Data Driven Music Understanding

Dan Ellis

**L**aboratory for **R**ecognition and **O**rganization of **S**peech and **A**udio
Dept. Electrical Engineering, Columbia University, NY USA

http://labrosa.ee.columbia.edu/

1. Motivation: What is Music?
2. Eigenrhythms
3. Melodic-Harmonic Fragments
4. Example Applications

Lab
ROSA
Laboratory for the Recognition and
Organization of Speech and Audio

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK
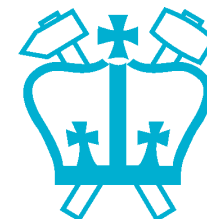
LabROSA Overview
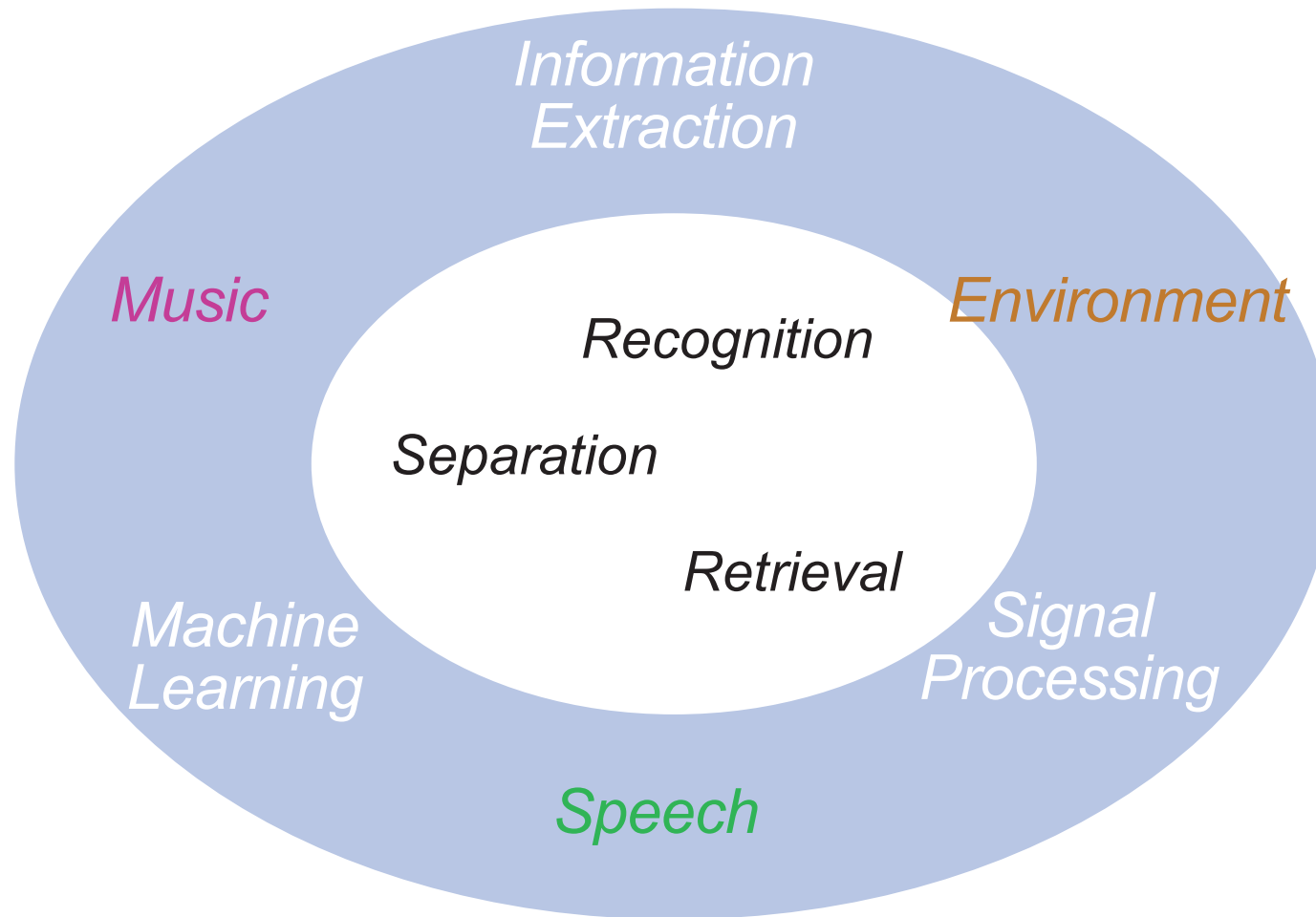
1. Motivation: What is music?

