SYSTEMATIC AND QUANTITATIVE ELECTRO-ACOUSTIC MUSIC ANALYSIS (SQEMA)

Tae Hong Park, David Hyman, Peter Leonard, Wen Wu

Tulane University
Music Department

ABSTRACT

In this paper we present the SQEMA (**S**ystematic and **Q**uantitative **E**lectro-acoustic **M**usic **A**nalysis) methodology. The primary aim of SQEMA is to create a comprehensive framework to assist in the analysis of electro-acoustic music. Our proposed methodology is based on two main strategies: (a) exploitation of MIR and salient feature extraction techniques and (b) employing a systematic analysis paradigm to segment a complex piece of music into smaller and more manageable parts. Our initial studies show potential in using MIR techniques which seem to be underemployed for electro-acoustic music analysis. The literature in electro-acoustic music analysis provides a plethora of promising methods. However, the research also seems to be plagued by lack of consensus, confusion in nomenclature, and absence of a standard modus operandi, as far as strategies for electro-acoustic music analysis are concerned. We propose a methodology that attempts to provide a framework to address some of the issues found in existing fragmented systems, while incorporating useful portions of existing techniques. SQEMA employs a stepwise segmentation approach focusing on techniques that yield quantifiable information pertaining to form and content. One of its primary goals is informing aesthetic interpretation. In this paper we undertake a thorough analysis of *The Machine Stops* using our proposed methodology.

1. INTRODUCTION

The motivation for creating the SQEMA methodology arose from our response to an apparent lack of "standard techniques" for electro-acoustic music analysis and our on-going research & development of the EASY software system [8]. The current literature in electro-acoustic music analysis reveals a dearth of quantitative techniques and a lack of standardization on many levels. There could be several reasons for these problems. Firstly, electro-acoustic music is a young genre. As Mary Simoni writes, "The information age in music making has flung these composers squarely back to a period in human creativity that rivals that of the medieval period" [5]. This state of affairs is manifested in the frequent absence and enormous variety of scores or visual representation formats for electro-acoustic music, as well as a lack of consistency in nomenclature. Even the term electro-acoustic music has different spellings and names including computer music, electronic music, musique concrète, tape music, and acousmatic music. Secondly, this music frequently focuses on timbral features rather than more traditional compositional paradigms, which rely heavily on pitch, harmony, and rhythm. Existing analysis methods, which hail from the world of "classical music," are vastly inadequate for the task of unmasking musical structure in electro-acoustic music. This is especially true for electro-acoustic music falling into the tape music category where timbre is often of central concern. Timbre, unlike the other basic components of sound (pitch, dynamics, and time) is not yet well understood and is subject to much confusion [7]. These issues make the analysis of electro-acoustic music especially difficult and force the listener to rely on the sound itself often with little or no other form of reference to the piece.

We hope to begin addressing the need for a standardized electro-acoustic music analysis framework which focuses on techniques that yield quantifiable data utilizing Music Information Retrieval (MIR). Current research in electro-acoustic music analysis does not yet seem to exploit MIR techniques. This is somewhat surprising as not utilizing these techniques can loosely be likened to not taking advantage of an available automatic pitch-tracker or chord transcriber when studying tonal music in the absence of a score. Since so much of electro-acoustic music is created with the assistance of complex technologies, it seems that we should also take advantage of the most current available technologies for exploring this music. To this end we build on our development of the EASY toolbox, a software system employing MIR techniques to assist in electro-acoustic music analysis. The EASY system allows for several multidimensional visualization formats of extracted sonic features and also employs segmentation and clustering algorithms to gain insights into a piece. As we will show later in the paper, this allows us to locate sonic relationships which are sometimes very difficult to identify even when closely listening to a piece.

SQEMA employs a stepwise, top down approach to the decomposition portion of analysis. After several listenings to familiarize oneself with a piece, the process moves top-down as follows: from high-level segmentation to mid-

level segmentation and finally to the identification and analysis of events. Once all of this information has been extracted from the music, SQEMA moves into a bottom-up, side-to-side, and diagonal reexamination phase. This includes identifying motives and musical streams and determining formal structure. Finally, we apply the accumulated information to assist in aesthetic interpretation.

