The Application of Psychoacoustic Audio Analysis Techniques to Electroacoustic Music for the Purpose of Visualization

Anderson Mills

The University of Texas at Austin

147th meeting of the Acoustical Society of America

New York, NY

2004–05–25

Introduction

- Objective
- Motivation
  - Electroacoustic Music
  - Use of Human Hearing Models
  - Lack of Source Segregation
- Human Hearing Model Resources
- Progress
  - Brainstorming Results
  - Graphical Score Prototype

Objective

The Goal

create algorithms which use models of human hearing to extract audio properties from recorded electroacoustic music

- identification of appropriate audio properties
- design of audio property algorithms
- validation of those algorithms against human performance

Motivation

- Electroacoustic Music
- Use of Human Hearing Models
- Lack of Source Segregation

Motivation

Electroacoustic Music — Unusual Type of Music

- definition — involves electronic technology for the compositional manipulation of sound
- not restricted by physics
- the blurry lines of perception
- no Western musical score

Motivation

Why use models of human hearing?

- Physical representations lose aspects of what humans hear.
  - missing fundamental
  - streaming
  - perceived loudness

Motivation

Why not segregate sources?

- Source Segregation
  - Holy Grail — translate a continuous pressure variation into Western musical notation
  - first separate the energy from different auditory sources
- Sans Source Segregation
  - the nature of electroacoustic music
  - example — Dr. Scheirer’s tempo algorithm

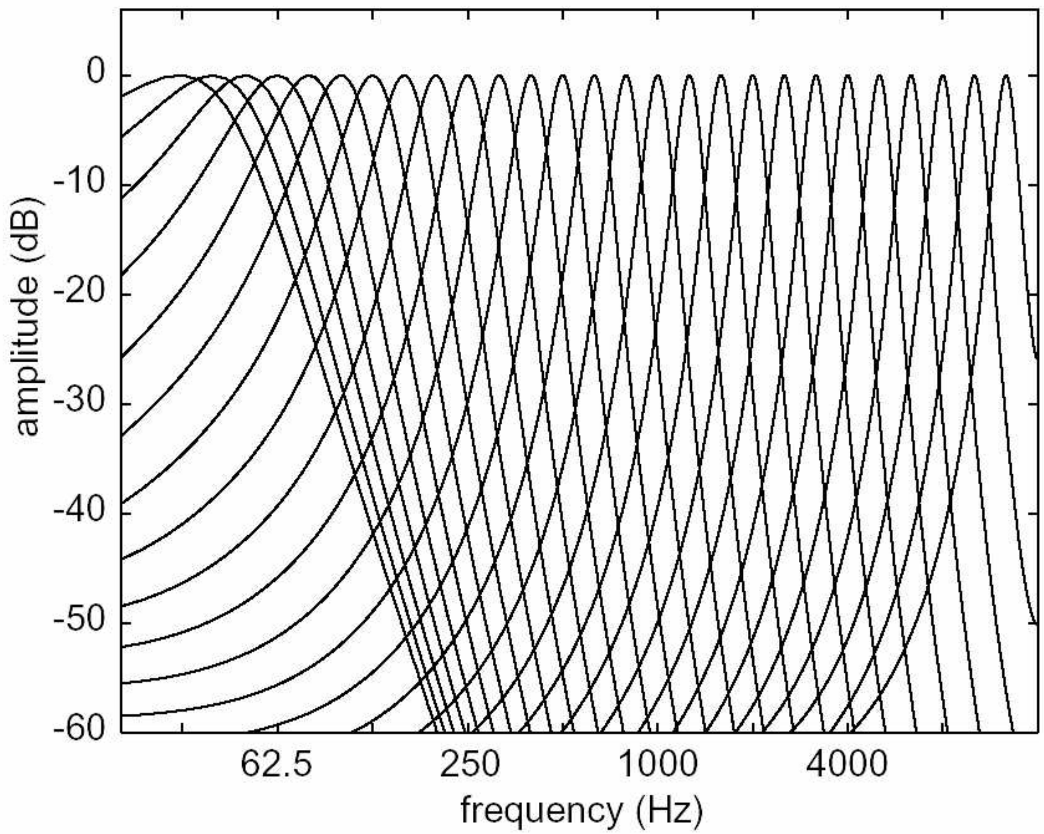
Resources

Human Hearing Model Algorithms

- Dr. Patterson’s Auditory Image Model (AIM)
- Drs. Meddis and Hewitt’s Correlogram
- Dr. Ellis’ Weft

