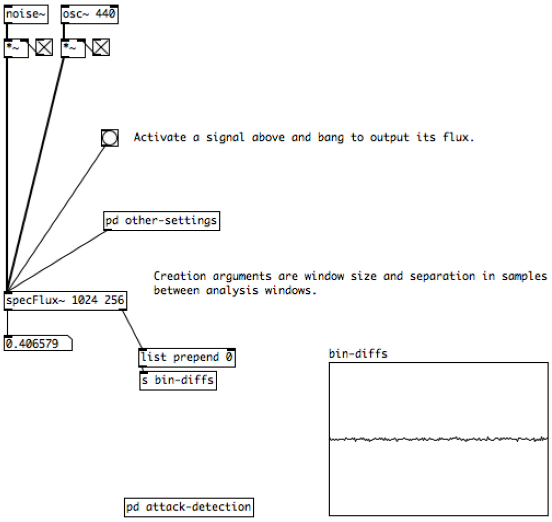
Figure 6. The Reality-Virtuality Continuum based on Milgram et al. (1994)

Milgram's "Reality-Virtuality Continuum" (1994 p. 283) presents a compelling series of questions regarding the integration of a musical instrument based performance system within a virtual reality environment: where does this combination fall on the continuum and what are the implications for the composer/improviser? Are we dealing with the augmentation of the listener's real environment through a virtual environment? The performer's experience is potentially and solely based within a virtual environment. As we consider this question, one may consider Smalley's ideas surrounding spectromorphology (1997) and space forms (2007) to guide the development of new mappings between real and virtual worlds. There are a number of highly suitable definitions directly applicable within our developing virtual environment and may inform our practical examples.

Tonal Pitch Space	Subdivision of spectral space into intervallic steps
Spectral Space	The impression of space and spaciousness produced by the range of audible frequencies
Microphone Space	Intimacy of the image is magnified

Table 1. Useful definitions from Smalley (2007 p. 55) for establishing new mappings within our virtual performance space

In previous system iterations, abstract tonal pitch space data was used to drive spatialization parameters. In this iteration, the performer is able to physically interact with sound objects within a virtual space using instrumental gestures. One can relate this to Smalley's notion of agential space, "a space articulated by human (inter)action with objects, surfaces, substances, and built structures, etc. Combines with utterance space to create enacted space" (Smalley 2007 p. 55). The following practical examples illustrate how one can extract and exploit the "energy–motion trajectory" and "spectromorphological expectation" (Smalley 1997) of each note-event performed by the improviser/composer using time-frequency analysis tools in Pd.



Activate a signal above and bang to output its flux.

Creation arguments are window size and separation in samples between analysis windows.

Figure 7. William Brent's TimbreID library for Pure Data (Pd) provides useful computations for low-level timbral features

The mappings use a new implementation of Brent's bark-scale based timbral analysis Pd externals (Brent 2016). These externals are especially useful for performance systems design aimed at extracting and reifying the various energy profiles associated with timbral structures found in each individual audio input signal. Bark spectrum versions of lower level features, such as spectral flatness, centroid, and flux, are more useful to a designer focusing on perceivable change in musical structures, particularly changes representative of a performer's energy-motion output modalities.

Supporting Video Examples

The following section presents a series of practical mapping strategies exploiting tonal, spectral, and microphone space outlined above.

Driving Tonal-Spatial Mappings Using Physics in Unity

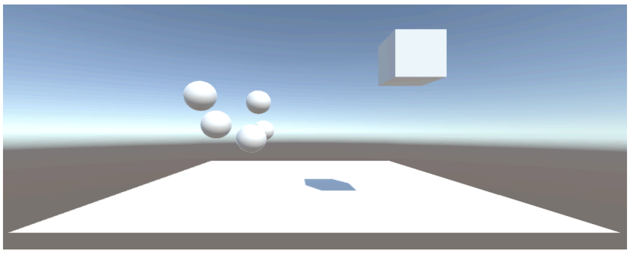
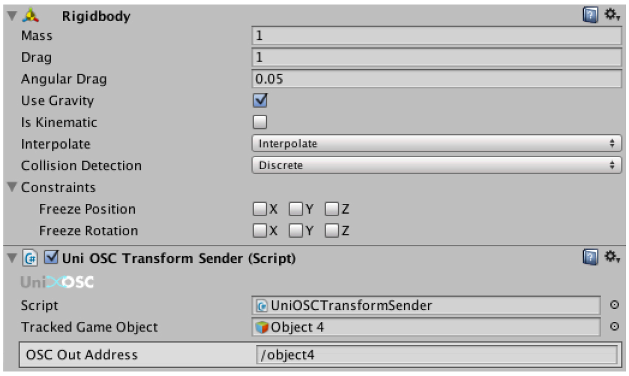