We test SQEMA in performing a thorough analysis of *The Machine Stops* by Tae Hong Park. This short piece explores gradual timbral change, covering the musical area between a single sinusoid and noise. It offers a rich palette of features to examine in both the time and frequency domains and, therefore, presents ample opportunity to assess the efficacy of the SQEMA methodology. We will begin with a brief survey of existing electro-acoustic music analysis methods. We will then discuss the SQEMA methodology in detail, followed by our analysis results for *The Machine Stops* using our methodology. Finally we will discuss our conclusions and future work toward further developing SQEMA.

2. BRIEF SURVEY OF EXISTING ELECTRO-ACOUSTIC MUSIC ANALYSIS TECHNIQUES

Our research covered a wide range of writings on electro-acoustic music analysis, with focus on contemporary essays that attempt to provide formal techniques for electro-acoustic music analysis. The following issues appear frequently throughout the literature: nomenclature, visual representation, aural perception, and generality/specificity. Our goal in this section is not to downplay the important contributions of the authors, but rather to note some palpable problems which our methodology tries to address.

2.1. Nomenclature

One of the greatest barriers to an efficient and coherent discussion of these matters is the issue of nomenclature. Mark Applebaum articulates the problem humorously in his 8 channel speech-based piece *Pre-composition*, in which we hear him describing electronic sounds with words like, "nasal filtered thing" and "like a swooshing sound" [1]. Across our readings, we found scarcity of agreement on the definitions of various terminologies and a misuse of others. Some words are attributed multiple meanings; at other times similar words are used to mean different things. For example, although Lelio Camilleri states that one of his goals is "the creation of an analytical lexicon" [2], he cites Smalley's noise-note continuum, the midpoint of which is the node. Smalley defines the node as "an event having a more complex texture than a single pitch" [14]. Unfortunately, the word node, borrowed from acoustics, already has a meaning as the point in a standing wave where there is ideally no motion. In fact, nodes are

part of any resonant system, which includes the "note" as utilized by Smalley. In this particular case, it seems that changing the meaning of a word from acoustics is perhaps more in the spirit of alliteration rather than precision of meaning. As electro-acoustic music is heavily technology driven, many of these "new" words are borrowed from science and engineering. Therefore, care must be taken to maintain as much clarity as possible to avoid ambiguity and perhaps even avoid adding confusion to the already complex field of electro-acoustic music analysis.

On the other hand we found certain nomenclature quite valuable and well defined such as the concept of reduced listening, the notion of the sound object and source bonding [10,15]. The codification of a lexicon geared toward the particular issues present in electro-acoustic music will be important in the development of a comprehensive system for electro-acoustic music analysis.

2.2. Visual Representation

An important difference between "classical music" and electro-acoustic music is the lack of standardized scoring techniques for the latter. In the absence of a score, waveforms and spectrograms can provide important information encoded in musical signals, and can be useful in the analysis of any music. Several extensions of signal analysis such as David Hirst's use of MQ plots [6] and Mara Helmuth's multidimensional analysis method [5] begin to draw correlations across the frequency and time domains. Both of these methods, however, focus on the outer layer of the spectrogram/waveform and do not address encoded information.

Pierre Couprie presents an insightful discussion of graphical representation, praising the effectiveness of iconic representation where "the link between graphical qualities and the sound criteria they represent are relatively intuitive" [4]. A great example of this type of technique can be seen online in Rainer Wehinger's rendering of a visual listening score accompanying György Ligeti's *Artikulation* [18]. Couprie hints at future work which would incorporate symbolic and iconic representation into one paradigm. Additionally the EASY toolbox offers methods for visualization of timbral features [8].

2.3. Aural Perception Analysis Techniques

Due to the lack of standardized nomenclature and visual representation formats, other than the waveform and spectrogram, much of electro-acoustic music analysis relies heavily on aural perception. For our purposes, the innate subjectivity of perception and the current difficulty in accurately quantifying perceptual data places severe limitations on using only a perceptual approach for analysis and research. Schaeffer's typo-morphology and Smalley's spectro-morphology both rely on aural perception. This research is a necessary part of building the descriptive language of electro-acoustic music and sound

in general, but the outcomes are currently too subjective to be relied upon for a high level of clarity and objectivity in analysis. Moving toward standardized methods for quantitative sound description is a necessary part of a fully realized electro-acoustic music analysis system. As these methods develop, they must be incorporated into the larger frameworks. Other interesting work is being done in the automatic detection of mood and emotion in music [16], but this important research is still in its earliest phases. The perceptual approaches hold great value in the aesthetic evaluation and interpretation of a piece. However, they will hold even greater significance when backed by quantifiable data.

