CREATING MUSICAL STRUCTURE FROM THE TEMPORAL DYNAMICS OF SOUNDSCAPES

Ryan Maguire

Dartmouth College
Digital Musics
6242 Hallgarten Hall
Hanover, NH 03755

ABSTRACT

Acoustic signal mixtures from the environment often manifest interesting structural properties. These features can be used to generate musical forms, but access is a problem. Here, Probabilistic Latent Component Analysis (PLCA) has been used to automatically extract independent sources from a mixed signal of environmental sounds. Excitation information corresponding to the temporal evolution of independent acoustic sources in the original environment is extracted. Given this information, we map the temporal dynamics of each independent sound source to a variety of parameters in a musical composition. Thus, the temporal flow of auditory events in an environmental soundscape is used to determine musical structure. After a discussion of PLCA, related composers and music compositions are highlighted. Thereafter, an approach to mapping the temporal dynamics of a given soundscape to a musical setting using source separation techniques is illustrated.

1. INTRODUCTION

Recent Research in Music Informatics has explored the use of Probabilistic Latent Component Analysis (PLCA) and related techniques to isolate independent components from mixed musical signals [1][6][7]. This has applications in automatic thumbnailing, retrieval, and structural segmentation, as well as music composition [8][11]. Recent compositions have used PLCA to cull specific sonic components from source recordings and then utilized these samples as musical source material [9].

In this paper, we propose using PLCA as a tool to generate musical structure rather than surface material for a composition. Natural soundscapes share traits with musical sound such as repetition, development, persistence, and contrast. Using PLCA to analyse field recordings, one can access the ebb and flow of auditory events in a natural soundscape. The individual sources can then be used to determine the structure of a composition, mapping extracted features to independent synthesizers, instruments, or parameters. In this way, the structure of a natural soundscape can be directly utilized in the creation of musical form.

2. 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][4][6][7][8][11]. These techniques are well documented in the Music Informatics literature.

Probabilistic Latent Component Analysis 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:

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

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 [11]. Each of V, w_k , z_k, and h_k are normalized to sum to 1, as they correspond to probability distributions.

Sparse settings are commonly favoured in the literature. These solutions can be encouraged by imposing constraints using prior distributions, alphaZ and alphaH. A detailed explanation is beyond the scope of this paper, but can be found in [11].

3. MUSICAL POINTS OF REFERENCE

Several compositions have drawn on PLCA and related techniques for their material. Topel and Casey discuss five compositions which utilize extracted components to articulate characteristics of their source audio at specific moments [9]. For example, *Violine* (2011), by Spencer Topel decomposed 12 – 90 second audio clips of J.S. Bach's Chaconne in d minor from Partita No. 2 for solo violin into individual notes or pitch classes and then matched extracted components to a live violin signal. Another interesting example is *Strange-Charmed* (2011), by Michael Casey and Simon Atkinson, which used independent subspace analysis (ISA) to generate an expanded set of sound materials from source sounds such as Geiger counters, insects, and band-pass filters [2][3][4].

978-1-4673-0382-8/12/$31.00 ©2012 IEEE

4. STRUCTURE FROM ENVIRONMENTAL SOUND

In this work, we have used PLCA to capture the dynamic flow of independent sound sources in a soundscape and then used this material to control numerous synthesizers, thus generating a musical work with its temporal structure derived from the source soundscape. This differs from previous uses of PLCA and SIPLCA which have focused on using the extracted features directly as audible material for composition. We have instead abstracted the extracted features and used them as a higher level parametric control. Rather than listening to the features themselves, one experiences their presence indirectly.

As a preliminary exercise in this technique, a recording of a mountain soundscape was analysed [5]. The source audio was 2 minutes and 2 seconds long. There are a variety of different bird calls which enter and fade, crescendo and decrescendo independently at different distances, times, and rates. Cicadas also fill the signal with their rhythmic call, more or less noticeable at different times.

In our implementation, standard PLCA with k = 6 was used. This number was an estimate of the number of independent sound sources audible in the recording. This resulted in 6 basis functions w (see Figure 1), each with a distinct correspondent time profile h (see Figure 2). Each w has a distinct frequency profile. This can be considered to be the spectral "signature" of each individual source. Similarly, each corresponding time function h represents when the given spectral profile is most present in the source recording.

