Managing Musical Complexity with Embodied Metaphors

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

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

ABSTRACT
This paper presents the ideas and mapping strategies behind a performance system that uses a combination of motion tracking and feature extraction tools to manage complex multichannel audio materials for real-time music composition. The use of embodied metaphors within these mappings is seen as a means of managing the complexity of a musical performance across multiple modalities. In particular, we will investigate how these mapping strategies may facilitate the creation of performance systems whose accessibility and richness are enhanced by common integrating bases. A key focus for this work is the investigation of the embodied *image schema* theories of Lakoff and Johnson alongside similarly embodied metaphorical models within Smalley's influential theory of electroacoustic music (*spectromorphology*). These metaphors will be investigated for their use as grounding structural components and dynamics for creative practices and musical interaction design. We argue that pairing metaphorical models of forces with environmental forms may have particular significance for the design of complex mappings for digital music performance.

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 project explores how digital technologies are changing the way we create and think about creative practices specifically through the use of embodied metaphors in rich sensory environments to interact, transform, and immerse performers and audience members. To date, we have explored certain basic embodied grounding perspectives on performance and general creativity, where digital performance technologies provide synchronized sensory feedback in response to physical and figurative performance gestures. Our contention is that metaphorical mappings based on familiar patterns of bodily sensations and actions may provide intuitive access points for the control and integration of complex musical and sonic materials, enhancing accessibility and managing complexity. Through a combination of theoretically–informed and practice–led design iterations, we are seeking to extend structural models of embodied relations that will prove useful to creative arts practitioners and interaction designers in developing, questioning, and extending their digital practices.

2. EMBODIED METAPHORS FOR THE CREATION OF SPATIAL MUSIC

2.1 Initial System Iterations
The initial iterations [6] of this project were motivated by the desire to connect the rich structural potential of familiar guitar performance gestures with the control of a broader field of sound structures opened up by contemporary digital processing technologies (specifically, timbral shaping and spatialization processes). The earliest versions of the system (implemented in Pure Data) used pitch tracking and feature identification tools to map performance gestures to macrostructures that connect music pitch, dynamics, and time with spatialization and sound shaping processes (see figure 1, below).

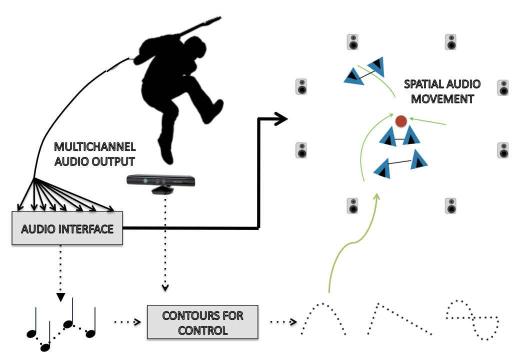

Figure 1: Initial system design: pitch, amplitude, and motion tracking to control spatialization and other processing in Pd

This system could therefore be seen as forging *gestural narratives* between performance gestures and control mappings for digital signal processing. It accomplished this by exploring embodied perspectives on pitch-based-schemas as the basis for the resulting mapping strategies. A tonal hierarchy model from Lerdahl [10,11] provided a spatial center–periphery mapping for pitch materials, further animated by a *boids* flocking algorithm [12,14][1], which organized the dynamic movements of the system's spatialized audio outputs [1,14].

[1] Boids for Max by E. Singer (1997) http://ericsinger.com/cyclopsmax.html.

The innovation within our own usage of the *boids* algorithm is its metaphorical mapping of musical forces, connecting melodic attraction and inertia values from Lerdahl's tonal model [10,11] to in–kind spatial flocking parameters (see figure 2, below). In this mapping, melodic intervals that are close to tonal center (e.g. triadic) will cause movement of voices towards the center. More tonally distant materials (e.g. chromatic) will cause movement towards the periphery. Many musicians will be familiar with these general structuring principles of tonal relationships in common practice music and will recognize the flocking behaviors as broadly predictable and related to their own understanding of musical macrostructures. Thus, the use of embodied correlates between input and output provides a means of managing emergent complexity via *conceptual mapping* [4,7,14,16].