Resources

Dr. Patterson's Gammatone Filterbank

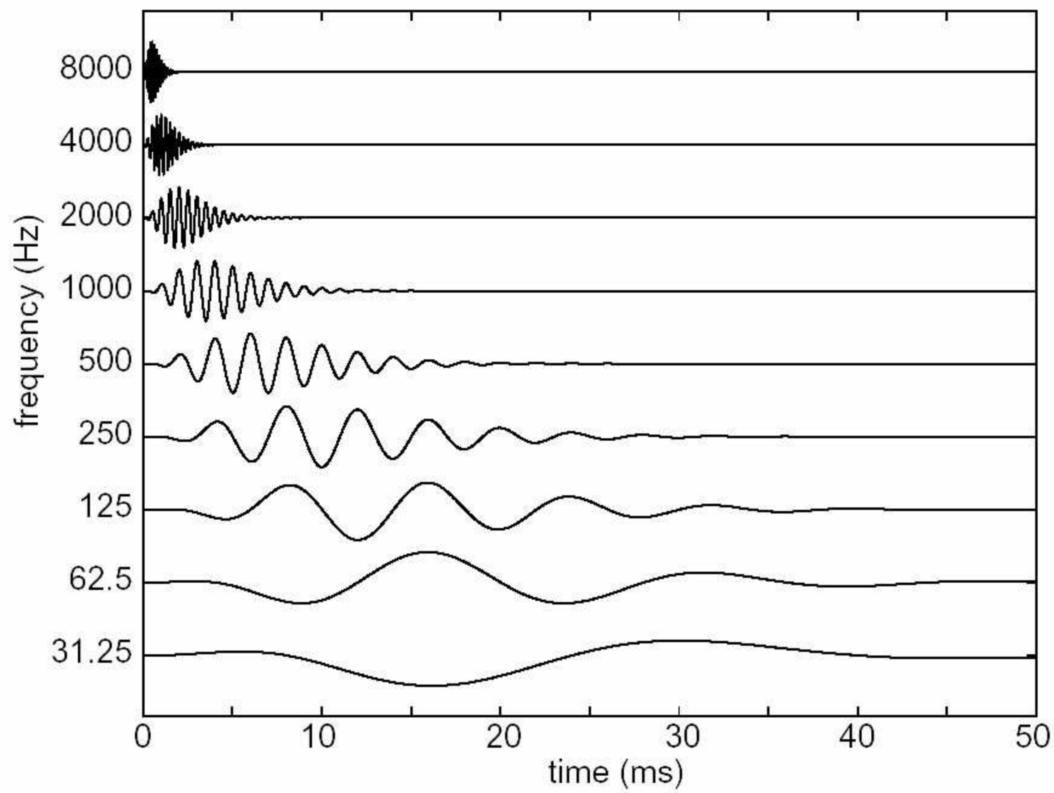


Overall frequency response of the cochlear filterbank, plotted on a logarithmic frequency scale (every second filter is shown).

K.D. Martin, *Sound-Source Recognition: A Theory and Computational Model*, PhD thesis, Massachusetts Institute of Technology, June 2000.

