A Theoretical Framework for Electro-Acoustic Music

Mary H. Simoni (msimoni@umich.edu), Benjamin Broening (broening@umich.edu),
Christopher Rozell (crozell@umich.edu), Colin Meek (meek@umich.edu), Gregory H. Wakefield (ghw@umich.edu)
Center for Performing Arts & Technology, School of Music – College of Engineering
MusEn Project: http://musen.engin.umich.edu
University of Michigan
Ann Arbor, MI 48109-2085
USA

Abstract

Developing a theoretical framework for electro-acoustic music presents an array of problems not present in the analysis of Western tonal and post-tonal music. We present a theoretical framework characterized by the interaction of a perceptual and an analytic model. The perception of a composition is used to inform the analytic model. The analytic model examines aspects of a composition ranging from macrostructures, such as form, to microstructures, such as the spectrum at a moment in time. The theoretical framework is exercised through the study of *Late August* by Paul Lansky.

1. A Brief Discussion of Analytic Methodologies for Western Tonal Music

While intellectual speculation on music can be traced to antiquity, music analysis, as we now understand it, emerged in the eighteenth century [Bent & Drabkin, 1987]. Many of the central concerns of early theorists, including phrase structure and modeling of large formal design, remain central concerns in the analysis of music today. Just as issues of pitch and rhythm were paramount in the early days of music analysis, so too are they similarly privileged in the analysis of contemporary music. Traditional analyses often organize elements of pitch and rhythm hierarchically: motive, theme, phrase, phrase group, and section. Schenkerian approaches delineate foreground, middleground, and background levels.

Attempts to directly apply traditional analytic methodologies to electro-acoustic music are problematic. While electro-acoustic and acoustic music share many musical attributes, these attributes function differently in each genre. How then are we to understand the formal organization of electro-acoustic music? Consider a musical phrase. As William Rothstein says, a phrase must contain "directed motion in time from one tonal entity to another" [Rothstein, 1989]. Since Rothstein's concept of a phrase is so deeply intertwined with tonal processes and thus not applicable to a great deal of electro-acoustic music, we instead think of a phrase as Roger Sessions put it, "the portion of music that must be performed, so to speak, without letting go, or figuratively, in a single breath" [Sessions, 1950]. Just as the notion of a musical phrase changes throughout history, so too must our understanding and analysis of musical structures.

In this paper, we offer suggestions to some of the problems of analytic methodologies for electro-acoustic music. We combine our understanding of music theory with information gleaned from research in signal processing to develop an analytic framework for the analysis of electro-acoustic music. Through this marriage of music theory and signal processing, we are able to enhance our understanding of electro-acoustic music.

2. The Perceptual Model

The analysis of music begins with the process of listening. Our perceptual model is centered upon the experiences of a listener in an ideal stereophonic listening environment (Fig. 1). The listener auditions the composition and discerns the musical elements of the composition. The process of identifying the musical elements in an electro-acoustic composition is borrowed from research in auditory scene analysis. Auditory scene analysis is the process whereby all the auditory evidence that comes over time, from a single environmental source, is assembled by the listener as a perceptual unit [Bregman, 1990].

Bregman observes that the auditory system apparently keeps an 'open slot' available for the return of a sound. His observation may be extended to the perception of music as evidenced by our tendency to listen for the return of previously stated musical events. In the case of auditory analysis of an electro-acoustic composition, these auditory streams may be any musical event such as pitch clusters, a timbre, or a class of timbres.
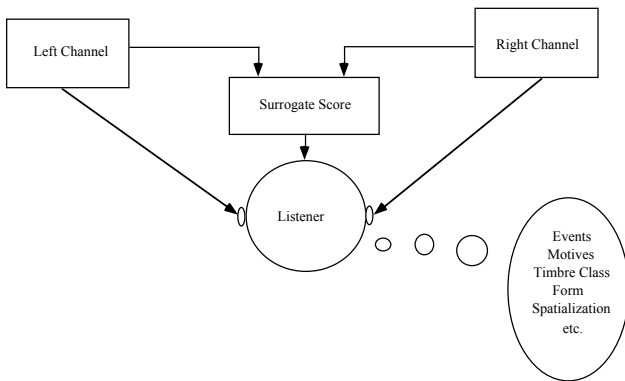
Fig. 1: The Perceptual Model