- What does music evoke in a listener's mind?

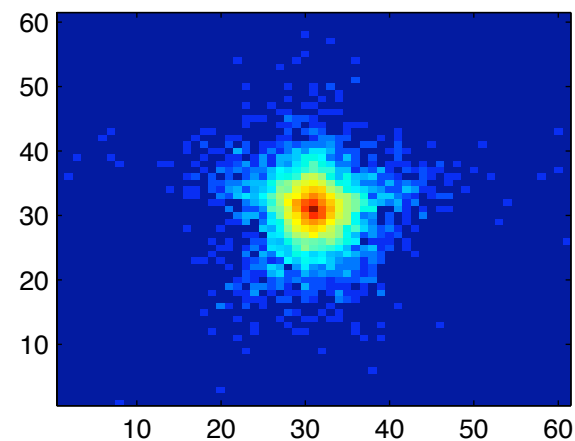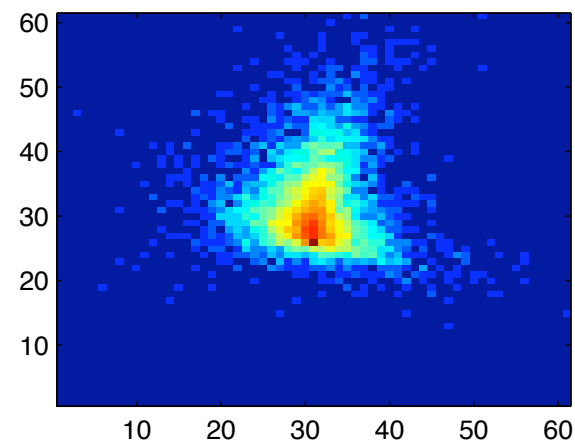- Which are the things that we call "music"?

Oodles of Music

- What can you do with a million tracks?

Re-use in Music

- What are the most popular chord progressions in pop music?
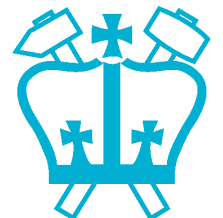
Scatter of PCA(3:6) of 12x16 beatchroma

Lab ROSA
Laboratory for the Recognition and
Organization of Speech and Audio
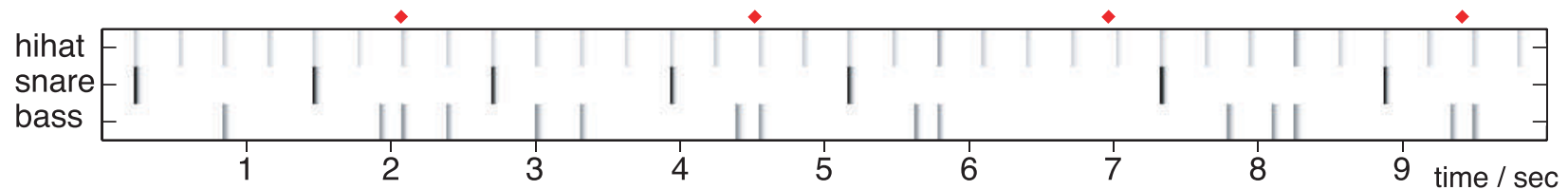
COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Potential Applications

