Exploring Pitch and Timbre through 3D Spaces: Embodied Models in Virtual Reality as a Basis for Performance Systems Design

Graham, R., Bridges, B., Manzione, C., & Brent, W. (2017). Exploring Pitch and Timbre through 3D Spaces: Embodied Models in Virtual Reality as a Basis for Performance Systems Design. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (n/a ed., Vol. n/a, pp. 157-162). (Archive of NIME Proceedings). Aalborg University. https://doi.org/10.5281/zenodo.1176207

Link to publication record in Ulster University Research Portal

Published in:
Proceedings of the International Conference on New Interfaces for Musical Expression

Publication Status:
Published (in print/issue): 01/06/2017

DOI:
10.5281/zenodo.1176207

Document Version
Publisher's PDF, also known as Version of record

General rights
Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.

Download date: 01/04/2023

Exploring Pitch and Timbre through 3D Spaces: Embodied Models in Virtual Reality as a Basis for Performance Systems Design

Richard Graham
Stevens Institute of Technology
Hoboken, New Jersey, USA
rgraham1@stevens.edu

Brian Bridges
Ulster University
Derry/Londonderry, N. Ireland
bd.bridges@ulster.ac.uk

Christopher Manzione
Stevens Institute of Technology
Hoboken, New Jersey, USA
cmanzion@stevens.edu

William Brent
American University
Washington, D.C., USA
brent@american.edu

ABSTRACT

Our paper builds on an ongoing collaboration between theorists and practitioners within the computer music community, with a specific focus on three-dimensional environments as an incubator for performance systems design. In particular, we are concerned with how to provide accessible means of controlling spatialization and timbral shaping in an integrated manner through the collection of performance data from various modalities from an electric guitar with a multichannel audio output. This paper will focus specifically on the combination of pitch data treated within tonal models and the detection of physical performance gestures using timbral feature extraction algorithms. We discuss how these tracked gestures may be connected to concepts and dynamic relationships from embodied cognition, expanding on performative models for pitch and timbre spaces. Finally, we explore how these ideas support connections between sonic, formal and performative dimensions. This includes instrumental technique detection scenes and mapping strategies aimed at bridging music performance gestures across physical and conceptual planes.

Keywords

Gesture, embodied, schemas, mapping, metaphor, spatialization, timbre, feature, tracking.

ACM Classification

J.5 [Arts and Humanities] Music, I.5.2 [Design Methodology], H.5.5, [Information Interfaces and Presentation] Sound and Music Computing.

1. INTRODUCTION

This paper will first present an overview of previous approaches to instrumental technique detection tools, specifically geared to accommodate stringed instruments. A new release of a timbral analysis library will be discussed in relation to how each module may be combined to detect specific physical performative gestures, thereby extending parametric control.

Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). Copyright remains with the author(s).

NIME'17, May 15-19, 2017, Aalborg University Copenhagen, Denmark.

The paper will conclude with a discussion on an original virtual reality (VR) performance artwork that incorporates the aforementioned timbre–feature identification tools in order to expand on how performative–gestural components of timbre may be integrated with other control data such as pitch tracking in a shared embodied–conceptual space. This exploration of an integrated approach to mapping from musical structures to spaces and forces is also discussed in relation to more general applicability within performance systems design.

2. GUITAR TECHNIQUE DETECTION

Guitar technique detection was a common area of research during 2010 through 2012, particularly within the context of augmented instruments [4,10,16,17]. Reboursière et al. created a toolbox for augmented guitar performance using Max/MSP [16], which was subsequently extended to identify left and right-hand playing techniques by combining a system of divided guitar pickups (one pickup per string) with spectral and temporal processing [17]. The general scope of these projects was to extract data from instruments in real-time and then afford the user to resample, scale, and map data in order to parametrically control a variety of signal processes in creative practice.