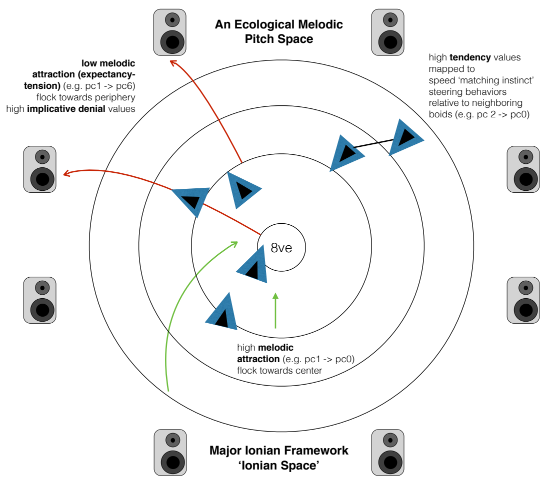

Figure 2: Dynamic tonal–spatial mappings of initial system iteration, 'animating' a tonal hierarchy and attraction model derived from Lerdahl [10,11] using a *boids* algorithm to control the spatialization of voices

These approaches have provided the basis for a further investigation of musical mappings from the perspective of a wider range of embodied metaphors. Our initial approach investigated relatively obvious parallels between various control parameters and embodied analogues, such as the mapping of melodic contours to spatial trajectories. We have since sought to refine and extend our models to consider the structural theories of embodied cognition, namely the *image schema* theories of Lakoff and Johnson [8,9]. We will now present these embodied image schema theories and related discourses from the perspective of informing the design of digital musical instruments, interfaces, and mappings [5,6,16].

2.2 Image Schemas, Musical Relations, and Embodied Mapping Strategies

2.2.1 Unifying mappings from embodied image schemas: performance gesture ecologies

We have previously [6] discussed how a variety of approaches to musical structure, gesture tracking, and control mappings may be unified within an embodied model. We propose that such a usage of embodied models can be of significant help for designers and creative practitioners in unifying such diverse domains of performance gestures and control mappings. We refer to this type of integrated model as a *performance gesture ecology*: a shared gestural space that conveys spatial–relational dynamics of performance gestures, resulting mapping strategies, and resulting sound structure. More specifically, our system combines physical motion–tracked movements with the *figurative gestures*[2] [6] of musical macrostructures, such as pitch contours and rhythmic groupings (see figure 1). The shared gestural space is the unifying 'glue' connecting input, mapping, and output domains, facilitating accessible and immersive designs for NIME. Schacher et al. [13] use the term *ecological relationships* for a related idea, the blending of human interaction gestures and system responses.

Brower [2,3] and Johnson [8] have previously applied image schema theories to tonal music. They have also been investigated for their practical significance in music software design [16,17]. Some of the key basic image schemas are illustrated below (figure 3).

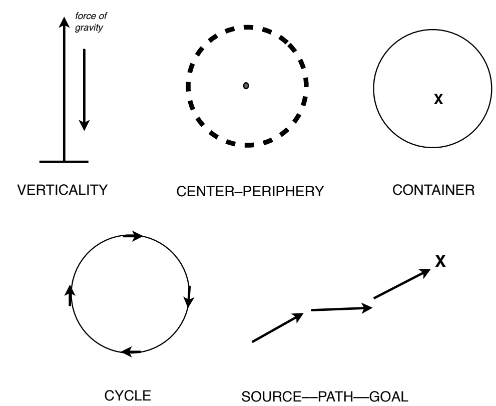

Figure 3: Key embodied image schemas

These image schemas may be applied to goal–oriented tonal tension/release (*center–periphery*, *container*, *source–path–goal*) or cyclical structures, along with hierarchical pitch relationships (*verticality*). [2,3] suggest common practice tonal structures may be based on these image schemas]. Our previous work [6] has attempted to unify these ideas with cognitively–based theories of tonality [10,11], culminating in the tonal–spatial mappings discussed above, which adopt *center–periphery*, *cycle,* and *verticality* schemas.