Figure 8. Individual audio channels per register are mapped to 3D objects in Unity

In this example[1], we explore the notion of pitch height, first by simply mapping our basic space cone to basic z-axis locations, from the ground plane to the top plane or the FOA cube. Individual audio channels are mapped to sound objects within the virtual space. Their position on the vertical plane is determined by the basic space value of the note performed by the instrumentalist. In the second example[2], we apply the computations from the melodic syntax external for Pd to control flocking behaviors within Unity in 3D space. The incorporation of elevation cues using our FOA cube is incredibly effective in reinforcing the superimposed tonal-spatial model within the virtual reality environment, particularly given the ability to assign mass, drag, and gravity based interactions using physics within Unity.

Using Spectral Flux to Determine Rigidbody Attributes

In this example[3], spectral flux values are scaled and mapped to determine the mass and drag of each individual sound object within the virtual performance environment. Data is captured and sent over open-sound control (OSC) to control Unity parameters using a C# script.



Rigidbody
Mass: 1
Drag: 1
Angular Drag: 0.05
Use Gravity: ☑
Is Kinematic: ☐
Interpolate: Interpolate
Collision Detection: Discrete
Constraints
Freeze Position: ☐X ☐Y ☐Z
Freeze Rotation: ☐X ☐Y ☐Z

☑ Uni OSC Transform Sender (Script)
Script: UniOSCTransformSender
Tracked Game Object: Object 4
OSC Out Address: /object4

Figure 9. Instrumental features extracted in Pd may be scaled and mapped to control Rigidbody attributes in Unity using Open Sound Control (OSC)

Using Attack Detection, Amplitude Envelope, and Spectral Flux to Determine Energy-Motion Profiles

In our previous design (2015), we used note attack detection in combination with a basic envelope follower to determine gravitational attraction or centricity relative to a central point within the performance space. This was a limited albeit somewhat effective mapping strategy from the point of view that it produced a perceptually clear and distinct musical effect relative to a performer's onset energy when performing any note on any register of the multichannel guitar system. The continuant portion of the event was then tracked to allow the system to determine the ideal termination phase, at which point a new series of signal processes may be applied to the real-time instrumental audio input source.

Attack or Onset	Sustain or Continuant	Release or Termination
Emergence	Prolongation	Disappearance
Departure	Passage	Arrival
Launching	Plane	Goal

Table 2. Smalley's basic energy-motion profiles and theorized embodied associations by Graham and Bridges (2015)

---

[1] See Example 1 in KEAMSAC Playlist: http://bit.ly/2dtUa74

[2] See Example 2 in KEAMSAC Playlist: http://bit.ly/2dtUa74

[3] See Example 3 in KEAMSAC Playlist: http://bit.ly/2dtUa74

In this iteration[4], the size and trajectory of sound objects will continue to be determined by the amplitude envelope of their respective instrumental register. Our propulsion idea may be further developed within a physics-sensitive space relative to peak amplitude at note onset, the amplitude envelope at continuant and termination phases, and the spectral flux weighted using the bark-frequency scale.

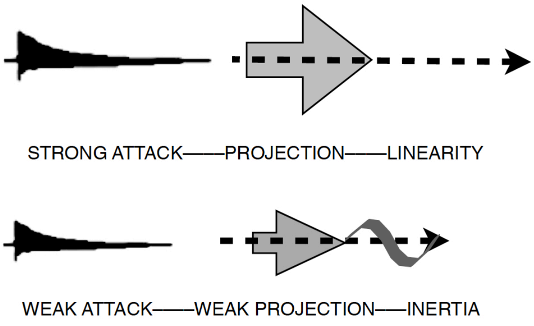


Figure 10. Attack-Projection-Linearity from Graham and Bridges (2015)