The listener may be aided by a two-dimensional monophonic time-frequency representation that displays the entire composition. For purposes of data reduction, we view a stereophonic composition as a monophonic signal. In some cases, the listener will require a stereophonic time-frequency representation. The two-dimensional time-frequency representation serves as a surrogate score during the analysis. The surrogate score displays time on the x axis and frequency on the y axis (Fig. 2). The intensity of color at any point corresponds to the energy present in a particular time and frequency region. In our research, the surrogate score assists the listener in organizing auditory streams as a time-ordered series. The listener may mark salient musical events on the surrogate score to assist in the analysis of the composition. These marks indicate time slices that require further investigation in the analytic model.
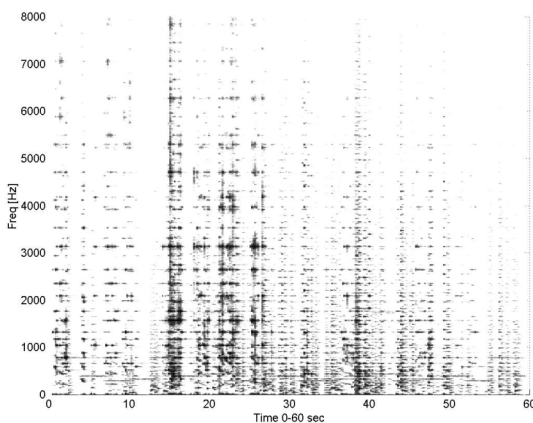

Fig. 2: An excerpt from the surrogate score of
Late August by Paul Lansky
(time = 0:00 – 1:00 minute)

3. The Analytic Model

The listener in the perceptual model informs the processes which take place in the analytic model (Fig. 3). Musical events perceived by the listener in the perceptual model are identified and classified. The classification determines if knowledge of music theory or signal processing should be employed to advance the analysis. Musical intelligence required by an analysis includes a well-developed ear for aural melodic and harmonic analysis and a thorough understanding of music theory. Signal processing techniques that assist in the investigation are spectrum analysis and pitch tracking.
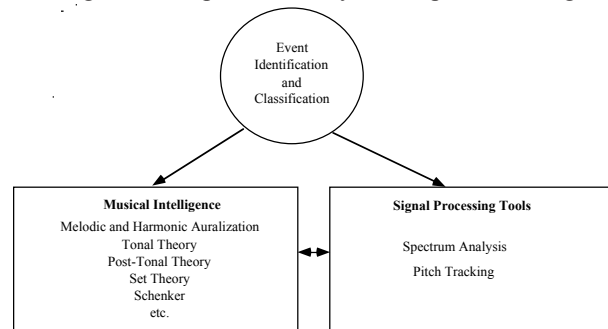

Fig. 3: The Analytic Model

The interaction of the perceptual and analytic models is best characterized as recursive. For example, information gleaned from the perceptual model informs the selection of techniques employed in the analytic model. Outcomes of the analytic model may suggest additional listening to refine the processes of the analytic model.

3.1 Musical Intelligence

The skilled listener makes observations regarding musical attributes such as phrases, timbre, tonal centers, and form. These observations are marked on the surrogate score and become identified with a particular time slice or point in time. These markings assist in the identification of motives, which may be timbres or classes of timbres, phrases, and sections. The markings indicate which time slices of a composition should be further investigated using signal processing techniques such as time-frequency analysis and pitch tracking.

3.2.1 Time-Frequency Analysis

We are interested in first-order trends in the time-frequency representation of the signal and use a spectrogram to these trends. To produce a spectrogram of a time slice, the audio signal is split into overlapping segments that are windowed using a Hanning window. The short-term, time-localized frequency content of the audio signal is found by taking the discrete Fourier

transform of each segment. A three-dimensional mesh plot of the resulting data is produced showing the frequency content of the signal as it changes over time. Specific parameters for the Fourier transform such as window size, window overlap and viewable frequency range are adjusted depending on the resolution needed for the task at hand.

