A Novel Representation for Rhythmic Structure

Vijay Iyer (1)
Jeff Bilmes (2)
Matt Wright (1)
David Wessel (1)
(1) Center for New Music and Audio Technologies, U.C. Berkeley
vijay, matt,wessel@cnmat.berkeley.edu, http://www.cnmat.berkeley.edu/People
(2) International Computer Science Institute, U.C. Berkeley
bilmes@icsi.berkeley.edu, http://www.icsi.berkeley.edu/~bilmes

Abstract

We have developed a representation for musical rhythm and rhythmic structure based on concepts derived from African and African-American musics. Included in the representation are a model for expressive timing against an isochronous pulse, and a cellular approach to musical organization. In our implementation, the representation and its data structures are controlled and modified in real time using MAX. The richness of control over many meaningful musical quantities distinguishes our representation from those in more common usage, such as music notation programs, sequencers, and drum machines.

1 Introduction

Many authors [2][4][7][8] have attempted to address the universal issue of rhythm perception and cognition. While these efforts feature rigorous approaches to data analysis and modeling, such work often contains musical assumptions specific to Western European classical music, even if the music under study lies outside of that tradition. It is clear that one cannot transcend one's own cultural context in discussing something as culturally contingent as music. Subsequently, the present work embraces the cultural perspective of its authors, and assembles a model based upon their specific musical knowledge and experiences with the musics derived from West African and African-American cultures. The ensuing multiscale representation of temporal structure in music features crucial musical concepts that appear in many musics of the world -- particularly African, African-American, and many Asian musics, but also European classical music to a lesser degree.

Anthropologists, historians, and ethnomusicologists have studied the presence of important historical links and strong conceptual commonalities between West African and African-American cultures, including their musics [11][14]. The best-known among their common musical traits are the prominence of improvisation as a mode of expression, a tendency towards percussivity, a prevalence of antiphonal (call-and-response) relationships, and a developed stratification of contrasting rhythmic layers. This last characteristic is crucial to our work, and it tends to come in tandem with a lower priority for hierarchical organization of large-timescale musical structure. Such music may de-emphasize these hierarchical relationships in favor of referential, associative, or functional relationships [5]. This trait may be contrasted with the notion of deeply recursive intra-musical hierarchies suggested by Lerdahl and Jackendoff [6] for Western tonal music. So, for example, in many examples of reggae, hip-hop, and funk music, the stress might be more on cyclicality, on reference to a shared body of knowledge, or on variable relationships to a composite percussive pattern (see the discussion of "groove" below), and less on the recursive grouping of sections into progressively larger chunks. Thus, many African-derived musical structures can be less "deeply" organized on a large scale (in the sense of the "depth" of a recursive tree) than Western tonal structures, but are frequently "deeper" (i.e. more stratified or polyrhythmic, containing more constituent units) on a small timescale.

Another important feature of African-derived musics is a “bottom-up” cognitive approach to musical organization. For example, rhythmic textures often arise from the superposition of various cyclic musical patterns. A prime instance of this trait occurs in Afro-Cuban *rumba*, which features fixed cyclic clave and wood-block ostinati, relatively stable repetitive low- and mid-range conga drum patterns, and a variable, heavily improvised *quinto* (high conga drum) part, all combining to form an extremely rich emergent texture. This bottom-up approach may also occur at a higher hierarchical level. Musical pieces may have a number of different repeated sections or “spaces” that cycle for arbitrary lengths of time; the transitions among these spaces are often cued in an improvisatory fashion, quite possibly without a preordained large-scale temporal structure or a linear notion of time. The music of James Brown provides many examples of this type. We may characterize these bottom-up methods of musical organization as “cellular,” large musical structures are assembled from small, fully-formed constituent units.

Yet another musical element crucial in African-derived and many other world musics is a concept of groove. This elusive quality may be described roughly as a complex relationship to a collectively-determined and relatively isochronous pulse. In groove-based music, the steady pulse is the chief structural element, and it may be articulated in a complex, indirect fashion. One could say that, among other functions, a groove *gives rise to the perception of lifelike, steady pulse* in a musical performance. In groove contexts, musicians display a heightened, seemingly microscopic (~5 ms) sensitivity to musical timing. Different kinds of rhythmic qualities, such as apparent accents, emotional mood, etc., are created by playing notes slightly late or early relative to their metric location. We claim that this variety of expressive timing against an isochronous pulse contains important information about the inner structure of groove. While numerous studies have dissected the nuances of expressive *ritardandi* and other tempo-modulating rhythmic phenomena [3][10][12], to our knowledge there have been few careful quantitative studies that focus on expressive timing with respect to an isochronous pulse. In groove-based contexts, even as the tempo remains constant, this fine-scale rhythmic delivery becomes just as important a parameter as, say, tone, pitch, or loudness. All these various musical quantities combine dynamically and holistically to form what some would call a musician’s “feel.” A careful study of these kinds of music requires similarly integrative treatment of these parameters.