- Compression

- Classification

- Manipulation

2. Eigenrhythms: Drum Track Structure

- To first order,
  all pop music has the same beat:
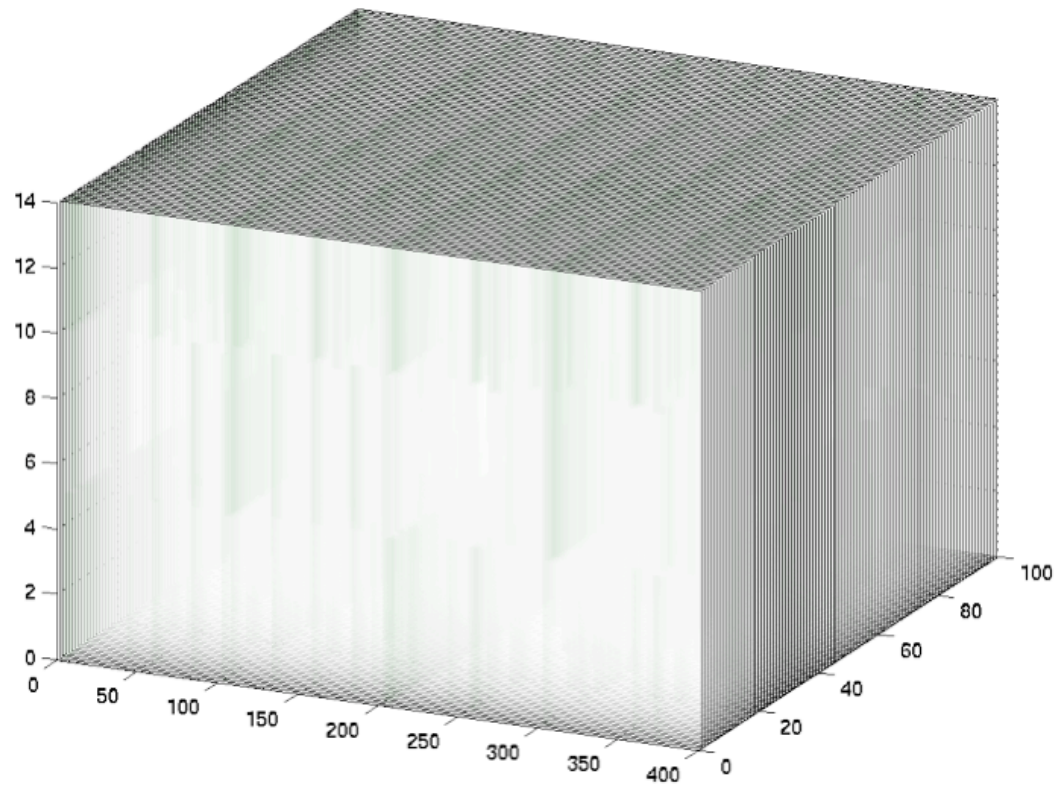

- Can we learn this from examples?

Basis Sets

- Combine a few basic patterns to make a larger dataset

data $\quad$ **X** $\quad$ **=** $\quad$ **W** $\quad$ **×** $\quad$ **H** $\quad$ *patterns*