Previous performance systems designed by Graham [5,6] utilized Vetter's implementation of Philip McLeod's SNAC (special normalisation of the autocorrelation; 2008) function for Pure Data (Pd) to provide more accurate pitch tracking [21]. This in turn permitted the recording of more subtle pitch changes absent of a clear note onset of attack, such as glissando or slides between sequential note events. By extension in the more holistic, structural domain, Graham's real-time computational models for melodic syntax [12], which include pitch extrapolation within a defined pitch class set, were improved dramatically [5,6]. This pitch space model was then used to drive a spatialization system for multichannel loudspeaker arrays. In this iteration, Graham and Bridges have expanded their initial focus on embodied dynamics within pitch space to include timbral morphology and to expand connections between theories of sound structure in electroacoustic music, such as spectromorphology and space forms [19,20], and embodiment [9]. The present stage of the work prioritizes a move towards more circumscribed ideas of timbre through the incorporation of spectral feature extraction software within our performance system's design. Overall, our goal is to provide the means by which performers or audience members can probe the notion of timbre by using a performative musical environment grounded by familiar embodied structures and dynamics.

3. MAPPING TIMBRE FEATURES

3.1 Introduction of timbreID 0.7.0

This new iteration of the system is the first to integrate Brent's timbreID [1], a set of timbral feature analysis and classification tools. The current release of timbreID is the first major update since 2010. It includes several new analysis objects and implements the FFTW library [3] for Fourier and Cosine transforms. Many of timbreID's new objects are Bark-frequency versions of previously available spectral features (e.g., barkSpecCentroid~, barkSpecSpread~, etc.). While Bark-based objects may prove useful for some aspects of guitar technique classification, the work described here relies on detailed high-frequency content lost in Bark spectrum conversion, so these objects were not employed. However, many of timbreID's basic spectral features were useful in classification, and a new time-domain analysis object, [waveSlope~][1], was valuable for distinguishing between gradual and sudden note onsets. In addition, the new FFTW implementation allowed freedom of choice for analysis window size based purely on timing considerations without restriction to power-of-two sizes or the need for zero padding. Beyond feature extraction objects, timbreID's onset detection object, [bark~], can be used to accurately trigger analysis timing. [bark~] offers a variety of parameters for fine-tuning detection to reduce false positives, and reports note onsets with lower latency than [sigmund~]'s "notes" function in its default configuration.

Brent's tools highlight the question of whether we need to develop a more accessible or intuitive way to map (and to talk about) timbre. Such tools are necessary for discussing timbre at a higher level, making it conceivable to map data, representative of performative gestures connected to timbral nuance, to parametrically extend a performer's sonic palette through digital signal processing. It is our aim to use timbreID to improve the recognition of a set of specific instrumental techniques in our performance system's design. Basic timbre feature data from timbreID analysis objects can be combined to form larger feature vectors that correlate with specific physical performance gestures and guitar techniques, such as legato, alternate picking, sweep picking, palm muting, finger-style, and so on. The work described here focuses on distinguishing between picking versus finger-styled performance gestures and palm muting versus open-string performance gestures on a multichannel electric guitar.

3.2 Instrumental Features and Emerging Mapping Strategies

In a series of initial tests, we were able to reliably classify finger-style versus picked notes based on a compound feature vector including spectral brightness, spectral spread, spectral centroid, spectral flatness, and waveform slope. Our analysis windows were 512 samples long (11.61ms at a 44.1kHz sampling rate), and were carefully timed to capture the initial note onset. Spectral brightness and centroid provide different perspectives on the amount of high-frequency content present. During the initial attack segment, more high-frequency content was present in finger-style note onsets, resulting in higher values for brightness and centroid. These results stand in opposition to what one might expect: fingerpicking produced more high-frequency transients than flat-picking. The finger-style technique used in our test recordings was close to slap/pop technique, which may account for the increased high-frequency energy. By contrast, the picking technique employed was kept quite flat to ensure a clean note attack. A rudimentary classification example is given[2], where even with spectral centroid alone, it is possible to distinguish between finger-style and picked note events. The inclusion of other features, and time-varying feature data over the course of the initial onset makes classification more robust. Spectral spread values were also higher for finger-style note attacks, indicating that spectral energy is less tightly concentrated near the centroid frequency when using this playing style. With regard to spectral flatness, finger-style notes produced broadband energy across a range of frequencies, resulting in a higher flatness value in comparison to measurements for picked notes.

