Creating Immersive Electronic Music from the Sonic Activity of Environmental Soundscapes

Eli Stine
University of Virginia
Charlottesville, United States
ems5te@virginia.edu

ABSTRACT

The interactions between different living and non-living agents and their surroundings within an environment are complex, multi-dimensional, and self-organizing. Within the context of electroacoustic music composition, this web of organic activity presents a fascinating potential to dictate or inform sonic gestures, textures, and the formal structure of a work. The extraction of these interactions for musical use, however, is easier said than done, and requires a method for translating a representation of an environment to a sound-mappable model. In the case of this work, the representation of the environment used is a stereo field recording, a directional sonic record of the (audible) events within an environment. A deconstruction of this model may then be mapped onto other sounds (or sound generators), ultimately creating a "sonification of a sound", a map from a transcription of the sonic interactions of an environment onto an entirely different sound world. Using methodologies at the intersection of bioacoustics and music information retrieval, the author designed and implemented a software system, *AcousTrans*, that facilitates such a mapping process. Segmented events within a stereo sound recording are intelligently mapped onto multi-channel sound events from another corpus of sounds using a k-nearest neighbors search of a k-dimensional tree constructed from an analysis of acoustic features of the corpus. The result is an interactive sound generator that injects the organicism of environmental soundscape recordings into the sequencing, processing, and composing of immersive electroacoustic music. This work was created within the context of bioacoustic analysis of intertidal oyster reefs, but is applicable to any environmental soundscape that may be effectively decomposed using the described method.

CCS CONCEPTS

• **Applied computing → Sound and music computing**.

KEYWORDS

bioacoustics, algorithmic music, music information retrieval, composing systems

ACM Reference Format:
Eli Stine. 2019. Creating Immersive Electronic Music from the Sonic Activity of Environmental Soundscapes. In *Joint Proceedings of the ACM IUI 2019 Workshops, Los Angeles, USA, March 20, 2019* , 4 pages.

1 INTRODUCTION

The history of applied musical engagements with the relationship between sound and the environment is rich and lengthy. Classical examples include the Ancient Greek "Harmony of the Spheres", a description of the harmonious musical proportions of the planets, a creation story in the Vedic Sanskrit texts that takes sound as fundamental, and various other archaic "acoustemologies" (sonic ways of knowing the world) [2][12]. More modern, applied instrumental music examples include John Cage's *Atlas Eclipticalis*, a work that makes use of star charts to inform the structure of a composition [11], the musical settings of bird calls and other animals in the works of Machê and, famously, Messiaen [13], and the millimetrization technique of Schillinger [17], all of which seek to map systems outside of the concert hall onto instrumental music.

More precise, algorithmic methods include those designed by Iannis Xenakis, or Gordon Mumma in his *Mograph* series, which took as sonic material seismological data [14], along with computer-assisted, data-driven works such as John Luther Adam's *The Place Where You Go To Listen* [1] or Carla Scaletti's *h->gg*, a work which utilizes data from the Large Hadron Collider (LHC). Acousmatic music composers such as Natasha Barrett [4] and Hans Tutshku [15] make use of models of natural systems (notably geological and hydrological systems) to generate sound materials and dictate musical structures used within their multi-channel acousmatic works. Digital musical instruments built using biosensors or motion sensors take advantage of the nuanced, organic data that may be derived from the human body to control sound synthesis and audio file manipulation in real-time (such as in the work of Atau Tanaka [25]).

2 RELATED WORK

There exist many different software tools, in a variety of disciplines, which seek to deconstruct and/or map onto sound organic systems [8]. Some approaches engage with natural computing and artificial intelligence: creating musical prescriptions using Cellular Automata, L-Systems, or flocking simulations (such as in the author's own *Murmurator* system) [24]. The field of auditory display and sonification, when making use of natural system data, also engages methods for best representing in sound the organic interactions of natural data, with many different sonification softwares available (such as [26]).

Decoding and transcribing the events within a recording is under the purview of both Music Information Retrieval (MIR) and bioacoustics. Pertinent MIR tasks include automatic transcription, track separation, and speaker diarization, each of which seek to automatically reveal structural decompositions of acousmatic sound. Within acoustic ecology and bioacoustics, techniques have been

---

IUI Workshops '19, March 20,2019, Los Angeles, USA.

developed to assist in the decomposition of environmental soundscapes, revealing their underlying sonic components (such as in the work of sound recordist Bernie Krause and composer Hildegard Westerkamp).