2.2.2 Unifying mappings from musical movement and force models

Smalley's influential theories of electroacoustic music [15] could be seen as anticipating such an embodied–structural usage through their implicitly embodied discourse. These theories see archetypal envelope forms (termed *energy–motion profiles*) as the key generative dynamics of electroacoustic music. Smalley theorizes these generative dynamics in terms of embodied associations such as force and coherence of path or apparent causality. For example, attack phases are considered to have associations with departure, launching, or emergence; continuation phases provide impressions of spatial passage; terminations or release phases are associated with arrival or closure (see table 1). These ideas may further inform the design of mapping strategies for hyper instruments.

[2] Figurative gestures may pertain to a fictive movement or motion perceived by the listener, including musical metaphors and abstracted musical figurations such as melodic and chord structures.

Table 1. Smalley's [15] basic *energy–motion profiles* (envelope profiles and stages) and theorized embodied associations

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

This idea of music's structure as being based on *force metaphors* finds striking and sustained parallels in Johnson's [8] later work. Johnson theorizes music metaphorically as *music–as–moving–force*, finding embodied associations in various musical events. For example, the musical *motion* of melodic sequences may traverse a *path*, forces generate musical events, and musical cadences may be related to a spatial metaphor of temporary cessation of that motion. Another corollary of the force–based model is that it draws attention to an image schema's manner of *execution*; Johnson [8] notes various *qualitative dimensions of movement*, making associations between the manner of execution and their resultant path.[3]

Two key cases stand out for the present purposes. *Projection* denotes an extremely energetic movement such as a sudden rate–change movement with less inertia that results in an event that continues to sustain itself for a longer period of time. *Linearity* denotes whether a resulting path is more coherent and incoherent, relating to the manner of its execution. Smalley [15] uses similar concepts (*motion launching*, *contour energy/inflection*) to describe emergent structure in electroacoustic music. These associations may be seen as further reinforcing the idea of embodied frameworks for gestural control and mediating mappings in musical interaction design, specifically related to the effort–force dynamics of performance gestures and their results.

3. APPLYING EMBODIED MAPPINGS

3.1 Tracking and Mapping Qualitative Dimensions of Movement

Instrument designers can use audio analysis techniques in systems design to track and map the energetic nature of such sounded performance gestures. For example (see figure 4, below), clearly detected (higher energy) attack transients may reset voices to a central (grounding) spatial location, but may be projected quite rapidly towards the periphery of a multichannel performance space (high *projection* and *linearity)*.

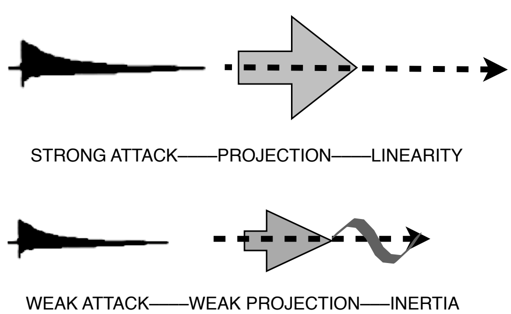

Figure 4: Different *attack–projection–linearity* profiles

[3] For example, a source–path–goal schema may be initiated via a motion that requires greater effort to overcome inertia.

Less energetic attack transients may produce an effect more like *emergence*[4]. That is to say, a less energetic attack may result in a shorter projection range (it may not travel as far) and less linearity of eventual path (for example, a more pronounced avoidance behavior by a flocking algorithm).

A more extensive consideration of gestural force or inertia may allow designers to develop systems that manage the complexity of larger–scale sound processes. Our earlier use of the *boids* algorithm as a mediating mapping to control spatial trajectories and tendencies conforms to this idea, combining physical and figurative gestural elements using force metaphors[5]. Following these principles, one may map other physical gestures to preexisting figurative structures, producing greater degrees of expressive and conceptually clear control for a variety of gestural–to–textural relationships.