Resources

Basilar Membrane Motion

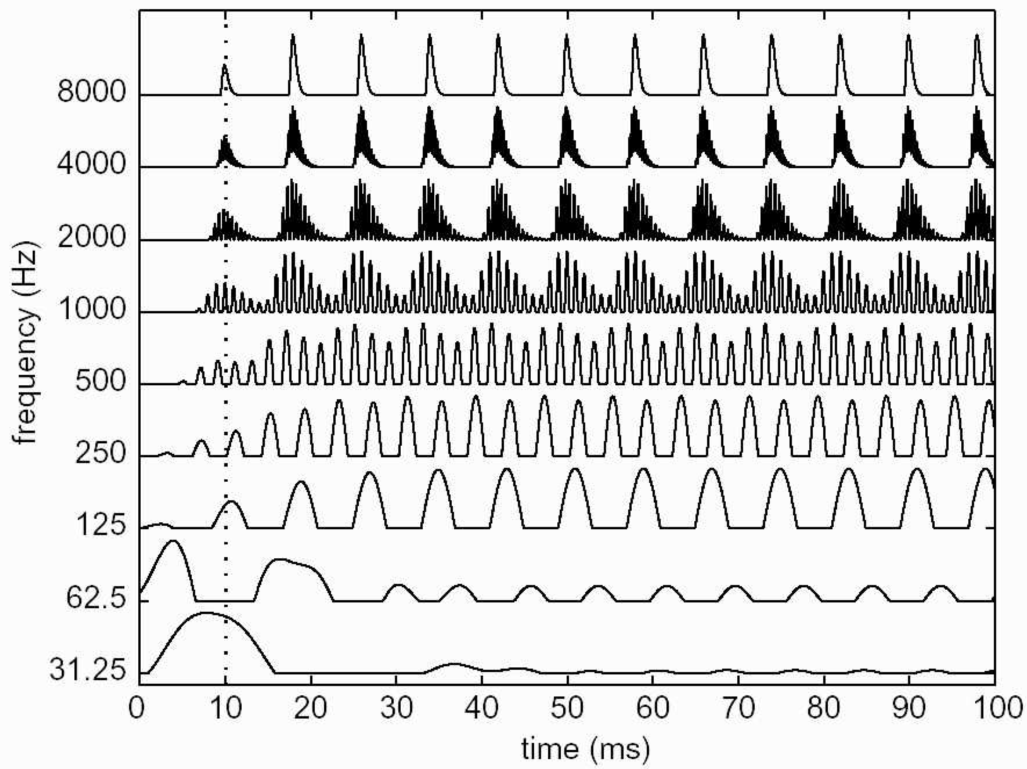


Impulse responses of nine cochlea bandpass filters (one filter is shown per octave). Their amplitudes have been normalized to a uniform scale for display purposes.

K. D. Martin, *Sound-Source Recognition: A Theory and Computational Model*, PhD thesis, Massachusetts Institute of Technology, June 2000.