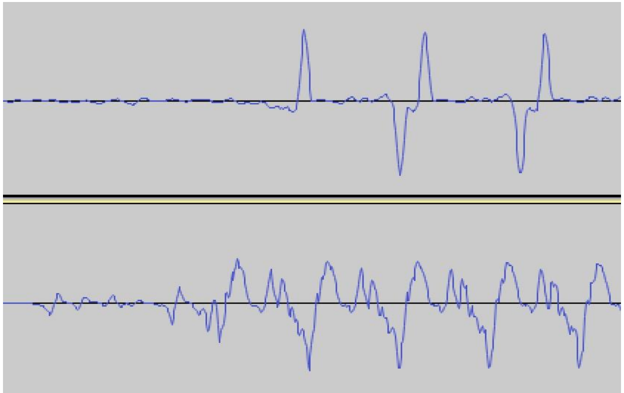

Figure 1. Onsets of picked (upper) and finger-style (lower) notes at a window size of 11.61ms.

Finally, waveform slope was reliably higher for finger-style attacks. [waveSlope~] calculates the best fit line through the absolute value of the waveform and reports its slope. Thus, during the attack segment of a note event, the slope will be positive as amplitude rises to the peak, negative during the initial decay, and near zero during a steady-amplitude sustain segment. [waveSlope~] can also normalize the amplitude of the signal per analysis window to remove variations due to dynamics. We chose to activate this normalization in order to more meaningfully compare quiet and loud events. Under this analysis technique, we observed that finger-style notes had a more gradual onset than the immediate attacks of picked notes. However, finger-style notes produced higher slope values because their attack segment spanned a greater length of time within the 11.61ms analysis window *(Fig 1)*. The immediate onset of the picked notes produced an amplitude envelope that essentially jumped straight to a sustain segment. With most of the analysis window filled by a steady-state waveform (and only a small amount of silence before the onset), the resulting slope measurements were nearly flat[3]. At a window size of 11.61ms, the attacks of palm-muted and open string note events are very similar across many features. It was necessary to increase window size in order to successfully distinguish these events. With a window size of 985 samples (22.34ms at a sample rate of 44.1kHz) we were able to distinguish between palm-muted and open events using waveform slope, spectral spread, and spectral roll off (see table 1, next page).

[1] waveSlope~ example video 1: http://bit.ly/2kP9sce

[2] timbreID examples: http://bit.ly/2pwQntW

[3] waveSlope~ example video 2: http://bit.ly/2kSdfBK

Table 1. Finger Picked v. Picked Note Events

Finger Picked v. Picked Note Events - settings: analysis window of 512 samples and a 44.1kHz sample rate	
waveSlope~	Values are *HIGHER* for finger style.
specBrightness~	Values are *LOWER* for picked events. *Boundary frequency set to 6800Hz.*
specSpread~	Values are *LOWER* for picked events.
specCentroid~	Values are *LOWER* for picked events.
specFlatness~	Values are *LOWER* for picked events.

The increased window size allows enough time to observe the beginning of the amplitude envelope's sudden release due to muting. This produces a negative waveform slope value, as opposed to the very flat slope of an open string sustained event. Spectral energy is more tightly concentrated near the centroid for palm-muted events, resulting in a lower spectral spread. Spectral rolloff locates the frequency below which a certain concentration of spectral energy is found. We used a concentration threshold of 85%, and found significantly lower rolloff frequencies for the palm-muted events, a technique where high frequencies are dampened by physically interfering with string vibration[4].

