

Iterative Meditations: The use of Audio Feature Extraction Tools on Acousmatic Music Composition

Aluizio B. O. Neto^{1*}

¹CEGeME Lab – Universidade Federal de Minas Gerais

aluizioneto@ufmg.br

Abstract. *This piece explores some possibilities of using Music Information Retrieval and Signal Processing techniques to extract acoustic features from recorded material and use this data to inform the decision making process that is intrinsic to music composition. By trying to identify or create sound descriptors that correlate to the composer’s subjective sensations of listening it was possible to compare and manipulate samples on the basis of this information, bridging the gap between the imagined acoustic targets and the actions required to achieve it. “Iterative Meditations” was created through an iterative process of listening, analyzing, acting and refining the analysis techniques used, having as end product the musical piece itself as well as gathering a collection of tools for writing music.*

1 Program notes

A collage of different soundscapes from urban and natural environments and musical material, assembled by means of digital sound processing techniques. Iterative Meditations is the result of an experiment with the gathering of audio samples and the use of sound feature extraction tools on the creative process of writing an acousmatic piece. The composer intended to develop a continuous repetitive cycle of listening, analyzing, acting and refining the analysis techniques used, having as end product the musical piece itself as well as a set of tools for music writing.

2 Extended Abstract

2.1 Background

Acousmatic music arose from the novel possibilities brought by recording technologies. Pierre Schaeffer’s train [1] carried a sonic revolution that made 20th century’s concept of “Music” a very hard thing to define. Using sound as raw material and manipulating it directly on a physical level by means of analog and/or digital devices allowed for a new way of interacting with this art form and new modes of listening, separating sound from its source. That doesn’t mean the sources were not identifiable, but that they were physically detached from the performed piece, being reproduced through loudspeakers by means of amplified electrical currents.

Audio feature extraction plays a major role on almost all fields related to Music Information Retrieval. The piece presented here was written as a initial effort in researching the use of audio feature extraction techniques in the creative process of acousmatic music composition.

*Supported by CAPES.

Samples were previously recorded or gathered from different sources, being it mono, conventional stereo or binaural. The material was digitally manipulated with the use of free and open source audio software tools and the whole process was informed by data extracted from the original acoustic material. The use of sound descriptors allowed to select, transform and organize sounds based on feature similarity or an intended acoustic goal. Low level descriptors such as spectral flux, spectral centroid, MFCC’s, estimated f0 and others were combined in order to find parameters correlated to the composer’s perception, and used to guide the decision-making process of creating the piece. In this first experiment, higher level descriptors were intentionally avoided in order to keep the understanding of the process closer to the physical level, since the intention was to find simple transformations that could achieve the targeted outcome imagined by the composer. Visualization software was used to combine the extracted features, waveform and spectrogram into a cohesive graphical way of analyzing the piece while writing it. Figure 1 shows a screenshot of a section of it with two channel spectrogram combined with spectral centroid and amplitude follower curves.

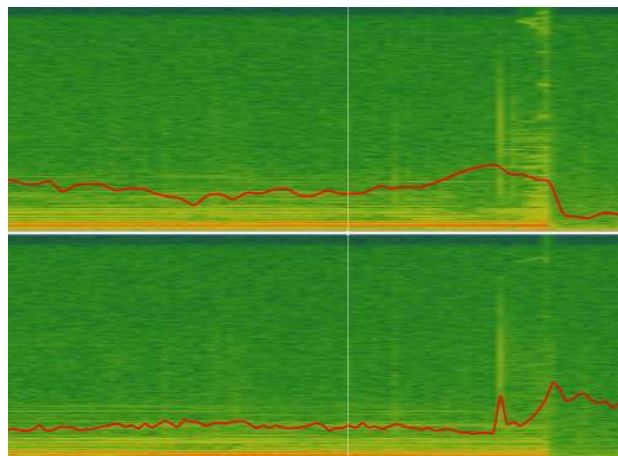


Figure 1: Screenshot of feature visualization

2.2 Methodology

Seems appropriate to describe the methods used in this project by starting from the beginning. The seed for the writing of this piece was a recording of a sharp tone extracted from a glass object hitting a metallic surface. The first sound heard in the final mix was created by combining and manipulating different samples that were chosen by matching their sound features to those of the original