2.4. Comprehensiveness: Specificity and Generality

The most pressing problem noted in our research revolves around the fact that none of the approaches delineated in the readings presents a comprehensive system for analyzing a piece of electro-acoustic music. What we discovered is a fragmented set of tools, each of which deals, whether effectively or ineffectively, with only a portion of the analytical process. On the one hand techniques such as MQ plots are too specific to allow for thorough readings of complex music; on the other hand the more subjective frameworks such as spectro-morphological description are too general to offer tools for discovering quantifiable information. The SQEMA methodology, presented in detail below, aims to bridge this gap by approaching electro-acoustic music analysis in a comprehensive manner by addressing specificity and generality.

3. SQEMA: METHODS AND PROCEDURES

3.1. Core Concepts

The proposed strategy for electro-acoustic music analysis begins with a top-down approach using feature extraction methods whenever possible. More details are revealed as we progress downward through the analysis levels. We are currently limiting our analysis to music that does not focus on pitch, harmony, and rhythm. For those types of pieces which can be analyzed with more traditional techniques, there is already ample literature. Although there is no clear evidence at this time that our proposed method is "better" than other existing methods, let alone that it will work for every piece, the consensus by our group was that some type of a top-down model should be employed when analyzing a musical work. This motivated us to begin drafting a systematic architecture reflecting a procedural analysis model utilizing a rigid, top-down, less detailed to more detailed feature extraction strategy. This strategy loosely adheres to a divide-and-conquer analysis paradigm. In the penultimate phase, which we call reexamination, one makes associations, connections, and correlations via an intra/inter-level analysis, combining top-down, bottom-up,

diagonal, and side-to-side omnidirectional comparisons. These comparisons are, wherever possible, based on measurable extracted features from previous levels. At the final analysis phase we also take into account the aesthetic aspect of music analysis. This step allows for greater liberty when interpreting a given composition. By advocating the use of quantitative results to inform qualitative discussions, we hope to promote coherent, consistent, and reasoned interpretations of a given work.

3.2. Feature Extraction and MIR

To automatically extract quantifiable and objective information from music, we utilize the EASY system which is based on MIR techniques. Although the EASY system is still in its early development phase, it already provides numerous objective feature measurements and also provides helpful data visualizations that reveal encoded sonic structures on a micro and macro level. The long term goal is to fully exploit research results from MIR and apply them to electro-acoustic music analysis including analysis/machine learning techniques from automatic musical instrument recognition [9], genre classification [17], emotion detection [16], and rhythm analysis [11].

3.3. Implementation and Methodology

The entire analysis procedure is divided into six large steps as shown in Figure 1 and includes: I. Multiple Listenings, II. High-Level Analysis, III. Mid-Level Analysis, IV. Event Level Analysis, V. Reexamination, and VI. Aesthetic Interpretation. At levels II, III, and IV quantitative analysis is conducted via feature extraction techniques, clustering results from feature vector spaces, and observation of visualizations of extracted information. The features listed in tables 1 and 2 are by no means exhaustive and much work needs to be done to address correlation of feature vectors to musical attributes.
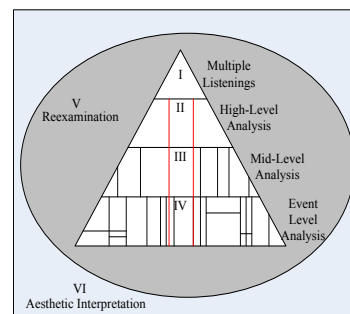
Figure 1. SQEMA model

The top-level analysis procedure simply involves multiple listenings to the piece to familiarize oneself with the music. The first listening is conducted without visual aids (eg. waveform, spectrogram, or score if available) and

subsequent listenings can be done with the aid of visuals. The first step is to encourage deep and focused listening from a purely aural perspective without visual sensory influence as suggested by Erik Christensen [3].

Aim
Initial Section Segmentation
Procedure
-*Energy levels:* time envelope, crest factor, silence detection
-*Frequency distribution*: spectral spread/roll-off/centroid, low freq. energy, etc.
-*Texture*: spectral smoothness/flux, noise content, MFCC, etc.
-*Frequency continuum*: rhythm/pulse, roughness, pitch, frequency sensation, zero-crossing rate, etc.

