ELECTROACOUSTIC MUSIC: OVERCOMING ANALYSIS PARALYSIS

Kevin Dahan
Université Paris-Est Marne-La-Vallée
Laboratoire LISAA - EA4120
kevin.dahan@wanadoo.fr

ABSTRACT

This paper tries to bring attention to a specific part of electroacoustic music that is often overlooked both by more technical aspects and seemingly more creative concepts - analysis. The field is indeed complex, in that it requires both technical competencies and musicological proficiency, and that's perhaps why few papers address the complex issue of analysis. We begin by reviewing the different technical concepts that are available to the musicologist. Then classical conceptual tools are examined, and the particular problems of electroacoustic music are discussed. Finally, several research strategies are outlined, taking into account musicological needs as well as technical possibilities.

1. INTRODUCTION

While analysis has always been at the forefront of computer music research (both in the technical and musicological sense), the problems of correlating technical analysis techniques to aesthetically meaningful artifacts has seldom been addressed. Analyzing electroacoustic music is indeed a difficult task, and musicological discussions on this aspect are numerous [19] [6] [8] for example.

Since its beginnings, electroacoustic music was in a reflexive attitude towards itself: Schaeffer immediately tried to conceptualize a system in which sound structures were to be classified in aesthetical terms, while Stockhausen tried to demonstrate an equally organized system in Mikrophonie I. This exemplifies the fact that electroacoustic music to a certain extent needed a justification, be it perceptual or self-referential - in order to be accepted as music, a certain amount of codification was in order. Such systems prove their limits as not many composers and/or analysts are stricto sensu using them, but serve as a cornerstone for historical references and as comparison point.

We will firstly describe some of the technical tools available to the musicologist ranging from the classical ones to newer MIR techniques. An overview of the different approaches to electroacoustic music analysis will then be described in more musical terms, and conceptual challenging problems will be described. Finally ideas stemming from cognition are outlined to conclude the discussion.

2. TECHNICAL TACTIQUES...

While many aspects of electroacoustic and computer music are dependent upon technical tools (be it hardware, software, or any combination thereof), the resulting product - and particularly in computer music - is (digitized) sound, most commonly a “sound file”. This sound file (or the representation of it through loudspeakers) is the basic (and most easily accessible) object for the musicologist. Most commonly, the analyst will access its data through a series of representations which doesn’t define the sound itself, but only a particular view of it. We will briefly examine the basic representation techniques commonly used, and have a look at more modern technical approaches.

2.1. Time-based Representation Techniques

Representations of musical signals in which time is a projective dimensions are called time-based. This representation method is very convenient for the analyst and is therefore quite popular. Time-based representations can be subsequently divided in two distinct categories: amplitude-time representations and frequency-time representations, the latter being the most numerous.

2.1.1. Amplitude-Time

Amplitude-time representations are found literally everywhere. It is the most basic representation, in which relative amplitude is plotted against time. One advantage of this method is that it consumes virtually no resource to compute, as the amplitude information are directly accessible in audio-file formats [1]. This representation model is of a limited interest while analyzing, except in some particular cases, in which dynamics are the most important aspect of the work; it can, however, be of use when describing the general schema of a piece.

2.1.2. Frequency-Time

Frequency-time representations are the most used techniques for analyzing electroacoustic and computer music works. They seemingly permit an easy read, as time is usually projected in ordinate and frequency in abscissa, however the underlying mathematics are quite complex and “surprises” may happen, leading to incorrect interpretations. For more details on the technical method, please
refer to [18] and [17]. The signal is segmented in slices of variable length, the bigger it is, the better the discrimination between frequency elements is, but the temporal resolution gets worse. Possibilities exist to reduce this classic "trade-off" problems, involving shaping the signal by using windows and overlapping the slices.

Wavelets are another class of representations, and give a particular time-frequency graph, in which dimensions are relatively independent given the chosen resolution, since wavelet transform is equivalent to convolving a signal with contracted or dilated versions of a mother wavelet [10]. The main interest of wavelets for the musicologist is that they permit to ignore (albeit relatively) the interdependence problem of temporal and frequency domain. The most frequent use of WTs is that they allow a quite good detection of onsets and therefore can help by finely detecting sound attacks.