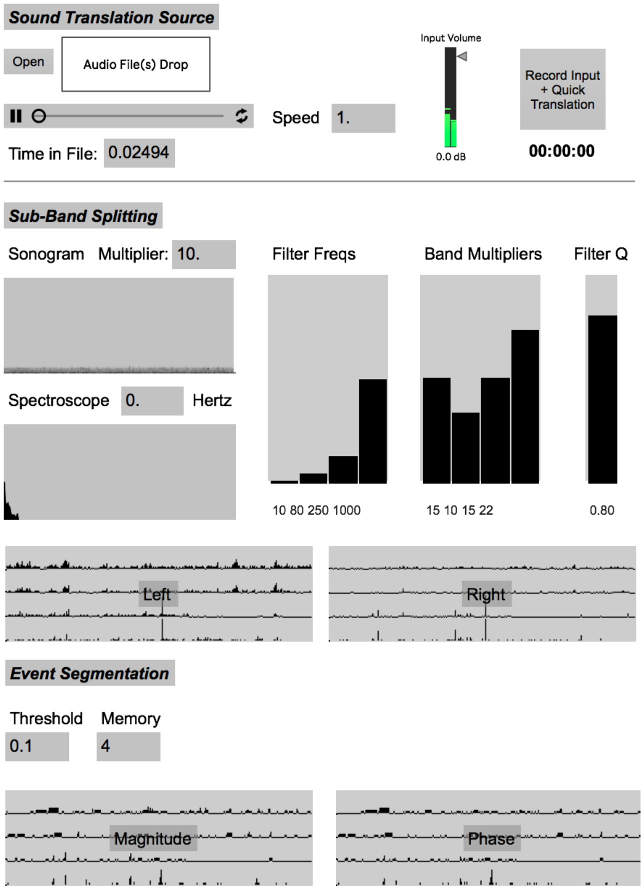
The specific method that this work builds from is concatenative sound synthesis, a synthesis technique that may be generally described as granular synthesis driven by audio analysis, and more specifically a process of selecting grains of sound from a file or corpus based on their best fit to some acoustic criteria [21]. There are a wide variety of projects developed over the past few decades that make use of concatenative synthesis, ranging from more or less artistic and scientific applications and from off-line systems to, more recently, real-time implementations [20]. These include cataRT [22], timbreID [5], developments into ”Soundspotting” [6] and ”Audio Mosaicing” [9] techniques, and the work of scientist-musicans Jean-Julien Aucouturier [3] and Aaron Einbond [10], among others. The author's system builds around the MuBu concatenative synthesis engine [18]: connecting it to an environmental soundscape event parser and an expressive electroacoustic sound mapper.

3 DESIGN

The goal of *AcousTrans* (*Acousmatic Translator*) is to allow a user to load in a source stereo audio file (field recording or other environmental recording) and a destination corpus of other audio files and interactively map the events, gestures, and structure of the source onto the destination. What results is a stereo or multi-channel audio file with gestural, rhythmic, and/or structural similarities to the source file, but with entirely different timbral characteristics: those of the destination corpus.

3.1 Filtering and Segmentation

AcousTrans operates by first taking in a user-selected stereo audio file of a soundscape (an intertidal oyster reef stereo hydrophone recording, for example) within the segmenter module (Figure 1). This audio file is then sent through N low-pass/band-pass/high-pass filters whose frequencies are tuned to the particularities of the soundscape (the specific threshold between the sub-soundscapes of wave movement, oysters, snapping shrimp, etc., within the reef, for example). Within the context of the intertidal oyster reef recordings a value of $N = 4$ was deemed both necessary and sufficient, although the system allows for N to be variable. Each of the N spectral sub-bands of the source audio file is then segmented using one of several different segmentation modes: amplitude-based peak detection, spectral flux-derived segmentation, or fixed-size segmentation. After segmentation, the result is an N-channel stream of events which encode the independent activity within each sub-band of the source file. The events are encoded as lists of intensity (average volume), duration, stereo localization (from centered to at either left or right channels), and then a subvector of acoustic features (including fundamental frequency (F0) estimation, energy, periodicity, loudness, spectral centroid, spectral spread, spectral skewness, spectral kurtosis). This multi-channel stream of events is then passed into the playback module (Figure 2).