2 Modeling Expressive Timing

One of the authors [1] has developed a tripartite model for expressive timing in performance of groove-based music. In addition to the salient moderate-tempo pulse or *tactus*, another important rhythmic unit is defined at the finest temporal resolution. It is called the temporal atom or *tatum* (in homage to the great African-American improvising pianist, Art Tatum), the smallest cognitively meaningful subdivision of the main beat. Multiple tatum rates may be active simultaneously. In Western notation, tatums may correspond typically to sixteenth- or twenty-fourth-notes, though they may vary over the course of a performance. Groove-based music is characterized in part by focused attentiveness to events at this fine level. The tactus and the tatum provide at least two distinct clocks for rhythmic synchronization and communication among musicians.

In Bilmes’ scheme, a performance displays musical phenomena that may be represented on three timescales. First, the musical referent or “score” corresponds to the most basic representation of the performed music as it would be notated in Western terms, using quantized rhythmic values (tatums) that subdivide the main pulse. All note-events are represented at this level. Secondly, at relatively large timescale, inter-onset intervals are stretched and compressed through tempo variation. This variation may be represented as a tempo curve -- a function of musical time vs. score time. However, particularly in percussive music, there is no real musical continuum separate from the note-events; score time is quantized in units of tatums. In fact, the tempo curve operates on tatums, modifying their durations such that their sequential sum corresponds to the integral of the tempo curve. In this way, the tatum may be regarded as a sampling rate.

Thirdly, the tatum-relative temporal *deviations* capture the aforementioned fine-grained expressive timing. They quantify the microscopic delays or anticipations of note-events relative to the theoretical tatum onsets. In other words, deviations represent the microscopic values by which note onsets differ from rigid quantization, *over a metronomic background*. Deviations take on continuous values from -0.5 to +0.5 tatums, so that all possible rhythmic placements are allowed. In the case of multiple simultaneous tatum rates, this may allow for a redundant representation, in that a given note-event may be described in a number of different ways. For example, an event may be represented as a deviated sextuplet or as a differently deviated sixteenth-note. We include this ambiguity purposefully, because such ambiguities occur frequently and naturally in the types of music under study, such as Afro-Cuban percussive music.

In his previous work, Bilmes used signal-processing techniques to extract each of these three quantities from a musical performance. The work demonstrated that analysis of deviations can shed some light on musicians' internal representations of the rhythmic content of their performances, particularly in ensemble contexts that feature fixed and variable rhythmic groups.

3 Representation and Implementation

More recently, we have elaborated upon the above model to develop a powerful representation for musical rhythm. Implemented in MAX (a graphical, object-oriented music-programming environment [9]), the representation includes features such as pitch, accent, rhythmic deviations, tempo variation, note durations (which are found to carry important rhythmic information and are therefore treated independently), and probabilistic processes. It may be used in conjunction with MIDI instruments or other synthesizers or sound modules.

To facilitate the use of the representation, we have developed various *Editors* and *Players*, for creating and musically enacting the data structures, respectively. One of our main goals has been interactive music performance, so our Players have been designed with real-time control in mind. A Player agent steps through the data structures, scheduling and playing the note-events. Multiple data structures are handled with ease thanks to MAX's parallel programming model. Players may also improvise by selecting from banks of rhythmic data or by creating new structures in real time. The Editors consist of graphical user interfaces for creating and modifying data structures in the representation.

The basic unit of our representation is the *Cell*, a data structure containing a *Duration*, a *Tempo_curve*, and any number of *Note_layers*. A given Note_layer contains either a discrete, regular *Tatum_grid* whose elements contain Notes, or a list of Notes occurring at fractional points of the cell duration. The presence of multiple Note_layers allows a rich variety of rhythmic possibilities at the most fundamental level, including multiple tatum rates, "a-tatum" rhythms [1], hierarchy, and stratification. A Note is a tuple containing data about the note type, loudness/velocity, note duration, and deviation (in the discrete case). Thus, expressive timing against a metronomic background stands on equal footing with other continually-modulating musical parameters.