Table 1. Level II: High-Level Analysis

The High-Level Analysis phase is summarized in Table 1. Its aim is to segment the piece into formal sections by analyzing salient features pertinent to energy levels, frequency distribution, texture, and frequency continuum. The Mid-Level Analysis procedure is fundamentally based on similar techniques used in the previous analysis level. In the mid-level phase, however, we concentrate on each individual section (eg. A-B-A) obtained in Level II and identify *subsections*, *divisions*, and *events*. Divisions are segments that do not have sufficient saliency or contrasting features to be subsections. Events are salient sonic occurrences that are hierarchically below the section, subsection, and division level such as a staccato wideband percussive sound with modal characteristics. Events, however, do not necessarily have to be limited to "short" occurrences and can, therefore, take on the role of a division, subsection, an entire section, or, in extreme cases, even an entire piece. The determination and assignment of labels to extracted sonic occurrences is conducted in phases IV and V.

Aim
Event Identification
Procedure
-Use Mid-Level Analysis results and analyze for event features and characteristics.
-Apply standard nomenclature in describing each event whenever possible.
-Analyze *density*: sonic events per given timeframe via onset analysis.
-Analyze *emphasis/saliency*: dynamics, register, timbre, temporal characteristics, and grouping/gestalt

Table 2. Event Identification

In level IV, Event Level Analysis, the primary aim is to analyze the identified events using clear nomenclature. For example, when describing an event that has a sense of pitch, one can elaborate on its harmonic structure (odd, even, missing fundamental, harmonic expansion/ contraction, jitter/shimmer characteristics, etc.) or its modal resonant structure. Other examples include investigating glissandi events by determining their rate of change/slope, span in Hz, frequency continuum between pulse, roughness, pitch, and sense of high frequency. Section 4.4 shows more examples of how we proceeded in analyzing events.

The next step in our analysis is Level V, Reexamination. Here we take into account all components (sections, subsections, divisions, and events) and analyze them in intra and inter-level cross-examinations. At this stage it is possible to find motifs, streams, and groups of events that have single or repetitive characteristics. This will help us identify developmental ideas, various structures such as retrogrades, inversions, and fragmentation.

The final phase of our analysis is Level VI: Aesthetic Interpretation. Using all the information from previous analysis levels one can now interpret the piece and render narrative and aesthetic elaboration. At this point in the analysis process, we acknowledge that our current systematic, data-driven methodology is, by itself, insufficient for delivering a comprehensive aesthetic interpretation. In order to accomplish a balanced musical analysis, technical detail-oriented analysis techniques should be combined with more subjective approaches. We believe that our methodology, as presented above, offers tools for analysis with the potential to yield the detailed information needed to support coherent connections and articulate aesthetic interpretations about a musical piece.

## 4. PRELIMINARY ANALYSIS RESULTS

4.1. Level I: Multiple Listenings

Following the steps laid out in the methodology, we began the process of analysis by listening to *The Machine Stops* a number of times without the use of visual aids. This is helpful even when one has prior familiarity with a piece. Upon second and third hearings of the piece, a general structure started to emerge – the piece seemed to contain three large sections.

4.2. Level II: High-Level Analysis

Our initial, High-Level Analysis of the piece yielded the following three sections with the aid of the spectrogram and waveform views in EASY: In the first forty seconds one hears a steady and gradual build-up in energy; this is followed by a sudden, dramatic shift in sonic character with the introduction of a sinusoid of static frequency and amplitude which creates a strong sense of sectional demarcation. After the sinusoid modulates through a number of variations, concluding in a dense, broadband section, a sinusoid reemerges creating a sense of transition;

a final, third section of the piece seems to repeat material from the beginning section but in retrograde, with variations.
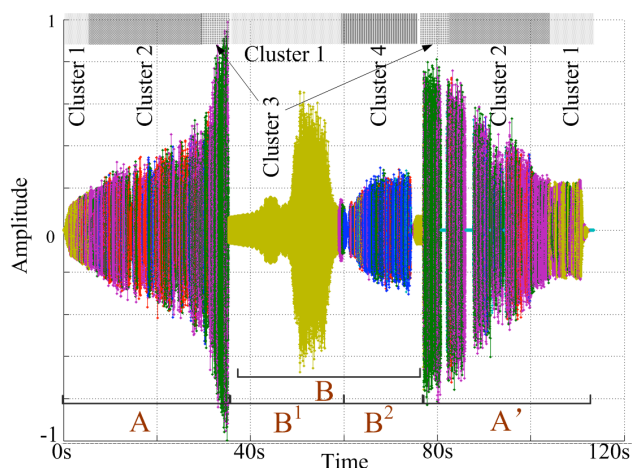
Figure 2. Clustering using EASY toolbox