tone. First, the sensation of pitch from the sharp glass sound was correlated to an approximate fundamental frequency of the sound (maximum common divider from the most prominent pure tones present in the original sample). From this f_0 -related data and the spectral distribution of the original sound, some notes from the piano were chosen in order to keep approximately that same sensation of pitch and brightness, while enhancing the timbre of the original sound and bringing it to a more familiar musical sound. A similar process happened when working with the short-term, transient characteristics of that first sound. The attack-time and spectral decay of the original sample were used in order to select and manipulate other sounds that would be combined with it afterwards to achieve the imagined acoustic goal. These newly selected samples as well as their transformed counterparts were then used as raw material for the subsequent part of the piece, and so forth, until its completion, in what could be described as a “fractal-like” or recursive procedure.

Audio analysis and visualization tools were used to achieve the methods described in the previous paragraph. The author mainly used VAMP plugins [2] installed in the Sonic Visualiser package[3]. Secondary tools developed in Matlab® and Python programming languages were also used for data extraction and audio manipulation throughout the writing of the piece. A subset of such secondary tools comprise some methods found as part of the Expan Toolbox [4], developed at the CEGeME Lab at UFMG, and an implementation of the “Deterministic plus Stochastic Model” proposed by Xavier Serra [5].

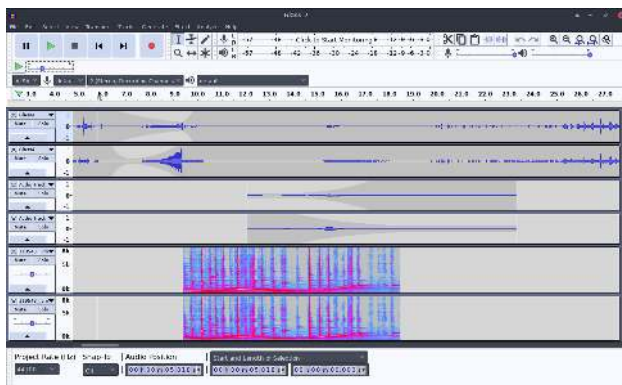


Figure 2: Multi-track visualization and editing

Audacity® [6] was chosen as a host for basic editing and the use of plugins, as well as a platform for simple multi-track waveform and spectrogram visualization (Figure 2). Some tools for spectral blurring, time stretching, frequency shifting and envelope modeling were used through its interface in order to speed up the process and automate some of the work. To achieve complete control of the processes applied, only free and open source plugins were used.

2.3 Conclusion and Future Work

This is a proof of concept piece, in which the author tried novel computational approaches for informing the creative process and combining different methods and tools in the creation of a musical piece. The techniques developed and systems used here are to be described in greater detail and published as a complete article in the near future. The development of more pieces intended to further explore the ideas presented here is also planned for the forthcoming months. Ideally, these tools and methods should be combined into a cohesive package that would bring together all the stages of the process in a unified interface that could ease the dialog between the steps of the creation cycle described in the first paragraph of this text (Program Notes).

References

- [1] Pierre Schaeffer. *Étude aux chemins de fer. Cinq études de bruits, Studio d'Essai, présenté le, 5, 1948.*
- [2] Chirs Cannam. The vamp audio analysis plugin api: A programmer's guide. *Available online: <http://vamp-plugins.org/guide.pdf>, 2009.*
- [3] Chris Cannam, Christian Landone, and Mark Sandler. Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 1467–1468. ACM, 2010.
- [4] Thiago Campolina, Maurício Loureiro, and Davi Mota. Expan: a tool for musical expressiveness analysis. In *Proceedings of the 2nd International Conference of Students of Systematic Musicology*, pages 24–27. IPeM-Institute for Psychoacoustics and Electronic Music Ghent, 2009.
- [5] Xavier Serra et al. Musical sound modeling with sinusoids plus noise. *Musical signal processing*, pages 91–122, 1997.
- [6] D Mazzone and R Dannenberg. Audacity [software]. *The Audacity Team, Pittsburg, PA, USA, 2000.*