Resources

Drs. Meddis and Hewitt's Cochleogram

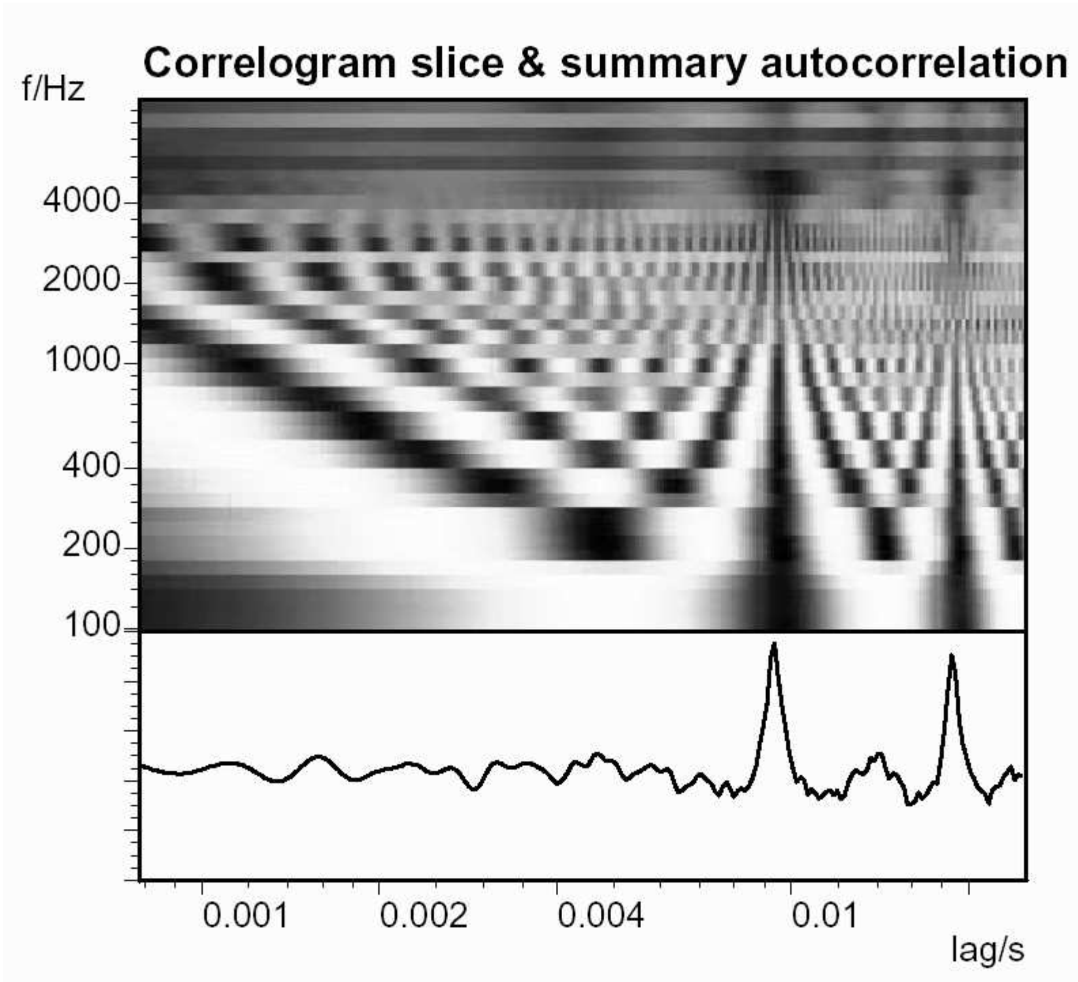


Responses of nine cochlea bandpass filters (one filter is shown per octave) to the 125 Hz sawtooth signal after half-wave rectification and light smoothing intended to model inner hair cell transduction. The output amplitudes have been normalized to a uniform scale for display purposes.

K. D. Martin, *Sound-Source Recognition: A Theory and Computational Model*, PhD thesis, Massachusetts Institute of Technology, June 2000.