3.2.2 Pitch Tracking

Additional signal processing techniques may be applied to sections of electro-acoustic music that are comprised primarily of pitched events. The pitch class of a note may be determined using pitch recognition software called *PTrack* [Wakefield & Pardo, 1999]. The term pitch class [Forte, 1973] refers to the successive numbering of pitches using integers. For the purposes of this paper, we use [0 1 2 3 4 5 6 7 8 9 t e] to represent pitches [C C# D D# E F F# G G# A A# B]. The pitch content of the audio signal is calculated in a manner similar to the process described in 3.2.1. In each segment, frequency components that have a strong presence in the signal are separated into a series of twelve bins with each bin corresponding to each pitch of the chromatic scale. For example, a signal with frequency components at 110 Hz, 220 Hz, 440Hz would place all frequency components in the "A" bin. Each pitch is investigated to ascertain the degree of harmonically-related partials present in the signal. The presence of harmonically-related partials is assumed to be evidence of the existence of that pitch class in the signal. The pitch class with the greatest evidence for existence is selected as the single-pitch summary for that time segment. By assembling these summaries, an approximation of the pitch classes present in a given time slice may be obtained. A plot referred to as a chromagram displays time as the x-axis and pitch class as the y-axis.

4. Case Study: *Late August* by Paul Lansky

Paul Lansky's *Late August* [Lansky, 1990] is constructed from three timbre classes: processed speech; a plucked, percussive timbre; and a sustained choral timbre. The composition exhibits ten clearly defined tonal centers. Since the most prominent sonic elements of the composition are pitched, we draw on music theory literature that deals with pitch, applying some of the set-theoretical concepts developed by Milton Babbitt and Allen Forte [Babbitt 1961, Forte 1973].

The principal pitch-class set in *Late August* is set 5-35 [02479] also known as a pentatonic collection (Fig. 4). The two speakers whose conversation Lansky processed for *Late August* are Chinese, and Lansky chose the pentatonic collection because of its association with the East [Lansky 1999].
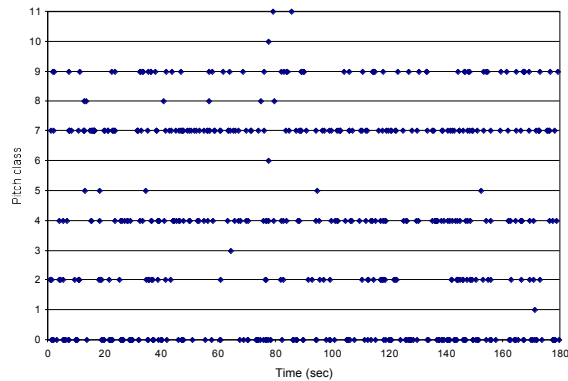

Fig. 4:   A chromagram of *Late August* by Paul Lansky clearly depicting the presence of set 5-35

Figure 5 shows a summary of the ten distinct tonal regions of the 13'45"composition and the corresponding set theory analysis.

Section	Start Time	Pitch-classes	Pitches	Number of tones in common with set 5-35	Tonal Center	Transpos-ition	Interval Class of Trans-position	Number of Occurrences of Trans-position
1	0:00	{02479}	C, D, E, G, A	5	C	T0	0	3
2	3:01	{t0257}	Bb, C, D, F, G	3	Bb	Tt	2	3
3	4:00	{357t0}	Eb, F, G, Bb, C	2	Eb	T3	3	2
4	5:07	{02479}	C, D, E, G, A	5	C	T0	0	3
5	6:15	{t0257}	Bb, C, D, F, G	3	Bb	Tt	2	3
6	7:15	{57902}	F, G, A, C, D	4	F	T5	5	1
7	8:15	{357t0}	Eb, F, G, Bb, C	2	Eb	T3	3	2
8	9:15	{8t035}	Ab, Bb, C, Eb, F	1	Ab	T8	4	1
9	10:15	{t0257}	Bb, C, D, F, G	3	Bb	Tt	2	3
10	11:15	{02479}	C, D, E, G, A	5	C	T0	0	3