Table 2. Palm Muted v. String Events

Palm Muted v. Open String Events - settings: analysis window size of 985 samples and a sample rate of 44.1kHz	
waveSlope~	Values are *HIGHER* for open string events
specSpread~	Values are *HIGHER* for open string events.
specRolloff~	Values are *HIGHER* for open string events.

This *modification gesture* [8] provides two different modes of performance: plectrum (pick) and finger style. In the context of a mapping strategy, the different type and implicit dynamism of the detected note articulations ('harder' with pick or 'softer' with finger) maps to a *tension–projection* and *linearity versus 'inertia'* dynamic [2,6]. The mapping of instrumental gesture data to signal processing parameters in earlier iterations of the system's design were focused primarily on spatialization. More energetic instrumental gestures were mapped to outward, directionally focused spatial projection. Less energetic attacks gave rise to a shorter spatial projection range and more '*inertial'*/chaotic behaviors. This was achieved by mapping a boids algorithm to determine the spatialized output of each guitar register, the behaviors of which were determined by the aforementioned tonal model [5,6]. In the context of this previous work, without timbreID, the more simplistic approach saw differing amplitudes of notes being tracked as opposed to differences in more nuanced articulation gestures.

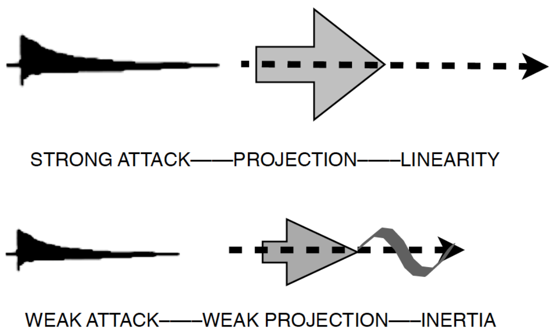

Figure 2. Different attack-projection-linearity profiles from Graham and Bridges (2015) [6]

In the present approach, two distinct modes are available, as opposed to a continuum, while still broadly conforming to the embodied associations of directionality and modification of linearity versus nonlinearity found in earlier iterations of the performance system [5,6]. The benefit of this approach is its accessibility for performers; very clearly defined changes in functionality are available which are nonetheless based on familiar affordances of the instrument [13].

3.3 Gesture, Timbre & Spatial Relationships

Potentially, the detection of palm muting provides an affordance–plus–mapping association with inertia and damping, creating a connection with the embodied concepts underpinning our previous *high-attack–strong projection–linearity* versus *low-attack–weak projection–inertia* implementation. In short, the less energy that is present within the system, the less energetic and linear the projection. More importantly, in a similar fashion to the modification gesture functionality discussed above, this approach–allows for more deterministic control of the output of the performance system.

By utilizing this data in combination with a deterministic force model (the Unity physics engine) and an artificial life model (boids), we believe the enhanced accuracy in tracking instrumental affordances and their application through environment–derived models will prove engagingly immersive for performers. This sense of immersion within a musical space that is both concrete (physical) and virtual (conceptual) is intended to support the development of further cross–domain mappings. Having previously focused on embodied–ecological models of tonal structure and melodic syntax, dynamics of note articulations, and their connection with spatial concepts,–we now find ourselves in a position to interrogate timbre as performative input and output domain.

We have previously [2, 5, 6] examined various ways in which Smalley's model of structure in electroacoustic music (*spectromorphology*) [2] are compatible with *embodied image schema* theory [9,11], a *gestural and embodied model of conceptual relationships*. We argue that spectromorphology is a dynamic, performative model of timbral structure; an enactive space comprised of connections forged on the basis of similarity between forms and the energetic associations of these forms (Smalley's *energy-–motion profiles*). We believe that three key cases of qualitative dimensions of movement [2,6]—*tension*, *projection* and *linearity*, which we regard as parallel terms for Smalley's [19,20] 'motion rootedness', 'motion launching', and 'contour energy/inflection', may provide grounding embodied narratives for describing timbral spaces, structures, and relationships.