In our 2015 iteration, our attack-projection-linearity mapping juxtaposed strong note attacks against weak note attacks, with strong note attacks resulting in a strong linear motion across the quadrants of the performance space and weak attacks assuming a random spatial trajectory. In this previous iteration, strong projections also morphed timbral events between monophonic and multichannel distortion or gain structures. The monophonic distortion structures presented a much more integrated scene, whereas the multichannel distortion presented a much more segregated scene. This notion was largely based on the previously discussed order to disorder (or a loose-tight) continuum. In this iteration, we extend these previous mappings by exploring transitions between Pitch Synchronous OverLap-Add (PSOLA) based granular synthesis and jitter-based/randomized granular clouds using Graham's [gramulator~] abstraction for Pd.

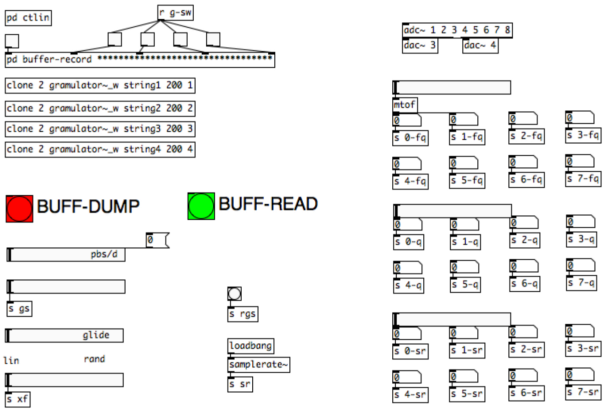


Figure 11. Graham's [granulator~] system for Pd (2016)

This granulation abstraction permits the composer to morph seamlessly between linear playback of a stream of time-stretched grains and randomly concatenated streams of grains in real-time. Streams of grains may be confined to a static spatial location within the performance space or they may be panned dynamically throughout he space, relative to the aforementioned tonal-spatial model. Streams of grains may also assume their own unique gain structure per stream (macro) or per grain (micro).

Using Spectral Flux to Determine Timbral Characteristics of Each Register

In this example[5], spectral space may be further explored using continuous values representative of energy per note-event mapped to all-pass filter frequency controls. This audio mixer plugin for Unity was created using Pd and Enzien Audio's online platform, "Heavy" (2016).

Exploring Microphone Space Intimacy in Relation to Musical Dynamics

This relationship between musical dynamics of a real-time instrumentalist and three-dimensional scene has an immediate impact on the perception of the immersive space and intimacy of the overall auditory scene. In this example[6], we utilize the ambisonic B-Format *zoom* function authored by David Malham and ported to Pd and Max/MSP externals by Matthew Paradis (2002). This function allows a user to zoom in on a defined point of a first-order soundfield positioned within the virtual space.

---

[4] See Example 4 in KEAMSAC Playlist: http://bit.ly/2dtUa74

[5] See Example 5 in KEAMSAC Playlist: http://bit.ly/2dtUa74

[6] See Example 6 in KEAMSAC Playlist: http://bit.ly/2dtUa74

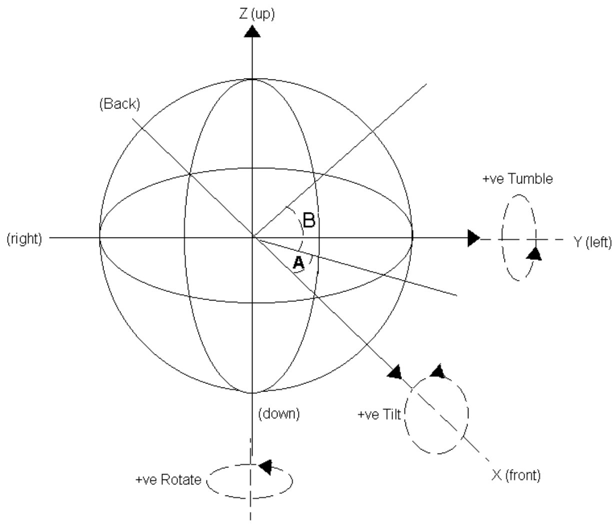


Figure 12. Soundfield manipulations from Malham (2008)

Malham's B-Format zoom implementation allows a user to choose a specific point within a three-dimensional soundfield and zoom in or away from that point. Aside from making sounds in the defined direction appear louder, this technique will also reduce the angular spread of sounds in the opposite direction. This is an effective tool for sculpting musical dynamics in general and more specifically for exploiting the aforementioned *microphone space* where the intimacy of the image is increased through magnification.

