Creating Musical Structure from Environmental Soundscapes

Ryan Maguire
Bregman Music and Audio Research Studio
Dartmouth College

Introduction

Recent research in Music Informatics has explored Probabilistic Latent Component Analysis (PLCA) and related techniques to isolate independent components from mixed musical signals. This has applications in automatic thumbnailing, retrieval, structural segmentation, as well as music composition. Here, PLCA has been used to extract sources from a mixed signal of environmental sounds. Natural soundscapes share traits with music such as repetition, development, persistence, and contrast. We can exploit these qualities, mapping each independent source to various musical parameters. Thus, the sequence of independent auditory events in an environmental soundscape is used to create musical structure.

Source Separation

Techniques for source separation from mixed audio include convolutive non-negative matrix factorization (NMF), independent subspace analysis (ISA), Probabilistic Latent Component Analysis (PLCA), and Shift Invariant Probabilistic Latent Component Analysis (SI-PLCA) [1]. These techniques are well documented in the Music Informatics literature.

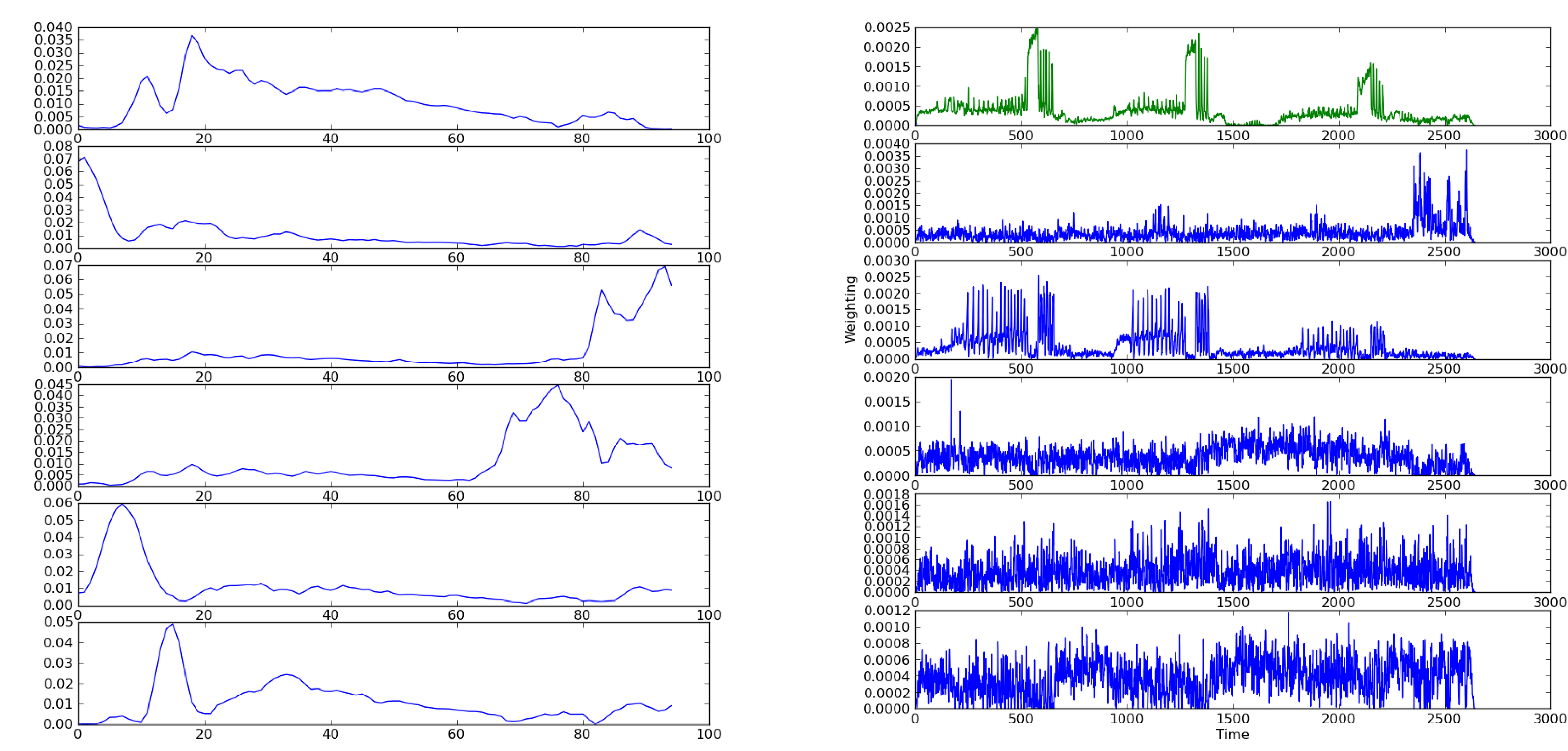
PLCA is a probabilistic variant of NMF. It decomposes a non-Negative matrix V into the product of two multinomial probability Distributions, W and H, and a mixing weight, Z. In the auditory Domain, V would be a matrix representing the time-frequency Content of an audio signal:

$$V \approx WZH = \sum_{k=0}^{K-1} w_k z_k h_k^T$$

where each column of W can be thought of as a recurrent frequency template and each row of H as the excitations in time of the corresponding basis. Z = diag(z) is a diagonal matrix of mixing weights z and K is the number of bases in W. Each of V, w_k, z_k, and h_k are normalized to sum to 1, as they correspond to probability distributions.

In this work, we use PLCA to parse out independent components from a field recording of a soundscape containing multiple sources.

Frequency Bins & Time Kernels



Bases w_k using PLCA with k = 6. Amplitudes across 95 frequency bins.

h_k using PLCA with k = 6.

From Sources to Music

We use PLCA to capture the amplitudes in time of independent sound sources in an environmental soundscape. We then use this information as control data for musical parameters, thus generating a musical work with its temporal structure derived from the source soundscape.

As a preliminary exercise, a recording of a mountain soundscape was analysed. There are a number of distinct bird calls and cicadas audible in the source signal. Implementing PLCA with k = 6 resulted in six basis functions w, each with a distinct time profile h (see *above*).

Given these unique amplitude-time kernels for each basis, w, we have independently mapped each h_k to a different synthesis parameter. For example, in our preliminary implementation, we used each frequency kernel to control the amplitude of an independent synthesizer. In this way, the temporal structure of the original components from the soundscape are used to create the temporal structure of a synthesized musical environment.

In a second implementation, we have attempted to also use the frequency information present in the sources to determine pitch content in our musical structure. Here we use SI-PLCA and utilize pitch tracking software to convert frequency content from our independent sources to note data. This relates to previous compositional uses of source separation [2].

Conclusions

Previous uses of PLCA in music composition have focused only on the small frequency patches generated. Here we have used PLCA to extract the temporal structure from a natural soundscape and then used this structure to determine the form of a musical work. This bears some similarity to the idea of cross synthesis with PLCA, wherein the basis functions extracted from one audio signal are combined with the time kernel of a different signal. Here, we are using the time kernel from our extracted sources to determine the structure of a musical work by controlling independent musical parameters with these separate kernels.

Future Directions...

Future possibilities include using the extracted temporal information to control an increasing variety of parameters or, alternatively, using this information to generate an acoustic composition or as an input for an algorithmically generated piece. One might also work with longer sound files, so as to utilize processes which occur more gradually in nature. An additional possibility would be to further our second implementation, integrating audio from the soundscape with the structure generated from the same source. To this end, one can envision the development of an app where a user could record a sound file from their immediate environment and then the application could use these techniques to turn their recording into a piece of music. Thus, it would be essentially a "musicification" of any given environment.

Selected References

[1] Smaragdis, P., and Bhiksha Raj. "Shift-Invariant Probabilistic Latent Component Analysis", *MERL Technical Report*, TR2007-09.

[2] Topel, Spencer S., and Michael A. Casey. "Elementary Sources: Latent Component Analysis for Music Composition", *12th International Society for Music Information Retrieval* Conference, Miami, USA, 2011.