Next we analyzed the piece using clustering/segmentation algorithms included in EASY. We ran the program multiple times using different feature dimensions. In line with our initial hypothesis, the following two trends emerged among the various analyses: the piece seemed to be comprised of either three or four sections. Referring back to one specific analysis via clustering, in which we used the parameters of spectral centroid, spectral roll-off, and spectral flux with six centroids, Figure 2 was produced which coincided well with what we had determined aurally.

Figure 2 shows that the piece has four sections rather than three. However, instead using four sections and labeling them A, B, C, and D, we decided to segment the piece into three sections, A, B, and A', based on results from level I and data generated in level II. The sections are demarcated as follows: A (0:00-0:40), B (0:40-1:27), and A' (1:27-2:07). The explanation for this conclusion will be reinforced in the next analysis level, where subsections and divisions are determined. This shows an important aspect of the methodology – information gained in later analysis levels can help fine-tune our analysis.

4.3. Level III: Mid-Level Analysis

In the Mid-Level Analysis phase we focus on each individual section (A, B, and A') to determine subsections and divisions, followed by identification of events. In section A, it was difficult to determine any clear subsections aurally, possibly due the section's gradual evolution with few abrupt changes. However, using the segmentation/clustering algorithms in EASY we were able to uncover a subtle, contrasting section at the beginning of A. Due to the subtlety of this segmental delineation, we

decided to denote the segment as a division rather than a subsection.

Next we continued on to identifying events within section A. The first evident event in A spans the entire length of the section and is an ascending, glissando-like sound mass, as corroborated via the spectral centroid. The second notable event in section A occurs at the beginning of the piece and can be described as a gradual shift from fragmented to more continuous sonic activity. The third salient event is comprised of a large number of staccato sounds with modal spectral characteristics.

As noted by the clustering output, section B (0:40-1:27) can be further segmented into two subsections: B^1 (0:40-1:13) and B^2 (1:13-1:27). The reasoning for this sub-segmentation can be explained by the substantially salient events in the beginning portion of B, which contrast the events found at the latter part of B. We determined that section B^1 can be further segmented into two divisions: B^1_i and B^1_{ii}, as delineated by clear events. The beginning of B^1_i is defined by its stark contrast to section A and begins with the introduction of a sinusoid (actually two sinusoids of identical frequency). A second event appears when the two sinusoids begin to exhibit beating patterns via detuning. Soon after, the appearance of overtones of the sinusoid marks the entrance of the third event. Section B^1_{ii} begins at the moment when the beating frequency shifts from the frequency range of rhythm to the roughness range. This is followed by another salient event – the sweeping of the sinusoids from DC to the Nyquist frequency and back down again. This sweeping activity increases in frequency until a "bouncing effect" is produced by the sweeping sinusoids. This "bouncing effect" marks the beginning of B^2. The frequency of the bouncing increases until the overall sonic atmosphere enters the area of roughness with the addition of pitched, contrapuntal glissandi. The final event in this section, the reintroduction of a sinusoid, can be seen as a transition to section A'.

Analyzing for subsections/divisions in section B, event identification and grouping became critical for defining the specific borders between subsections/divisions. As previously mentioned the segmentation/clustering algorithms in EASY generated four sections. At Level I, however, we hypothesized three sections. The final determination of section B was best explained by further segmentation into B^1 and B^2 as elaborated above. In a sense this is a compromise between the computational results and aural analysis.

Aural analysis shows that section A' is in essence a retrograde version of A (this was later confirmed by the composer). In contrast to section A, however, two distinct divisions were identified in A': A'_i and A'_{ii}. A'_i is denoted by four, dense blocks of broad bandwidth sound texture in the pulse/roughness range, followed by absolute silences (1:27-1:47). A'_{ii} is marked by a decrease in density and the entrance of filtering suggesting textures of the voice (1:47-

2:07) – it was confirmed by the composer that a type of formant filter was applied.