[4] timbreID examples: http://bit.ly/2kNWPv8

3.3.1 Structural Descriptions of Timbral Relations

The search for structural descriptions of timbral relations was a significant feature of research at the intersection between early computer music, psychoacoustics and music cognition. Grey and Gordon [7] provided a three–dimensional scaling solution for various resynthesized Western musical instrument tones, based on (1) spectral energy distribution, (2), relative synchronicity/asynchrony in higher partials[5] and (3) the presence or absence of low–amplitude, high–frequency energy during the attack phase. We argue that this three-dimensional cognitive model could be translated into an embodied–cognitive model [figure 3].

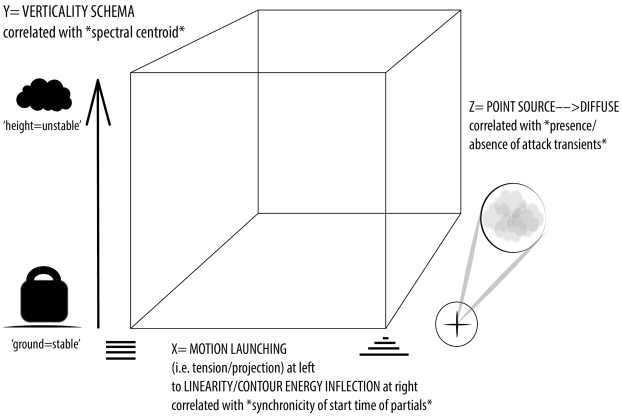

Figure 3. An embodied timbre–space; concepts from Smalley (1997) and Johnson (2008), after Bridges and Graham (2015)

In Grey and Gordon's classic spatial representation [7], spectral distribution was on the y-axis, synchrony was on the x-axis and presence of 'noisier' attack transients was on the z-axis. The assigning of spectral distribution to the y-axis is compatible with a basic embodied perspective on frequency, which treats it as a *verticality schema*. The degree of synchrony/asynchrony and presence/absence of attack transients could be seen as relating to embodied dynamics of order (or smooth change of state) and more chaotic behavior, with clear attack transients also contributing to a sense of clear spatial acuity/lower localization blur (thus necessitating a flipping of the z-axis). This perspective is somewhat similar to morphological notation [15], which treats spatiotemporal structures from electroacoustic music as present at front and more diffuse at rear.

Applying these ideas, along with embodied–cognitive narrative rubrics back to the classical timbre–space model of Grey, we have an embodied, and thus performative, timbre–space model. X is synchronicity of entry/exits of sound events, z is presence/absence of noise elements, and y is relative distribution of on a frequency–height axis [figure 3]. We can therefore posit embodied dynamics within the structure presented in [table 3]. Taken together, we believe that these theoretical contexts— (1) classic timbre–space studies, (2) spectromorphology, and (3) embodied image schema theory as a structural theory of embodied cognition—can support the interrogation and extension of tracked and mapped musical gestures within a virtual or mixed–reality space.

[5] Future work will focus on partial synchronicity.

Table 3. Mapping Strategies based on Embodied Dynamics

Embodied Dynamics - Potential Mapping Strategies	
Dynamic 1: *Temporal Synchronicity of Attack Envelopes*	X ranges from *motion launching* (rapid dynamic change, more synchronous entry) to *gradual contour energy* (asynchronous entry of partials).
Dynamic 2: *Spectral Energy Distribution: Height vs. Rootedness*	Y via the spectral centroid gives us two scales and dynamics: *contour energy* (verticality schema: pitch height) and associated *motion rootedness*; regions of stability.
Dynamic 3: *Spatial Clarity within Individual Sound Sources*	Z via presence or absence of attack transients articulates *motion rootedness or tension* (audible transient products of inertia) to ungrounded events (diffuse or sustained tones). This is related to a *diffuse–to–point source* spatial coverage schema.