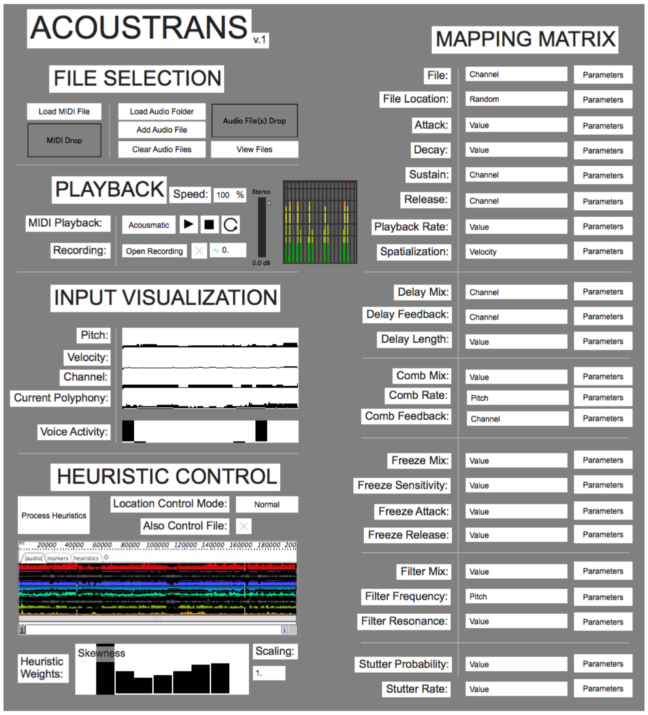
Speed 1.
Time in File: 0.02494
00:00:00
Sonogram Multiplier: 10.
Spectroscope 0. Hertz
Filter Freqs 10 80 250 1000
Band Multipliers 15 10 15 22
Filter Q 0.80
Threshold 0.1
Memory 4

Figure 1: AcousTrans segmenter module, the output of which is fed into the playback module (below).

3.2 Event Mapping

Different dimensions of these events are mapped to different parameters of sounds generated by *AcousTrans* using a mapping matrix (right side of Figure 2). For example, the intensity of a source event may be used to dictate the volume of a destination event, or the stereo localization value of a source event may be used to dictate the spatialization speed of the destination event. These destination events take as sound material a user-selected corpus of audio files. Further electroacoustic abstractions may also be applied including delay, comb filtering, spectral freezing, filtering, and a probabilistic repetition (stutter) effect, the parameters of each being either set by the user or driven by different event dimensions.

The acoustic features embedded in each source event may be used to select a similar sound within the user-selected destination corpus via concatenative synthesis. Using a k-nearest neighbors search algorithm on a k-dimensional tree constructed from the acoustic features of segments of each audio file in the audio file corpus, the subvector of acoustic features for a source event is mapped to the most similar sound within the destination corpus [23].



Speed: 100 %
MIDI Playback: Acousmatic
Recording: Open Recording 0.
Location Control Mode: Normal
Also Control File:
Scaling: 1.

File: Channel
File Location: Random
Attack: Value
Decay: Value
Sustain: Channel
Release: Channel
Playback Rate: Value
Spatialization: Velocity
Delay Mix: Channel
Delay Feedback: Channel
Delay Length: Value
Comb Mix: Value
Comb Rate: Pitch
Comb Feedback: Channel
Freeze Mix: Value
Freeze Sensitivity: Value
Freeze Attack: Value
Freeze Release: Value
Filter Mix: Value
Filter Frequency: Pitch
Filter Resonance: Value
Stutter Probability: Value
Stutter Rate: Value

Figure 2: AcousTrans playback module, including file selection, playback, input visualization, heuristic control, and mapping matrix submodules.

At the time of this writing, there is no similarity thresholding procedure (i.e. producing silence when a source event does not have a highly similar destination event), so a match, however distant, will always be made. The user may customize the weighting of the acoustic features used in the search via a multislider interface (bottom left of Figure 2), which can be useful to "tune" the system depending on the particular source and destination sounds (for example, de-emphasizing F0 estimation if only using sounds with no clear pitch center). The result of this process is an acoustic feature-driven mapping between the events in the source audio file and those generated by the system from the user-selected audio file corpus. Combined with the electroacoustic abstractions outlined above, this system can generate a diverse array of natural system-derived soundscapes. An overview of the functionality of *AcousTrans* as a system diagram may be viewed in Figure 3.