We deem such ideas as highly applicable to performance systems designs that seek to engage with complex sonic structure. Smalley [15] further models sound gestures and streams using various axes, such as *order* to *disorder* in terms of temporal and frequency–based continuity and coherence; a *loose–tight continuum* that encompasses degrees of continuity in materials, from coherent auditory streams to granulation to iterative, and from periodic to aperiodic materials. This continuum pairs the degree of order in temporal and frequency–based structures, providing a useful timbral–textural superstructure for organizing timbral transformations and, more particularly, relationships between momentary performance gestures and the contours of sound transformations.

3.2 Physical Gesture to Projection: Attack–Projection–Linearity/Inertia

Our force–projection mappings allow us to connect dynamism of initial performance gestures to the degree of projection and linearity or chaotic paths of the resulting sonic structures. As noted previously, a high–energy note event from a dynamic physical gesture will center voices in within a multichannel loudspeaker array, via an act of strong linear projection. **These cases are illustrated in audio example 1**[6]. A low–energy note event will produce a less pronounced projection effect and a less linear path. **These cases are illustrated in audio example 2**[7]. The coherence of projection for the strong attack case may trigger an increase *attraction and matching* values, causing more coherent flocking behaviors of spatialized voices. **These cases are illustrated in video example 1**[8].

It is useful to note how this model works alongside our previous tonal–spatial mappings. For example, in the original tonal mapping, more dissonant or distant materials (within a tonal hierarchy) were assigned to spatial periphery with less pronounced coherence of path between different members of the boids flock (due to lower *attraction* values). However, the new projection–linearity mapping provides a more definite statement of tonally dissonant materials that produces more clearly defined spatial trajectories. Softer ('more uncertain') playing of more dissonant materials will produce more moderate spatial diffusion of voices, which is less unidirectional.

[4] *Emergence* is a more subtle behavior, such as a low-dynamic attack profile.

[5] For example, spatial and timbral processing may be influenced by the velocity or acceleration of a physical performance gesture.

[6] https://soundcloud.com/spatialschemas/nime15_ex1

[7] https://soundcloud.com/spatialschemas/nime15_ex2

[8] https://www.youtube.com/watch?v=0JhV2A_k2ec

3.3 Physical Gesture and Mappings with Distortion Processes

Following the work of Smalley [15], we have introduced the idea of a *loose-tight continuum* as a superstructure for timbral–textural processes articulating order to disorder, as opposed to spatial audio processing only. We have noted how this model may apply in terms of temporal and frequency structure. Extending this via an embodied–relational metaphor, the control of a monophonic–to–polyphonic distortion processes may provide one example of such integration. For example, our system's design comprises a multichannel audio pickup, which allows an electric guitar's output to be separated or combined at will. This allows for distortion to be applied in a summed/monophonic or separated/polyphonic or multichannel manner. If the audio output is summed, a distortion process will entail non-linear intermodulation effects, producing audible sum and difference tones and beating effects.

An embodied approach to mapping this combination of distortion processes would see more dynamic attacks (more *tension–energy*) producing more pronounced monophonic distortion effects, accompanied by a spatial centering. This effect could be seen as a combination of a centering and reaffirmation of embodiment, as the resulting audible intermodulation effects present the electroacoustic instrument design in greater detail or granularity, as opposed to a disembodied multichannel presentation (activated by less dynamic attacks, i.e. less *tension–energy*), which minimizes such sonic interactions. Thus, in these mappings two co-existing loose–tight dynamics may be said to interact in sometimes opposing ways. A lower–dynamic event may produce a less acoustically–spatially centered result; a loose coupling. **These cases are illustrated in audio example 3[9].** A higher–dynamic event will produce a tighter resulting coupling in terms of spatial–textural effects (i.e. maintaining the instrument at center, with enhanced sense of a bodily *detail* or an increase in granularity. Overall, these embodied mappings are a way to balance the rich variety of effects of opposing schemas in ways that are still intuitively accessible.