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

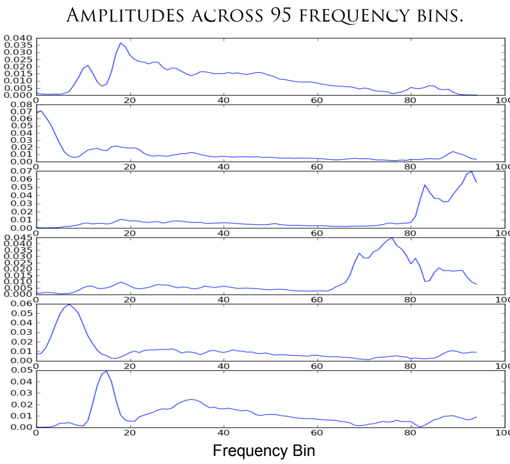

Figure 2. Bases w_k using PLCA with k = 6.

5. CONCLUSION

Previous uses of PLCA in music composition have used the extracted bases as sound material. 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.

Future possibilities include using the extracted temporal information to control other 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 and changes which occur more gradually in nature. Finally, one might combine this approach with the previous techniques by integrating audio from the soundscape with the structure generated from the same source.

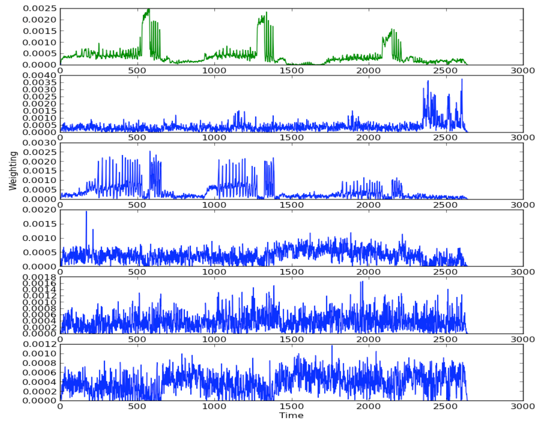

Figure 3. h_k using PLCA with k = 6.

REFERENCES

[1] Cardoso, Jean-François. "Source separation using higher order moments", *Proc. ICASSP-89, pages 2109-2112*, 1989.

[2] Casey, M.A. "Soundspotting: A new kind of process", *The Oxford Handbook of Computer Music*, 2009.

[3] Casey, M.A., and S. Atkinson. *Strange-Charmed: MIT EMS@25 CD, Track 9*. MIT Media Laboratory, 1999.

[4] Casey, M.A., and A. Westner. "Separation of mixed audio sources by independent subspace analysis", *Proceedings of the International Computer Music Conference*, pages 154-161, 2000.

[5] Freesound, "050823-suraksan_extract.wav", Accessed September 29th, 2011, http://www.freesound.org/people/sazman/sounds/7959/

[6] Smaragdis, P., Bhiksha Raj, and Madhusudana V.S. Shashanka. "Sparse and shift-invariant feature extraction from non-negative data", *ICASSP*, pages 2069-2072. IEEE, 2008.

[7] Smaragdis, P., and Bhiksha Raj. "Shift-Invariant Probabilistic Latent Component Analysis", *MERL Technical Report*, TR2007-009.

[8] Smaragdis, P., B. Raj and M.V. Shashanka "Supervised and semi-supervised separation of sounds from single-channel mixtures", *Proc. 7th Int. Conf. Ind. Compon. Anal. Signal Separation (ICA '07)*, p.414, 2007.

[9] 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.

[10] Tawaststjerna, Erik; Robert Layton (Translator)(1976-1986). *Sibelius*. Faber & Faber, Vol. I, 1865-1905, Vol. II, 1904-1914, London.

[11] Weiss, Ron J., and Juan Pablo Bello "Identifying repeated patterns in music using sparse convolutive non-negative matrix factorization", *Proc. Int. Conf. Music Inf. Retrieval*, pp.420-425, 2005.