Resources

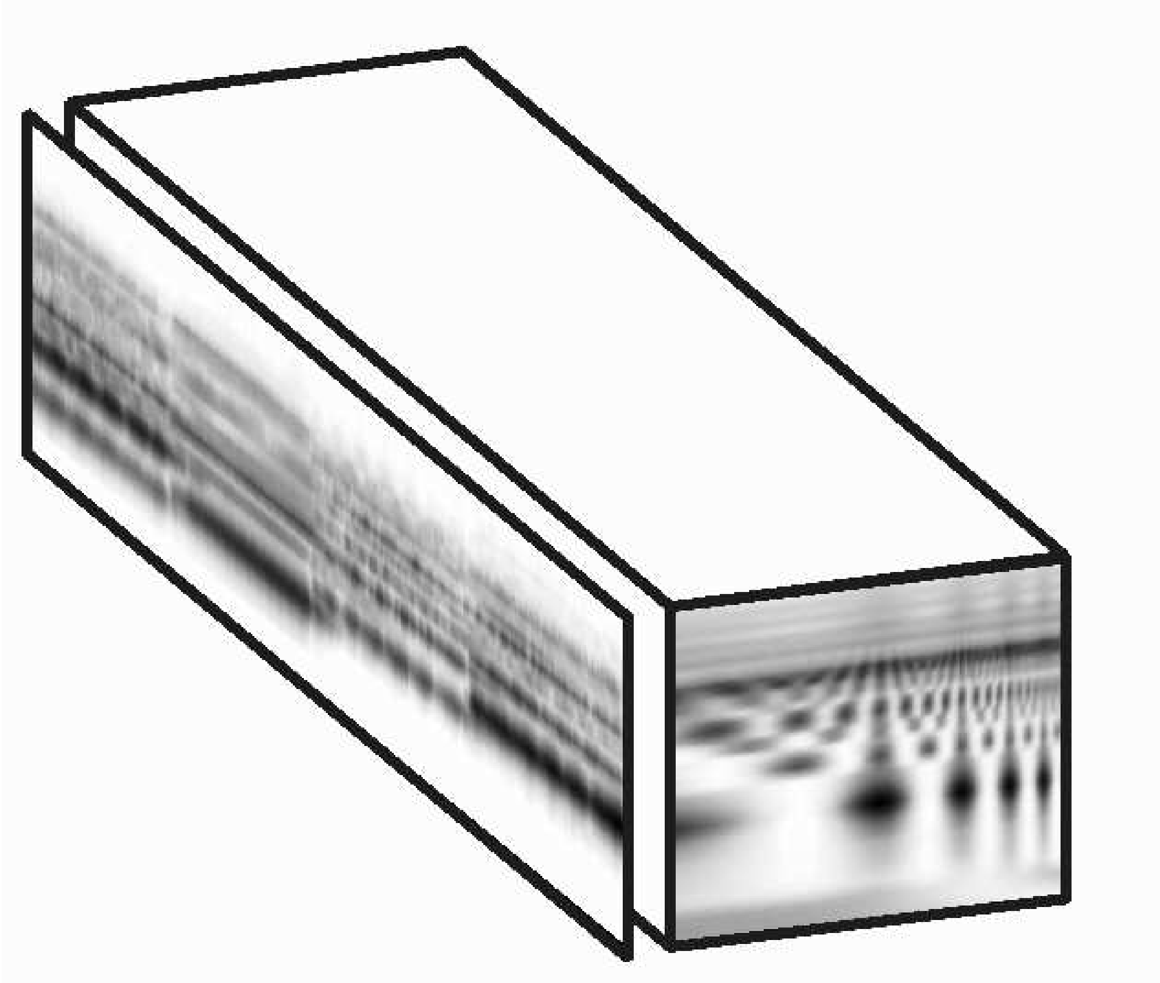
Drs. Meddis and Hewitt's Correlogram



D. P. W. Ellis, *Prediction Driven Computational Auditory Scene Analysis*, PhD thesis, Massachusetts Institute of Technology, June 1996.