4. CONCLUSION

We have presented a series of embodied metaphors that may prove useful in music performance systems and DMI design. These conceptual metaphors combine physical performance gestures and figurative gestures in a shared gestural space (a *performance gesture ecology*), which also encompasses the relational dynamics of a mapping strategy and overall creative output. In seeking accessible embodied metaphors to inform rich mapping strategies, we examined various cases from the image schema theories of Lakoff and Johnson [8,9], their extension via force metaphors, and the qualitative dimensions of movement and embodied associations in Johnson [8]. We have extended these structuring principles to musical mappings by comparing Johnson's [8] ideas with similar structural dynamics in Smalley's [15] theory of *spectromorphology*.

Given the breadth of musical and performative materials that these typologies encompass, we believe that comparative studies of image schema theories and Smalley's ideas may fruitfully inform the design work of the NIME community. Building on our tonal–spatial and force–based dynamics, we have further extended two embodied mapping strategies: (1) *attack–projection–linearity;* (2) *center–periphery*. These mappings emphasized embodiment through spatially centered intermodulation distortion effects and disembodiment through a spatially peripheral reduction of intermodulation distortion effects (reduced roughness): *loose/tight continuum*. While these cases are clearly not exhaustive in terms of spatial or textural dynamics, they demonstrate the potential of this theoretical combination, integrating the physical affordances of familiar performance gestures with the figurative gestures of tonal structures and the conceptual affordances of image schemas. This approach to systems design will arguably generate formulaic patterns for real-time music composition, which may be developed and refined as an extended performance vocabulary. Future iterations of the system will account for spectral considerations including sounds that do not have a clear transient or fast attack profile.

[9] https://soundcloud.com/spatialschemas/nime15_ex3

5. REFERENCES

[1] E. Bates. The Composition and Performance of Spatial Music. *Ph.D. Thesis*. University of Dublin, Trinity College, Dublin, 2010.

[2] C. Brower. A Cognitive Theory of Musical Meaning. *Journal of Music Theory*, 44, 2, pp. 323–379, 2000.

[3] C. Brower. Paradoxes of Pitch Space. Music Analysis, 27, pp. 51–106, 2008.

[4] C. Cadoz and M. Wanderley. Gesture–Music. In: M. Wanderley, M. Battier. *Trends in Gestural Control of Music*. IRCAM, Paris, 2000.

[5] M. Donnarumma, B. Caramiaux, and A. Tanaka. Muscular Interactions. In *Proc. NIME Conference*, pp. 128 – 131, Seoul, South Korea, 2013.

[6] 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.

[7] A. Hunt and R. Kirk. Mapping Strategies for Musical Performance. In: M.M. Wanderley and M. Battier. *Trends in Gestural Control of Music*. IRCAM, Paris, 2000.

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

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

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

[11] F. Lerdahl and C. Krumhansl. Modeling Tonal Tension. Music Perception, 24, 4, pp. 329–366, 2007.

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

[13] J.C. Schacher. "The Body in Electronic Music Performance." In *Proc. Sound and Music Computing Conference*, pp. 194–200, 2012.

[14] J.C. Schacher, Daniel Bisig, and Philippe Kocher. The Map and the Flock: Emergence in Mapping with Swarm Algorithms. *Computer Music Journal*, 38:3, pp. 49–63, 2014.

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

[16] K. Wilkie, S. Holland, and P. Mulholland. Evaluating Musical Software Using Conceptual Metaphors. In *Proc. BCS-HCI,* Cambridge, United Kingdom, September, 2009.

[17] K. Wilkie. S. Holland, and P. Mulholland. What Can the Language of Musicians Tell Us about Music Interaction Design? *Computer Music Journal.* Winter 2010, Vol. 34, No. 4, pp. 34–48, 2010.