Cells may be combined either in series or in parallel into larger Cells, or they may be cycled indefinitely. The system employs a novel method for handling large numbers of complex rhythmic structures, using features of the MAX collection object, a versatile and open-ended data structure. This method facilitates many of the bottom-up combination techniques mentioned above, such as the stratification of different-length rhythmic cycles, creation of composite beat schemes, repetition, rhythmic progression, and most importantly, improvisatory manipulation of these structures. Thus, a user may create a number of different Cells and select rapidly from among them in real time, superimposing, serializing, or otherwise manipulating the constituent units. Note that the overall design privileges hierarchy at the intra-cellular level, and emphasizes "heterarchy" or modularity at the multi-cellular level. This prioritization arises from African-derived concepts of musical organization, as described in the introduction.

Among the various tools for manipulating the data structures, we provide a way to exaggerate or de-emphasize rhythmic features by the use of non-linear (power-law) compression and expansion. This technique applies most effectively to deviation and accent data, where subtle expressive features may be either softened or enhanced via a continuous controller.

The richness of control over many meaningful musical quantities distinguishes our representation from those in more common usage, such as music notation programs, sequencers, and drum machines. In addition, as mentioned above, the representation supports creative applications in improvised performance with electronic instruments.

4 Future Work

We intend to use our implementation to collect rhythmic data from musicians for analysis, and to develop hypotheses and models of rhythm cognition. Although we will not go into the details here, we have also begun to employ probabilistic processes to construct a useful preliminary representation of rhythmic improvisation. If used wisely and in conjunction with careful treatment of other parameters, these controlled random processes can yield musically interesting output. Lastly, we have begun to develop more sophisticated Players that incorporate ideas about the body, kinesthetics, and embodied cognition [13].

References

[1] Bilmes, J. 1993. *Timing is of the Essence: Perceptual and Computational Techniques for Representing, Learning, and Reproducing Expressive Timing in Percussive Rhythm*. Masters' thesis, Massachusetts Institute of Technology.

[2] Clynes, M. and J. Walker, 1982. "Neurobiologic Functions of Rhythm, Time and Pulse in Music," In Clynes, M., Editor, *Music, Mind, and Brain*. New York: Plenum Press, pp. 171-216.

[3] Desain, P. and H. Honing, 1996. "Physical Motion as a Metaphor for Timing in Music: The Final Ritard," In Proc. ICMC, Hong Kong, pp. 458-460.

[4] Fraisse, P. 1982. "Rhythm and Tempo," In Deutsch, D., Editor, *The Psychology of Music,* New York: Academic Press, pp. 149-180.

[5] Honing, H. 1993. "Issues in the representation of time and structure in music," *Contemporary Music Review 9*, pp. 221-239.

[6] Lerdahl, F. and R. Jackendoff, 1983. *A Generative Theory of Tonal Music*. Cambridge: MIT Press.

[7] Longuet-Higgins, H. and C. Lee, 1982. "The Perception of Musical Rhythms," *Perception,* Volume 11, pp. 115-128.

[8] Longuet-Higgins, H. and C. Lee, 1984. "The Rhythmic Interpretation of Monophonic Music," *Music Perception,* Volume 1, Number 4, pp. 424-441.

[9] Puckette, M., 1991. "Combining Event and Signal Processing in the MAX Graphical Programming Environment," *Computer Music Journal*, Volume 15, Number 3, pp. 58-67.

[10] Repp, B., 1990. "Patterns of Expressive Timing In Performances of a Beethoven Minuet by Nineteen Famous Pianists," JASA Volume 88, pp. 622-641.

[11] Southern, E., 1983. *The Music of Black Americans*. New York: W. W. Norton & Co.

[12] Todd, N., 1989. "Towards a Cognitive Theory of Expression: The Performance and Perception of Rubato," CMR Volume 4, pp. 405--416.

[13] Varela, F., E. Thompson, and E. Rosch, 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MIT Press.

[14] Wilson, O., 1974. "The Significance of the Relationship between Afro-American Music and West-African Music," in *The Black Perspective in Music* 2, pp. 3-22.