Resources

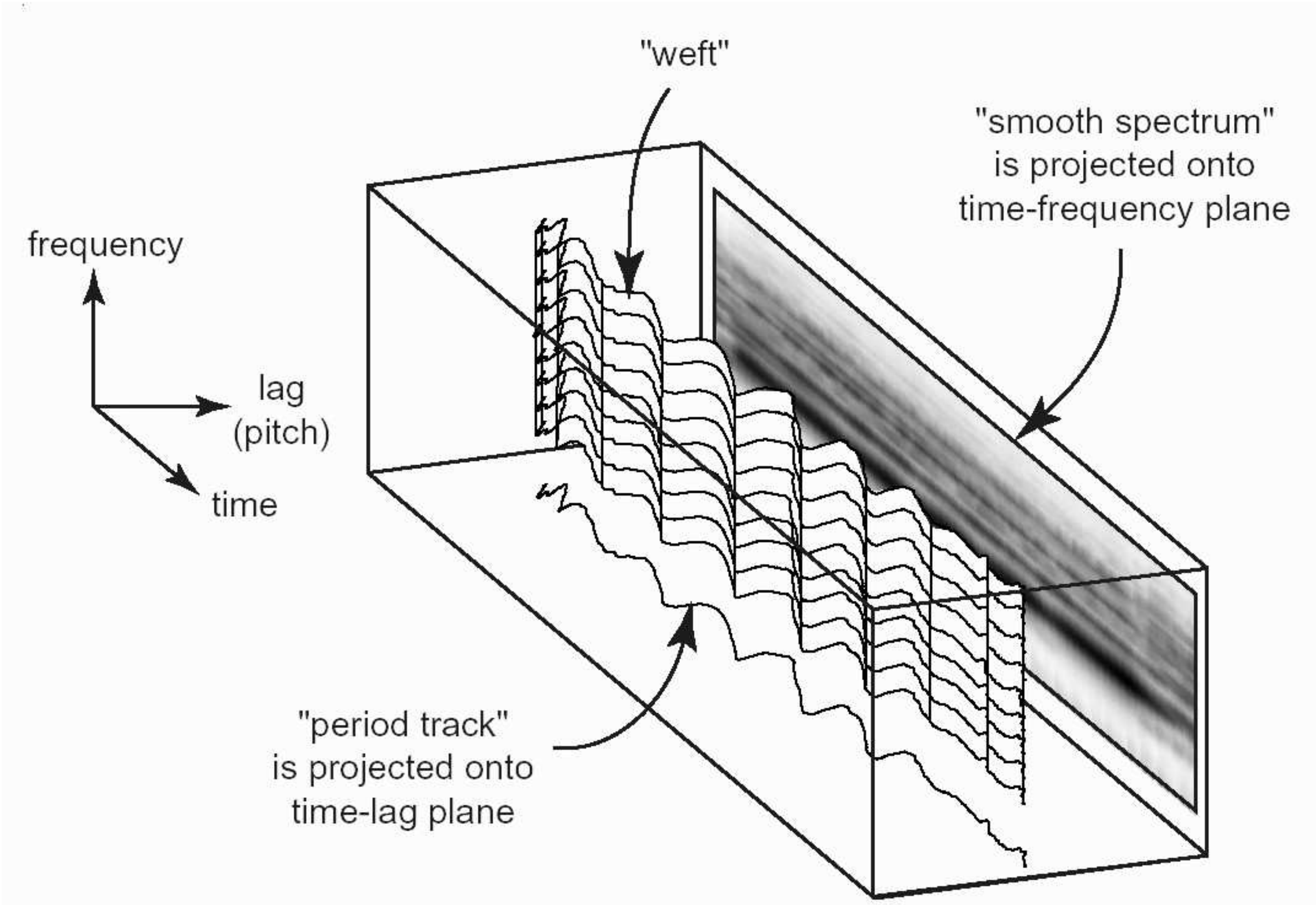
Drs. Meddis and Hewitt's Correlogram



K. D. Martin, *Sound-Source Recognition: A Theory and Computational Model*, PhD thesis, Massachusetts Institute of Technology, June 2000.

Resources

Dr. Ellis' Weft



K. D. Martin, *Sound-Source Recognition: A Theory and Computational Model*, PhD thesis, Massachusetts Institute of Technology, June 2000.

Progress

- Brainstorming Results
- Graphical Score prototype

Progress

Brainstorming Results

- Low Level

amplitude	loudness	spectral centroid
brightness	panning	SNR
presence of harmonics	energy	perceived energy
impulses	compression	