4.4. Level IV: Event Level Analysis

Observing the events identified during Mid-Level Analysis in closer detail, plus identifying new, additional events of shorter duration and/or bandwidth, offers critical information that is used at the reexamination level. For all of the events identified above, we extracted all data available using the EASY System. The amount of information extracted via event analysis is too great to detail here since it covers precise frequency measurements of sinusoids, harmonic relationships, modal resonant structures, shimmer/jitter characteristics, noise-content via inverse LPC filtering, and precise timing measurements including attack times. Also, though it is not applicable to this piece, we would at this level identify sound objects. For example, an attempt to determine sources used in musique concrète such as train whistles or crackling embers, or traditional instrumentation would occur here.

4.5. Level V: Reexamination

At Level V we perform intra-level and inter-level cross-examinations. We begin our reexamination by comparing components to one another within the same level, starting with those closest in time/frequency, then moving on to components not directly adjacent to one another. Following this step, components at different levels should be cross-examined. The goal is to determine relationships, associations, and correlations, which can lead to identifying motives, streams, phrases, passages, and other groupings.

4.5.1. Intra-Level Cross-Examination

During Intra-Level Cross-Examination, components of the same level are cross-examined for connections. The following is a list of intra-level connections that were most important to the thematic analysis of *The Machine Stops*.
(a) One of the most salient intra-level relationships in the piece is the quasi-mirror image of A presented in A'. A', as mentioned, presents a retrograde version of A with variations; by extension, all events in section A have retrograde siblings in A'. For example, the sweeping sound mass which increases in frequency in A, decreases in frequency in A'. (b) One important motivic variation which interrupts the retrograde trend is the reflection of the fragmented material at the beginning of A by the large silences at the beginning of A'. (c) In sections B^1 and B^2, an underlying idea runs through the duration of both, comprised of several adjacent events which combine into a single stream: the sinusoid presented at the beginning of B modulates through numerous variations including beating, clipping, sweeping of frequency, a "bouncing effect," transformation into roughness, concluded by a return to a pure sinusoid, transposed up a perfect 4th. (d) The sinusoid at the beginning of B has a symmetrical relationship to the sinusoid at the end of B which transitions into section A'.

4.5.2. Inter-Level Cross-Examination

At the inter-level cross-examination phase, we compare components of different levels to one another. The following are salient inter-level connections in the piece. (a) The most encompassing motive of the piece is based on an arching shape. The shape appears at various levels throughout the piece: at the Event Level the beating pattern produced by the two sinusoids in B^1_i exhibits a full-wave rectified sinusoidal pattern; the sweeping sinusoids which evolve into the "bouncing effect" exhibit arching shapes in their frequency trajectories; finally, the modally resonant sounds of short duration presented throughout the piece, when examined closely, prove to be miniature ascending and descending glissandi. Additionally, the entire form of the piece resembles the arch shape. (b) Comparing B^1 and B^2 with the entire piece yields an interesting finding: the entire piece, comprised of three large sections, is reflected by the three, distinct B segments (B^1_i, B^1_{ii}, and B^2), with an inversion of dynamics: the entire piece shows a general pattern of large, small, large amplitude, whereas the B sections exhibit a small, large, small amplitude pattern. (c) The modally resonant sounds of short duration, present throughout the piece, exhibit motivic qualities; the reintroduction of these events in separate sections serves to reinforce the thematic connections between sections. (d) One of the most important motives is presented by the frequent reoccurrence of streams which exploit the frequency continuum by demonstrating the perceptual boundaries between rhythm, roughness, and pitch. Examples include the event at the beginning of the piece in which rhythmic gestures transform into roughness, the sweeping sine wave in section B which transforms into rhythm then roughness, and the harmonic sound modulated by jitter in section B, which transforms from exhibiting harmonicity, to inharmonicity to broadband noise. This conceptual theme appears to be central to the piece due to the prominence that these events are given.

4.6. Level VI: Aesthetic Interpretation

In our aesthetic analysis of *The Machine Stops,* we approached the piece without the aid of the composer's notes. Drawing on an anthropomorphic notion that the piece followed the life cycle of a machine (as inferred by the title), we returned to our detailed notes of each level of the piece and attempted to find coincidences between the narrative we inferred and the sonic events which comprise the piece.