2.2. Non-time based Representation Techniques

Opposing time-based representations, other representation techniques are useful when analyzing electroacoustic music. Mathematically, "Non-time" is incorrect, but it serves as a shortcut for the musicologist, as they relate to very short time fragments, which are therefore context-independent over the course of a full-length piece or section.

2.2.1. Spectrum

Obtaining a spectral representation of a sound at a given instant is of prime usefulness for the analyst. The underlying mathematical technique is that of Fourier transform, giving a graphical representation of amplitude (frequently ordinate) versus frequencies (abscissa) of a given sound\(^1\). This type of representation is particularly useful for piece based on sound-synthesis techniques, with "composed" timbres [21].

2.2.2. Other

Other representations and/or techniques that can be of use when analyzing computer music or electroacoustic pieces include SNR (Signal to Noise Ratio, estimating the portion of "meaningful" information against noise in a given signal) and THD (Total Harmonic Distortion, giving a reading of the deviation of frequency components from the theoretical harmonics). LPC (Linear Predictive Coding) is also extensively used in speech and music analysis and synthesis, as it works by estimating formants in a given signal[14], resulting in a spectral envelope. Cepstrum is also used by performing a Fourier Transform on the logarithm of the Fourier Transform of the signal, giving a spectral envelope which track the rate of changes in different spectral bands[2]. This technique was further refined and developed within the framework of MIR, which we will now discuss.

\(^1\)More precisely, using a FT on a given sound gives two spectra: the magnitude spectra (amplitude vs frequency) and the phase spectra (phase vs frequency)

2.3. MIR techniques

Many recent developments in the field of computer music has been made in the sub-field of Music Information Retrieval. However tools and techniques developed in this domain are largely (not exclusively) based on two basic assumptions, which are the main characteristics of western, tonal (maybe written) music: musical signal is measured and pitched (as in symbolically representable). As such, fundamental questions for any electroacoustic music composer and analyst such as timbre, frequency space organization, spatial grouping of elements, are not a priority for MIR research. This is largely due to the fact that modeling of electroacoustic audition is more complex than that of tonal music - which allows the use of a preconceived formal grid of frequency space, pitch spacing, meter and tempo estimation and expectations. However, some techniques developed may be of use in the context of computer music.

2.3.1. MFCC

Mel-Frequency Cepstral Coefficients (MFCC) can provide information on redundancy in a given signal. It however assumes the division of frequency space as a non-linear scale (mel-scale) but this information is coherent with human audition. Embedded in this model are the non-linearity of loudness (decibel scale) as well as spectral masking [7], hence it may be usable in an electroacoustic music analysis setting, where many different timbres are simultaneously played.

2.3.2. Repeat Patterns

Many MIR systems rely on repetition in order to identify and extract information from audio streams [15] [23] [16]. Repetition in electroacoustic music is a particular phenomenon, as it rarely happens in the sense it occurs in classical music (as a da capo, for example, with limited variations) or popular music (with predefined forms such as AABA); but it may be a feature of certain composition systems, at different structural levels (as discussed further). Moreover, electroacoustic music is generally unmetered, and thus without "expressed" beats or alternating strong and weak beats, and thus most techniques are not fit for the task yet.

2.3.3. Sound Segregation

Another aspect of interest in MIR systems is the issue of sound segregation [12]. This problem is even more complex in electroacoustic music, where (a) timbres are unpredictable (i.e. rarely defined as "instrument-like", harmonically related spectra) and (b) several layers of seemingly related fragments may not define an unique sound entity.

This problem is strongly related to more musicological problems which will now be discussed.
3. ...MUSICOLOGICAL STRATEGIES

Given the problems of sources in some electroacoustic and computer music works [24] [3], a most common habit is to analyze solely by relying on the audio source of the work.

3.1. Perceptual Analysis

Perceptual analysis’ paradigm somehow relates to the same desire for emancipation from established, “traditional” codes of music. Instead of relying on complex methods and conceptual tools (such as for example Ruwet’s paradigmatic analysis or Schenkerian analysis), perceptual analysis decides to consider raw perception as the only valid analytical source. Perceptual analysis is therefore quite close to certain trends in electroacoustic and computer music composition, in which "color" (i.e. timbre and sound variations) is more important than structure.