Fig. 5: An Overview of *Late August* by Paul Lansky

The principal set's Interval Class Vector (ICV), 032140, offers a possible explanation for Lansky's choice of transpositions. Only a limited subset of the twelve possible transpositions is presented: T_0, T_3, T_5, and T_t. These transpositions correspond to the intervals available from within the set. The exact level of transposition corresponds to the inversion of the principal set:

Principal set:	0	2	4	7	9
Inversion:	0	t	8	5	3
Inversion reordered:	0	3	5	8	t

	0	3	5	8	t
Section No. where transposition appears	1, 4, 10	3, 7	6	8	2, 5, 9

The decision to dominate the pitch language with set class 5-35 creates a sense of homogeneity that is further strengthened by Lansky's choice of transposition levels. Pitch class 0 is present in each section of the composition. Each section has at least two tones in common with its preceding section. As a result, adjacent sections have a high degree of continuity. Section eight is the point of greatest tonal distance from the tonic, but still retains pitch class 0.
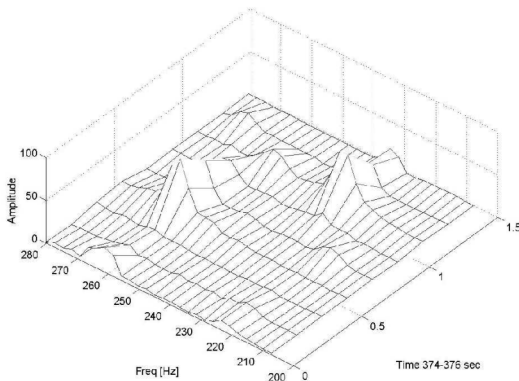

Fig. 6: A spectrogram showing the shift in tonal center from C (265 Hz) to Bb (233 Hz) at 6'14"-6'15". A linear amplitude scale is used to prominently display the change in pitch.

5. Summary and Considerations for Future Work

We have described an analytic process characterized by perceptual and analytic models that assist in our understanding of *Late August* by Paul Lansky. The visualization of musical signals augments, but certainly does not replace, what is perceived by the ear. As John Tyndall stated in 1875, "in the width of perception the ear exceedingly transcends the eye" [Tyndall, 1877]. Similarly, we have observed that a well-trained ear should guide the selection of signal processing techniques that are applied to an analysis. Furthermore, a skillful interpretation of that analysis is required to make meaningful assertions about the music.

Although visualization of musical signals is useful is the analysis of electro-acoustic music, it has its limitations. *Ptrack* is a single pitch estimator for a particular time segment and for that reason, is incapable of accurately representing segments where more than one pitch is present. Future work will investigate extensions of *Ptrack*, or the development of alternative approaches, so that we can extract multiple simultaneous pitches.

References

Babbitt, Milton. (1961). "Set Structure as a Compositional Determinant," *Journal of Music Theory*, Vol. 5.

Bent, Ian and Drabkin, William. (1987). *Analysis.* MacMillan Press, Houndsmills, Basingstoke, Hampshire.

Bregman, Albert S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound.* MIT Press, Cambridge, MA.

Forte, Allen. (1973). *The Structure of Atonal Music.* Yale University Press. New Haven, CT.

Lansky, Paul. (1990). *Late August.* New Albion. San Francisco, CA.

Lansky, Paul. (1999). Personal communication.

Rothstein, William Nathan. (1989). *Phrase Rhythm in Tonal Music.* Schirmer Books, New York, NY.

Sessions, Roger. (1950). *The Musical Experience of Composer, Performer, and Listener.* Princeton University Press, Princeton, NJ.

Tyndall, J. (1877). *Sound.* Third Edition. D. Appleton & Co., New York.

Wakefield Gregory H. and Pardo, Brian. (1999). "Signal Classification using Time-Pitch-Chroma Representations.*" Proceedings of the SPIE'99*. Denver. CO.