For instance, observing the arching form of the entire piece at high-level, we determined a three-section analogy to the life cycle, consisting of birth, youth into maturity, and death. Going deeper, to the Mid and Event Level

Analyses, we determined specific subsections and events which we associated with events in the life cycle. The sonic activity in section B, especially in subsection B^2, seemed to reflect the active behavior associated with youth, or the productive stage of the machine. Section A', we felt, reflected death extremely well, with the implication of frailty presented by long silences, dying gasps reflected by the shorter silences and formant filter, and a final fleeting attempt to hold on to life suggested by retardation of the rate of decay at the end.

5. ASSESSMENT OF PERFORMANCE

5.1. Ground Truth

When assessing the performance of a musical analysis methodology it is helpful to test for ground truth by comparing system-based conclusions about a piece using its score, software patches/programs that were used, or the explanations of the composer. Unfortunately, in many cases none of these is available. Analyzing *The Machine Stops* with the composer Tae Hong Park present, we had access to his comments, programmatic materials, and source code, which were helpful in determining analysis performance.

Beginning by comparing the High-Level Analysis results of the methodology versus the composer's explanations, we were pleased to see that the sectional segmentation coincided closely with the composer's descriptions. At the Mid-Level, the methodology again provided accurate sub-segmentation results according to the composer. At the Event Level, our methodology sometimes revealed unexpected details of which even the composer was not fully aware.

In comparing our aesthetic interpretations with those of the composer, we found many similarities, especially regarding the overall structure in its relation to the narrative. However, certain interpretations which we attributed to the piece were not intended by the composer. This in no way detracts from the performance of our analysis. In the end, the perception of music can be highly subjective. However, when approaching the final step of analysis, doing so with accurate, empirical musical data, as opposed to concentrating mainly on subjective descriptions, will yield more coherent, consistent, and hopefully more objective results.

5.2. Problems

Our proposed methodology is still in its early stages of development and has not yet been extensively tested. However, due to the paradigm of the methodology, namely the exploitation of quantitative strategies to extract salient sonic data, we are hopeful that the methodology will perform similarly across a variety of musical works as far as feature extraction is concerned. We should point out that not all the quantitative data was extracted automatically for several reasons: (a) the EASY toolbox is still under development; (b) machine learning and pattern recognition algorithms have not been fully implemented; (c) we observed potentially useful new features for which extraction algorithms have not yet been implemented to our knowledge. Other issues we encountered during the development of the methodology are outlined below.

5.2.1. Organizational Issues: Data and Time Consumption

The choice to analyze *The Machine Stops* was a practical one, due to the relatively short duration of the piece and the availability of technical and programmatic information from the composer. That being said, following the steps of the methodology, even with a short piece, proved to be a time-consuming endeavour due to the vast amount of data the methodology yields. The hierarchical nature of the methodology prescribes a specific order of operations, by which a piece is sub-segmented down to the Event Level. Clearly, the longer and more complicated a piece is, the longer the process will take.

5.2.2. Involvement of Composer

The involvement of the composer in analyzing his own work with a methodology he helped create may present problems such as biasing the analysis in favor of specific conclusions due to a priori knowledge and emotional engagement. However, we decided to move forward with this piece for the following reasons: We specified that the composer would not reveal any information during the testing phase; the piece was composed over 10 years ago, therefore, many of the details were forgotten or vague; there would be minimal involvement of the composer during the aesthetic interpretation phase; the consensus was that having a strong foundation for ground truth outweighed the risk of potential system bias, and that the necessity for quantitative support would minimize these issues when they arose.

5.2.3. Feature Extraction: Effectiveness and Meaning

Although the EASY system provides 26 feature extraction modules, a number of them proved to be less useful for analysis than we had hoped. The main issue was determining the correlation between feature vectors and musical outcome. This problem not only exists in single-dimensional feature analysis, but is also compounded in multidimensional feature spaces.

6. FUTURE WORK

6.1. Navigable Pyramid Map

The current methodology is based on extracting voluminous information from a piece. However, we do not yet have the means of efficiently organizing, representing, and navigating the complex multidimensional musical

structure of electro-acoustic music. We thus plan to implement a navigable pyramid map interface similar to Figure 1. This interface will allow the user to zoom into any level and component, select components, and automatically display other related components present throughout the piece. This will be achieved via automatic and/or manual tagging, labelling, and annotation strategies, which will hopefully lead to the creation of scores/maps for electro-acoustic music that is not based on pitch and rhythm. Additional visualization methods that can be incorporated into this pyramid map include iconic and symbolic representations and labels/tags derived from automatic mood detection and measurements.