3.2. Semiotic Analysis

To a certain extent, perceptual analysis is forced to use a listening score, such as the one constructed by Wehinger for György Ligeti’s Artikulation. As Ligeti in his foreword points out, such score may be over-interpreted, leading to confusion on the meaning of the musical work, as opposed to the interpretation of it, symbolized by the listening score.

This distance to observe, easy to understand in the framework of the listening score, is less obvious when relying on "technical" representations (i.e. non symbolic); however it is the same process in action. In some way, as we mentioned, representations such as sonograms can also be understood as a symbolic construct - a different view on a given object, which only gives one particular perspective.

3.3. Specificities of Electroacoustic Music Analysis

Two main problems emerge from the aforementioned takes on analysis: the lack of symbolic representation and its corollary, the lack of unified "grammar".

3.3.1. Lack of symbolic representation

Since Schaeffer, a number of proposals were made to allow composer to have at their disposal an unified notation technique for electroacoustic and computer music composition. The first obvious answer was to associate written (i.e. language) descriptors to certain sound qualities, such as "bell-like sound", "flute sound", and so on, however as pointed out by Trevor Wishart [6, chapter 3], psychological interpretation may vary greatly from a listener to another.

Use of technological tools allowing to transform sound into symbol was (is?) an obsession of many composers. The prime example of such a system is Xenakis’ UPIC. This apparatus permit to composer to obtain an aural version of a graphical sketch, an approach that influenced further developments [5] [20]. Tools like Acousmographe permit to annotate a sonogram using various (user-defined) symbols, graphs, and information. Following Schaeffer, Smalley developed spectromorphology, an hybrid approach, associating the structural bias of typo-morphology with the flexibility of symbolic naming; more recently [8] attempted to combine spectromorphology with a technological model.

3.3.2. Lack of Grammar

Composition is expressed through definition of a structure. Whereas classical composers had a reservoir of pre-defined forms, post-Webern composers chose not to rely on past forms, therefore creating unique structures using other techniques. It is the first problem that arises when trying to define an analytical grammar adapted to electroacoustic music.

In many electroacoustic works, general (macro-)structure stems from lower-level structures, themselves originating from even lower structures, and so on. This characteristic of electroacoustic music (which is even more obvious in recent works) is what I called morphologic contingency.

3.3.3. Summary of problems

Given the multiple difficulties mentioned, analyzing electroacoustic music seems like a succession of multiple tough choices among multiple technical possibilities and numerous perceptual biases. The first missing link is the problem of correlating the sound to its description. The second problem is related to another linkage problem, this time between the different levels of organization in an electroacoustic work, from the micro-level (grain-level) to the macro-level (the piece in its entirety).

From a certain point of view, technical and musicological ones are related, in that each asks for a trade-off to be made between accuracy and completeness, form and syntax, much resembling Heisenberg’s uncertainty principle.

4. CONCLUSION: BRIDGING THE GAP(S)

We reviewed a number of technical and musicological approaches to electroacoustic music analysis. Given the particularities of such music, traditional, well-known audio signal analysis techniques are sufficient only to help as a guide to determine potentially meaningful objects, whereas more complex developments in MIR are still too much tied to a particular bias towards western pitch-based music.

Computational Auditory Scene Analysis (CASA) [22] is of particular interest, since as we determined, analyzing electroacoustic music is strongly dependent upon perception. Applying this method as a filter gives some interesting results when working with "traditional" music [15]. The main problem is for the analyst to be able to characterize complex sounds in a way sufficiently precise so as
to be able to give a formalized description of it\(^2\) for the system to discriminate.

Following Kelso’s approach to dynamic cognition[9] and Varela’s concepts of enaction, some research is beginning to tackle the issue of the dynamic nature of music with novel tools, from modern cognition approaches [4] to neuro-cognition techniques [11]. At the moment each approach is limited in its use - tonality and creative mechanisms - but analysis of non-metered, non-pitched music is the logical convergence point. By combining a timbre-oriented MIR system with a formalized grammar of sound structures organization (and characterization), a more automated analytical system should be feasible, greatly simplifying the task of electroacoustic music analysis.

5. REFERENCES


\(^2\)We are currently working on a tool that could use Music-V like code to be used as descriptor for sounds to be isolated.