Conclusion: Moving Towards a Pitch-Timbre-Space Model for Virtual Reality Based Music Performance Environments

We have presented a developing interactive model for extended instrumental practice. Our system extracts note events, pitch, amplitude, and spectral data per register and utilizes this data to inform timbral and spatial processes within our developing physics-driven scene in Unity. This system will permit the user to engage with primitive and higher-level, more abstract musical ideas in an interactive environment governed by their musical choices and bodily movements. Future work will explore the development of a more detailed timbral space whereby sections of the abstracted virtual space may be divided detailed axes, sectors, or quadrants representing more complex spatiotemporal structures. Sound objects may also directly interact with one another based on the performer's physical position within the space. This approach would ensure more fluid performance interaction, whereby the performer is freed from the confines of the physical world, and is given free reign over spatial location behaviors based on commonly understood gravitational interactions and familiar everyday physical gestures.

Acknowledgments: We would like to extend our thanks to the faculty at SCENE Lab and Dr. Kelland Thomas at Stevens Institute of Technology for the continued support on this project.

References

Graham, R. (2012) Expansion of Electronic Guitar Performance Practice through the Development of Interactive Digital Music Systems. Ph.D. Thesis. Londonderry: Ulster University.

Graham, R. and Bridges, B. (2014a). Strategies for Spatial Music Performance: Practicalities and Aesthetics for Responsive Mappings" in Divergence Press. Issue 3. Online.

Graham, R. and Bridges, B. (2014b). Gesture and Embodied Metaphor in Spatial Music Performance Systems Design. In Proceedings of the International Conference on New Interfaces for Musical Expression. pp. 581–584. London: Goldsmiths.

Graham, R. and Bridges, B. (2015). Managing Musical Complexity with Embodied Metaphors. In Proceedings of the International Conference on New Interfaces for Musical Expression. pp. 103-106. Baton Rouge: Louisiana State University.

Graham, R. and Cluett, S. (2016). The Soundfield as Sound Object: Virtual Reality Environments as a Three-Dimensional Canvas for Music Composition. In Proceedings of the Audio for Augmented and Virtual Reality Conference. Los Angeles: Audio Engineering Society Annual Conference.

Graham, R. and Harding, J. (2015). SEPTAR: Audio Breakout Circuit for Multichannel Guitar. In Proceedings of the International Conference on New Interfaces for Musical Expression. pp. 241-244. Baton Rouge: Louisiana State University.

Johnson, M. (2007). The Meaning of the Body: Aesthetics of Human Understanding. Chicago: University of Chicago Press.

Lerdahl, F. (2001). Tonal Pitch Space. Oxford: Oxford University Press.

Malham, D. (2008). Spatial Hearing Mechanisms and Sound Reproduction. Retrieved from http://www.york.ac.uk/inst/mustech/3d_audio/ambis2.htm on September 30 2016.

Malham, D. (2002). B-Zoom. Retrieved from http://www.york.ac.uk/inst/mustech/3d_audio/vst/bzoom_help.html on September 30 2016.

Milgram, P. / Takemura, H. / Utsumi, A. / Kishino, F. (1994). Augmented Reality: A class of displays on the reality-virtuality continuum. Telemanipulator and Telepresence Technologies. vol. 2351(34). pp. 282-292. Washington: SPIE.

Reynolds, C. (1983). Boids Flocking Algorithm. Retrieved from www.red3d.com/cwr/boids/ on September 30 2016.

Roth, M. (2016). Enzien Audio / Heavy. Retrieved from http://enzienaudio.com on September 30 2016.

Smalley, D. (1997). Spectromorphology: explaining sound-shapes. Organised Sound, 2(02), pp. 107-126. Cambridge: Cambridge University Press.

Smalley, D. (2007). Space-form and the acousmatic image. Organised Sound 12(1): pp. 35–58. Cambridge: Cambridge University Press.