3.3.2 Embodied Structures and Texture–Music Affordances within our Performance System

Key embodied image schemas previously identified within common practice tonal structures include *cycle, path, verticality, center–periphery, container* and *balance* [5,6,9]. In the context of more directly musical gestural and textural schemas, we have a number of modifications. Thus, in our latest iteration, we map instrumental and musical affordances in the manner of broader bodily gestures, with archetypes of center–periphery and the qualitative dimensions of path schemas. Table 4, below, expands on these gestural types and sound–structure affordances.

Table 4. Using Bodily Gestures to Drive Parametric Change

Gesture	**Sonic Affordance**
Short, repetitive movements	Detached individual sound events, cycle–loops.
Expansive gesture	Clear path or projection outward (versus inertia).
Less expansive gesture	Weak projection (inertial) and chaotic path.

The dynamics between legato, pronounced–attack features and contextual recognition of less expansive gestures provide a range of gestural affordances conceptually compatible with the dynamic *embodied dimensions* noted in table 3. *Dynamic 1* (envelope profile) integrates small, repetitive movements with expansive gestures. *Dynamic 2* (spectral energy distribution) provides the vertical component of the projected path in the expansive gesture. *Dynamic 3* (presence or absence of transient detail) is provided by the joining of smooth or sustained cyclical structures and iterative ones.

Our present performance system combines the gesture–tracking and identification available from timbreID with a number of key process mappings. Significantly, with regard to the combination of our conceptual framework, each object is assigned unique mass and drag values determined by timbreID. This can be linked to *Dynamic 2 (Y Axis)* within the proposed embodied timbre space. The projection outward that determines the path and projection strength is determined by Bark-frequency cepstral coefficients, again another low-level timbral feature, which we can link directly to *Dynamic 2 (Y Axis)* concerning spectral centroid position. This example of a conceptual affordance circle illustrates the potential of this type of virtual reality system for investigating a wider range of integrated mappings for live performance systems design.

More generally, our new system sees a representation of gestural dynamics through the spawning of objects and meshes in VR. At a macro-level, different instrumental registers modulate between overlaid environments, which fade as the sound decays. The multichannel audio feed from an instrument is parsed and sent to timbreID audio externals. Newly detected events then create virtual objects with unique projection profiles within the Unity game development environment. As noted earlier, the *center–periphery* concept of localized presence to projection/diffusion can lends itself to a conceptual framework based on more or less energetic and expansive gestures. Such object creation and projection dynamics are even more apparent in a first-person perspective view, and may provide insight for both performers and audiences into music–structural correlates[6]. In terms of sound spatialization behaviors, the *strong–attack–projection–hold* affordance sees objects visibly and audibly developing less physical–deterministic trajectories as the environmental-tonal hierarchy model takes over spatial movements during the sustain/decay phase of each note. In addition, in textural terms, the change in granular settings from short, pointillist grains to longer, more gradually enveloped ones provides a related dimension of note–gestural control over processes which are spatiotemporal and textural. Additional looping functionality for the granulating buffers also demonstrates further potential for the integration of performance gestures detected by timbreID.

Following on from this idea, one of the most sophisticated potentials of the system is its ability to control FFT-based timbral shaping techniques. Beyond the present implementation *projection* mapping and the location and force–based data obtained from the interactions between performance gestures and the physics engine, there is the additional potential to map and select groups of spectral bins within a virtual scene through their intersection with note–objects and their associated force–dynamics or through combinations of gestures using the spawned note-objects also as 'timbre–shaping cursors' (discussed further in relation to the details of the control mappings and audio processing in section 4.1, below). Such an attack-projection–linearity model within spectral space would allow for more extensive timbral shaping possibilities with this system whilst still maintaining a connection with the original performative and conceptual frameworks.