6.2. Improvement of Feature Extraction

We found that one of the greatest challenges in using a data-driven analysis approach based on salient feature extraction techniques was the difficulty in mapping features to musically relevant attributes. Thus, we foresee much work to be done in determining which features are suitable for each specific analysis level. During the testing phase, we found that certain attributes such as "spectral sparseness" and "fullness" would have been useful in gaining more insights into timbral structures. Our current analysis concentrated on timbral feature vectors but showed that spatial cues could potentially offer additional analytical insights. Another feature that holds promise for music analysis is the *dissonance score*, which examines relationships of resonance structures and seems to capture musical tension [12]. We thus plan to develop and implement additional feature vectors that will render useful information for electro-acoustic music analysis.

7. SUMMARY AND CONCLUSION

We have presented a methodology for electro-acoustic music analysis, exploiting MIR techniques to extract quantitative data and providing a methodological top-down approach to the decomposition of sonic information, followed by a reexamination and aesthetic interpretation. We discussed initial results obtained by applying the methodology to *The Machine Stops*, detailing the findings of our analysis, problems with the methodology, and future work.

Although much work remains to be done, we feel carefully optimistic that the SQEMA methodology provides an alternative approach to electro-acoustic music analysis by incorporating data-driven and aesthetic paradigms into one comprehensive methodology. We are especially excited by the prospect of moving toward automation and improvement of our methodology by continually fine-tuning it via further testing and collaborative efforts.

8. REFERENCES

[1] Applebaum, Mark. "Pre-composition", *Intellectual Property* (CD), Innova, 2004.

[2] Camilleri, Lelio. "Electro-acoustic music: analysis and listening processes" (Unpublished paper).

[3] Christensen, Erik. http://www.timespace.dk /page.php?lang=en&id=3, accessed 2010.1

[4] Couprie, P. "Graphical representation: an analytical and publication tool for electroacoustic music". *Organized Sound*, Vol. 9, No. 1, 2004, pp. 109-113.

[5] Helmuth, Mara, "Barry Traux's Riverrun," in M. Simoni (ed.), *Analytical Methods for Electroacoustic Music*. New York: Routledge, Taylor, and Francis Group, 2006, pp. 187-238.

[6] Hirst, David, "The Use of MQ Plots in the Analysis of Electro-acoustic Music" (Unpublished paper).

[7] Krumhansl, C. L. (1989). "Why is musical timbre so hard to understand?," in S. Nielzen & O. Olsson (eds.), *Structure and Perception of Electroacoustic Sound and Music*, New York: Excerpta Medica, 1989.

[8] Park, T. H., Z. Li, W. Wu, "EASY DOES IT: Electro-Acoustic Music Analysis Toolbox", in *Proceedings of the 2009 ISMIR*, Kobe, Japan.

[9] Park, T. H. "Towards Automatic Musical Instrument Timbre Recognition," Ph.D. Thesis, Princeton University, 2004.

[10] Schaeffer, Pierre. "Traité des objets musicaux," Paris: Éditions, 1966.

[11] Scheirer, Eric, "Tempo and beat analysis of acoustic musical signals," *Journal of the Acoustical Society of America*, No. 50, 1998, pp. 588–601.

[12] Sethares, W. *Tuning, Timbre, Spectrum, Scale*. London: Springer, 2005.

[13] Simoni, M. "Introduction," in M. Simoni (Ed.). *Analytical Methods for Electroacoustic Music*. New York: Routledge, Taylor, and Francis Group, 2006.

[14] Smalley, Denis. "Spectro-morphology and Structuring Processes," in S. Emmerson (ed.), *The Language of Electroacoustic Music*, London: Macmillian, 1987, pp. 61-93.

[15] Smalley, Denis. "Spectro-morphology: Explaining Sound-shapes". *Organized Sound*, Vol. 2, No. 2, 1997, pp. 107-126.

[16] Tao Li, Mitsunori Ogihara, "Content-based music similarity search and emotion detection", ICASSP, 2004.

[17] Tzanetakis, G., Essel, G. Cook, P., "Musical genre classification of audio signals," in *Proceedings of ISMIR 2001*, Indiana, USA.

[18] Wehinger, Rainer. http://www.youtube.com/ watch?v=71hNl_skTZQ, accessed 2010.1