3.3 Implementation

AcousTrans is implemented in Cycling '74's Max 8 [7], taking advantage of ICST's Ambisonics externals to handle multi-channel audio [16] and IRCAM's MuBu for Max [18] and Programming Interface for Processing Objects (PIPO) Max externals to handle acoustic feature analysis [19].

4 APPLICATIONS

The author has composed several multi-channel electroacoustic compositions using material generated by *AcousTrans*. These include *Reef*, for octophonic fixed media, presented at the 2018 Coastal Futures Conservatory conference at University of Virginia, and *No Where*, for octophonic fixed media, presented at the 2018 Technosonics conference at University of Virginia, the 2019 National Student Electronic Music Event (N_SEME) at University of Virginia, and at the 2019 Society for Electroacoustic Music in the United States (SEAMUS) conference at Berklee College of Music. Examples of the system in action may be heard at www.elistine.com/software/acoustrans.

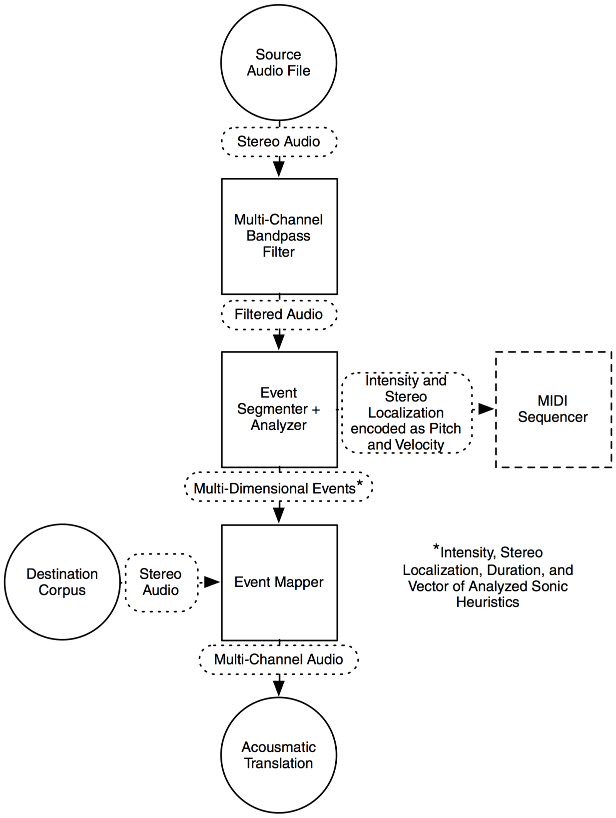


Figure 3: Systemic diagram of the inputs, processing, and intermediary data formats in AcousTrans.

5 FUTURE WORK AND CONCLUSION

Future work includes expanding the segmentation algorithm to be more sensitive to dense, low dynamic range, "lo-fi", environmental soundscapes, expanding the type and control parameters of the acoustic features control submodule, and distributing the software so that it may be used by other soundscape composers and electroacoustic musicians.

Ultimately, *AcousTrans* presents a methodology for intelligently mapping a multi-dimensional stream of gestures from one environmental soundscape to an entirely different, multi-channel electroacoustic sound world. It harnesses techniques from both bioacoustics and MIR to facilitate the generation of electroacoustic material derived from the activity of natural systems.

ACKNOWLEDGMENTS

Thanks to Environmental Sciences Ph.D. candidate Martin Volaric and the University of Virginia Environmental Sciences Department's Moore Award and the Jefferson Scholars Foundation for providing financial support for this research.

REFERENCES

[1] John Luther Adams. 2010. *The place where you go to listen: In search of an ecology of music.* Wesleyan University Press.

[2] Aaron S Allen and Kevin Dawe. 2015. Current Directions in Ecomusicology: Music, Culture, Nature. Routledge, Chapter 1.

[3] Jean-Julien Aucouturier and François Pachet. 2006. Jamming with plunderphonics: Interactive concatenative synthesis of music. *Journal of New Music Research* 35, 1 (2006), 35–50.