- Medium Level

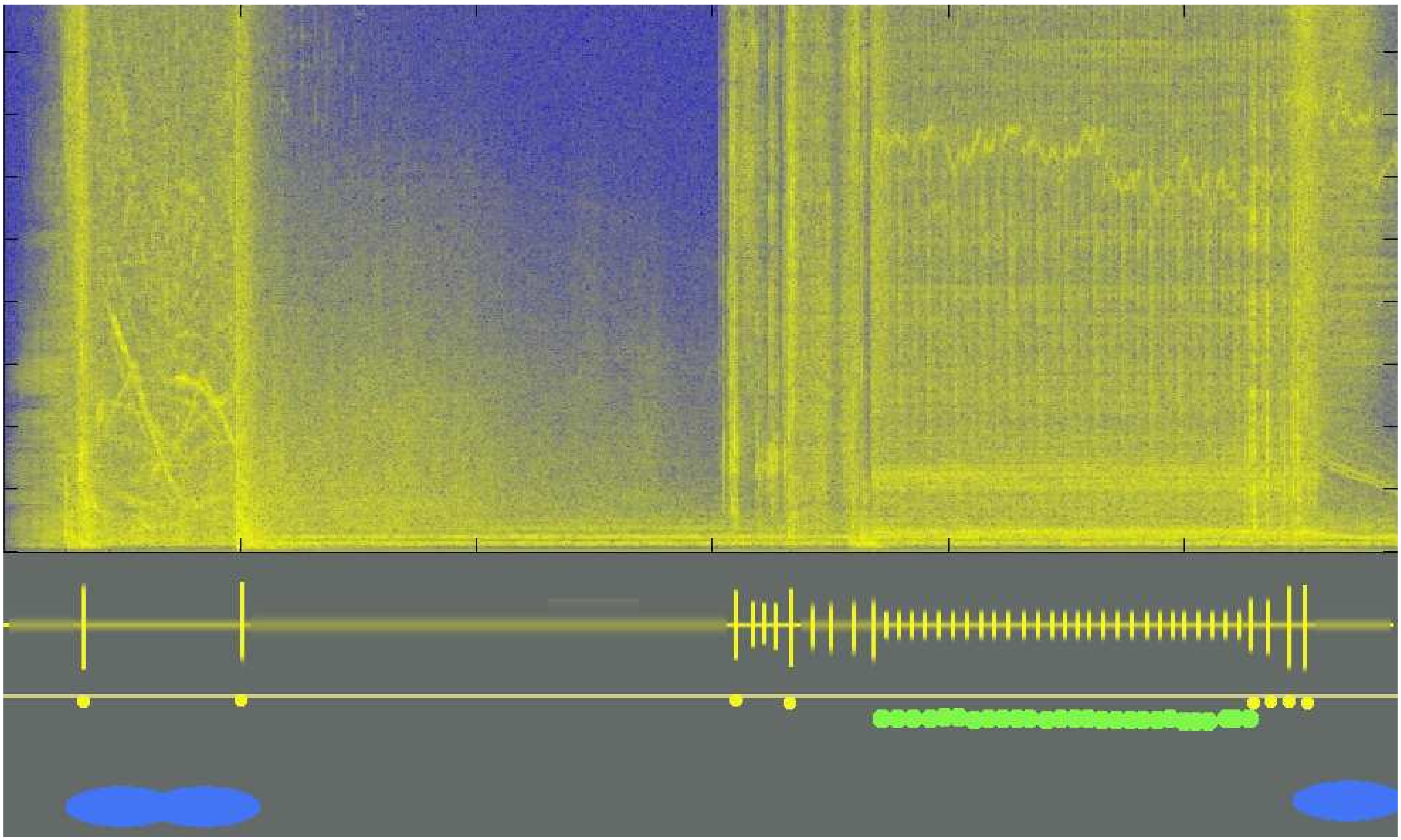
impulsiveness	**self-similarity**	backwardness
commodulation	noise/tonal-ness	fundamental frequencies
ecological sounds	event durations	number of sound sources
harshness	texture/event-ness	energeticness

- High Level

reverb	event types	predictableness
important moments	mood	time signature
number of layers	sound quality	presence of effects
conventionalness	musical key	text

Progress

Graphical Score Prototype



Conclusion

- create algorithms to extract audio properties from recorded electroacoustic music
- use human hearing models by Patterson, Meddis and Hewitt, and Ellis
- begun the process of designing the algorithms