4. Disrupt/Construct: Exploring Object, Place and Space in a Virtual Environment

Disrupt/Construct is a performance piece regarding the origin of object and place[7]. An improvising musician and visual artist explore assumptions about personal memory and the disruption of automatic trust in paramnesia. A history of the musician's gesture data is recorded from an augmented instrument and mapped to determine a variety of interactions between sounds and objects within a complex virtual scene. The visual artist has a degree of control over the unfolding virtual environment through a motion tracking system, which allows the accompanying performer to interact with the objects to trigger additional audio samples within a static timbre space. The audience views a projection of the artist's first-person view within the virtual environment in addition to viewing the improvising musician in the physical performance space. This piece seeks to reposition or re-contextualize performance systems design within the context of virtual reality environments while exploring where a music performance–and by extension the human performer–may be situated along the Reality-Virtuality Continuum [14]. The chosen textures and meshes are representative of objects and places from the performer's previous experiences. The performer is effectively mixing images and sounds from a variety of personal experiences to produce unique historical amalgams. These new events create a historical confusion necessary to disrupt assumptions held by the performer about their past. This notion may be extended to challenge assumptions regarding how musical data should be mapped as a form of parametric control. In essence, the virtual environment challenges or forges new relationships between the performer's musical responses to the objects and places of his past. This collaborative artwork aims to explore how one can reflect and re-contextualize the relationships within a performed historical ecology by extrapolating a virtual representation of its past, present and future.

4.1 Timbre and Pitch Spaces: Mapping Sound Objects in Virtual Reality

In this VR–based iteration, more abstract structures (the use of a tonal hierarchy model, via boids) exist at the periphery of the virtual space, representative of the fringes of our consciousness and, perhaps, a general sense of disembodiment experienced when interacting with abstract structures. In this piece, spheres represent objects of abstraction, whereas environmental objects (such as walls, rocks, or wood) are representative of more episodic events relative to the constructed virtual scenes. The deterministic embodied mappings tying these different objects and dynamic behaviors together highlight the usefulness of a virtual reality system in exploring various aspects of performance gesture and musical textures and macro structures.

Sound objects project outward from the position of the performer (center in the world by default), determined by Bark-frequency cepstral coefficients. At a certain amplitude threshold, the sound objects, in this case represented by spheres, will select a spatial location based on our tonal model from previous system iterations [2,4,5,6]. For example, resolving to the tonic will cause objects to flock to the central position of the world, in this case determined by the position of the performer. Tones outside of the chosen diatonic set will increase inertia and decrease centricity and attraction behaviors.

The modulation between these modes of spatialization presents a juxtaposition of an event directly related to a burst of (spectral) energy with a dynamic tonal spatial mapping.

Note-attack gestures spawn prefabricated objects of abstraction, which are directly tied to the narrative of a particular environment. Audio buffers store a unique audio loop per environment, which are analyzed and then organized into a three-dimensional granular and static timbre space, relative to user-defined spectral attributes. The physics-enabled prefabricated objects then act as playback heads or cursors, triggering audio samples of the static timbre space when they

[6] See FPC in video example: https://vimeo.com/202093804

[7] *Disrupt/Construct*: https://vimeo.com/202093804

collide with grain positions generated by the aforementioned spectral analysis. This, coupled with the modulation between each mesh based on the amplitude of each register, allows the performer to reposition objects within different virtual spaces.

5. CONCLUSION & FUTURE WORK

We have presented the next stage in our collaboration between theorists and practitioners through the design of new mappings for a revised performance system, which accommodates pitch and timbral structures within a three-dimensional mixed-reality space. We have outlined a new application of timbral detection tools to detect common instrumental vernacular, expanding the relationship between physical instrumental technique, selective gestures and digital signal processing. Furthermore, we have presented a new work that explores the notion virtual reality as an incubator for developing performance gesture ecologies. Given that Unity provides a sophisticated physics engine, the player/performer can interact directly with spawned rigid body objects with the virtual scene as they perform on their instrument. The performance system designer can also interrogate and refine a range of embodied concepts in mappings through visualization and performative exploration.