[4] Natasha Barrett. 2016. Interactive spatial sonification of multidimensional data for composition and auditory display. *Computer Music Journal* 40, 2 (2016), 47–69.

[5] William Brent. 2010. A timbre analysis and classification toolkit for pure data. In *International Computer Music Conference (ICMC).* 224–229.

[6] Michael Casey. 2012. Soundspotting: A new kind of process? In *The Oxford Handbook of Computer Music.*

[7] Cycling '74. 2019. Max. https://cycling74.com/products/max

[8] Leandro Nunes De Castro. 2006. Fundamentals of natural computing: basic concepts, algorithms, and applications. Chapman and Hall/CRC, Chapter 1.

[9] Jonathan Driedger, Thomas Prätzlich, and Meinard Müller. 2015. Let it Bee-Towards NMF-Inspired Audio Mosaicing.. In *International Symposium on Music Information Retrieval (ISMIR).* 350–356.

[10] Aaron Einbond, Christopher Trapani, Andrea Agostini, Daniele Ghisi, and Diemo Schwarz. 2014. Fine-tuned control of concatenative synthesis with catart using the bach library for Max. In *International Computer Music Conference (ICMC).* 1037–1042.

[11] Sabine M Feisst. 2009. John Cage and improvisation: an unresolved relationship. *Musical improvisation: Art, education, and society* 2, 5 (2009).

[12] Steven Feld. 1993. The Politics of Amplification: Notes on 'Endangered Music'and Musical Equity. *Folklife Center News XV/1 (1993): 12* 15 (1993).

[13] François-Bernard Mâche. 1992. *Music, myth, and nature, or, The Dolphins of Arion.* Taylor & Francis.

[14] Gordon Mumma. 2011. Gordon Mumma: MEDIUM SIZE MOGRAPH 1962. *Leonardo Music Journal* (2011), 77–78.

[15] Ketty Nez. 2003. An Interview with Hans Tutschku. *Computer Music Journal* 27, 4 (2003), 14–26.

[16] Jan C Schacher. 2010. Seven years of ICST Ambisonics tools for maxmsp–a brief report. In *Proc. of the 2nd International Symposium on Ambisonics and Spherical Acoustics.*

[17] J. Schillinger, L. Dowling, and A. Shaw. [n. d.]. *The Schillinger system of musical composition.* Number v. 1 in The Schillinger System of Musical Composition. C. Fischer, inc.

[18] Norbert Schnell, Axel Röbel, Diemo Schwarz, Geoffroy Peeters, Riccardo Borghesi, et al. 2009. MuBu and friends–assembling tools for content based real-time interactive audio processing in Max/MSP. In *International Computer Music Conference (ICMC).* 423–426.

[19] Norbert Schnell, Diemo Schwarz, Joseph Larralde, and Riccardo Borghesi. 2017. PiPo, A Plugin Interface for Afferent Data Stream Processing Modules. In *International Symposium on Music Information Retrieval (ISMIR).*

[20] Diemo Schwarz. 2006. Concatenative sound synthesis: The early years. *Journal of New Music Research* 35, 1 (2006), 3–22.

[21] Diemo Schwarz. 2007. Corpus-based concatenative synthesis. *IEEE signal processing magazine* 24, 2 (2007), 92–104.

[22] Diemo Schwarz, Grégory Beller, Bruno Verbrugghe, and Sam Britton. 2006. Real-time corpus-based concatenative synthesis with catart. In *9th International Conference on Digital Audio Effects (DAFx).* 279–282.

[23] Diemo Schwarz, Norbert Schnell, and Sebastien Gulluni. 2009. Scalability in content-based navigation of sound databases. In *International Computer Music Conference (ICMC).* 13–16.

[24] Eli Stine and Kevin Davis. 2018. The Murmurator: A Flocking Simulation-Driven MultiChannel Software Instrument for Collaborative Improvisation. In *International Computer Music Conference (ICMC).* forthcoming.

[25] Atau Tanaka. 2010. Mapping out instruments, affordances, and mobiles. In *Proceedings of the International Conference on New Interfaces for Musical Expression.* 88–93.

[26] Catherine M Wilson and Suresh K Lodha. 1996. Listen: A data sonification toolkit. Georgia Institute of Technology.