6. ACKNOWLEDGEMENTS

We would like to thank Dr. Thomas Maguire at Ulster University in Northern Ireland and Dr. Kelland Thomas and SCENE Lab at Stevens Institute of Technology in Hoboken, New Jersey, for supporting this research project.

7. REFERENCES

[1] W. Brent. timbreID. Web. <http://williambrent.conflations.com>. Accessed January 2017.

[2] B. Bridges and R. Graham, Electroacoustic Music as Embodied Cognitive Praxis: Denis Smalley's theory of spectromorphology as an implicit theory of embodied cognition. *Electroacoustic Music Studies Proceedings: EMS15 The Art of Electroacoustic Music*, 2015.

[3] M. Frigo and S. Johnson. FFTW library. Web. <http://www.fftw.org> Accessed January 2017.

[4] R. Graham. Expansion of Electronic Guitar Performance Practice through the Application and Development of Interactive Digital Music Systems. Ph.D. Thesis. Ulster University, Northern Ireland, 2012.

[5] R. Graham and B. Bridges. Gesture and Embodied Metaphor in Spatial Music Performance Systems Design. In *Proc. NIME Conference*, pp. 581–584, Goldsmiths, London, 2014.

[6] R. Graham and B. Bridges. Managing Musical Complexity with Embodied Metaphors. In *Proc. NIME Conference*, pp. 103-106. LSU, Louisiana, 2015.

[7] J. Grey. Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America,* 61, 5. 1977.

[8] A. Jensenius. Action–Sound: Developing Methods and Tools to Study Music-Related Body Movement. *PhD Thesis*, University of Oslo, 2007.

[9] M. Johnson. *The Meaning of the Body: Aesthetics of Human Understanding*. University of Chicago Press, Chicago, 2007.

[10] O. Lähdeoja. Une approche de l'instrument augmenté: La guitare électrique. *PhD thesis*, Ecole Doctorale Ésthétique, Sciences et Technologies des Arts, Paris 8, 2010.

[11] G. Lakoff and M. Johnson. *Metaphors We Live By*. Chicago, University of Chicago Press, 1980.

[12] F. Lerdahl. *Tonal Pitch–space*. Oxford University Press, Oxford, 2001.

[13] D. Levitin. Control parameters for musical instruments: a foundation for new mappings of gesture to sound. *Organised Sound*, 7(2), 171-189. 2002.

[14] P. Milgram et al. Augmented Reality: A class of displays on the reality-virtuality continuum. *Telemanipulator and Telepresence technologies* 2351, 34. pp. 282-292. 1994.

[15] K. Patton. Morphological notation for interactive electroacoustic music, *Organised Sound*, 12(2), pp. 123–128. 2007.

[16] L. Reboursière, et al. Multimodal guitar: a toolbox for augmented guitar performances. In *Proc. NIME Conference* 2010.

[17] L. Reboursière et al. Left and right-hand guitar playing techniques detection. In *Proc. NIME Conference*, 2012.

[18] C. Reynolds. Flocks, Herds and Schools: A distributed behavioral model. SIGGRAPH, 21, 4, pp. 25–34, 1987.

[19] D. Smalley. Spectromorphology: explaining sound–shapes. *Organised Sound*, 2, 2, pp. 107–126, 1997.

[20] D. Smalley. Space-form and the Acousmatic Image. *Organised Sound*, 12, 1, pp. 35-58, 2007.

[21] K. Vetter. helmholtz~ finds the pitch. Web. <http://www.katjaas.nl/helmholtz/helmholtz.html>. Accessed January 2017.