weights

Lab ROSA
Laboratory for the Recognition and
Organization of Speech and Audio

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK
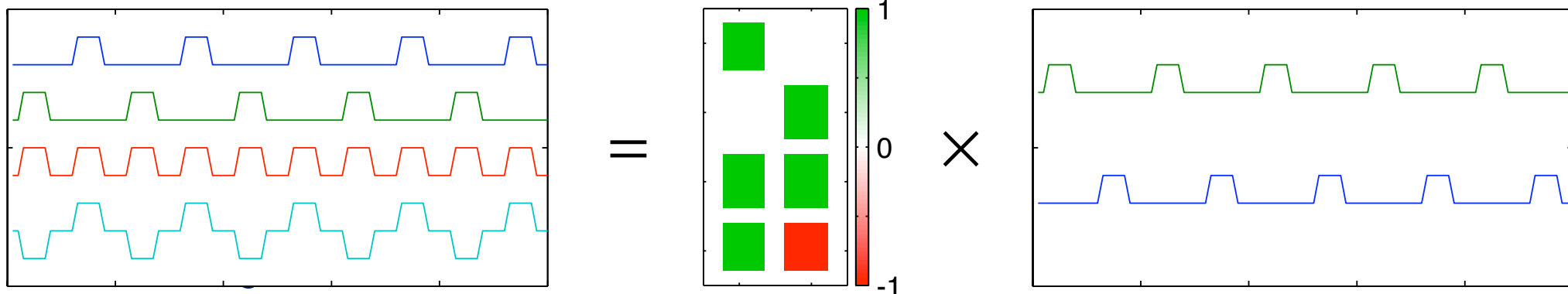
Drum Pattern Data

- Tempo normalization + downbeat alignment

NMF Eigenrhythms
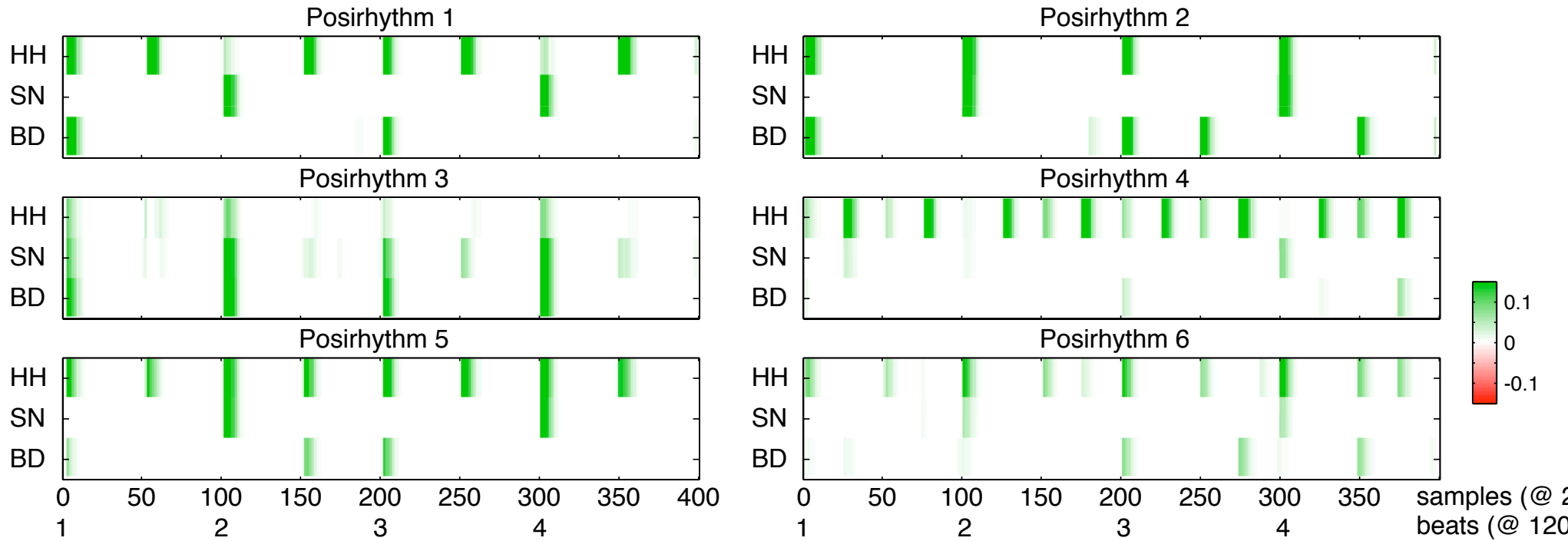
- Nonnegative: only add beat-weight

Eigenrhythm BeatBox

- Resynthesize rhythms from eigen-space
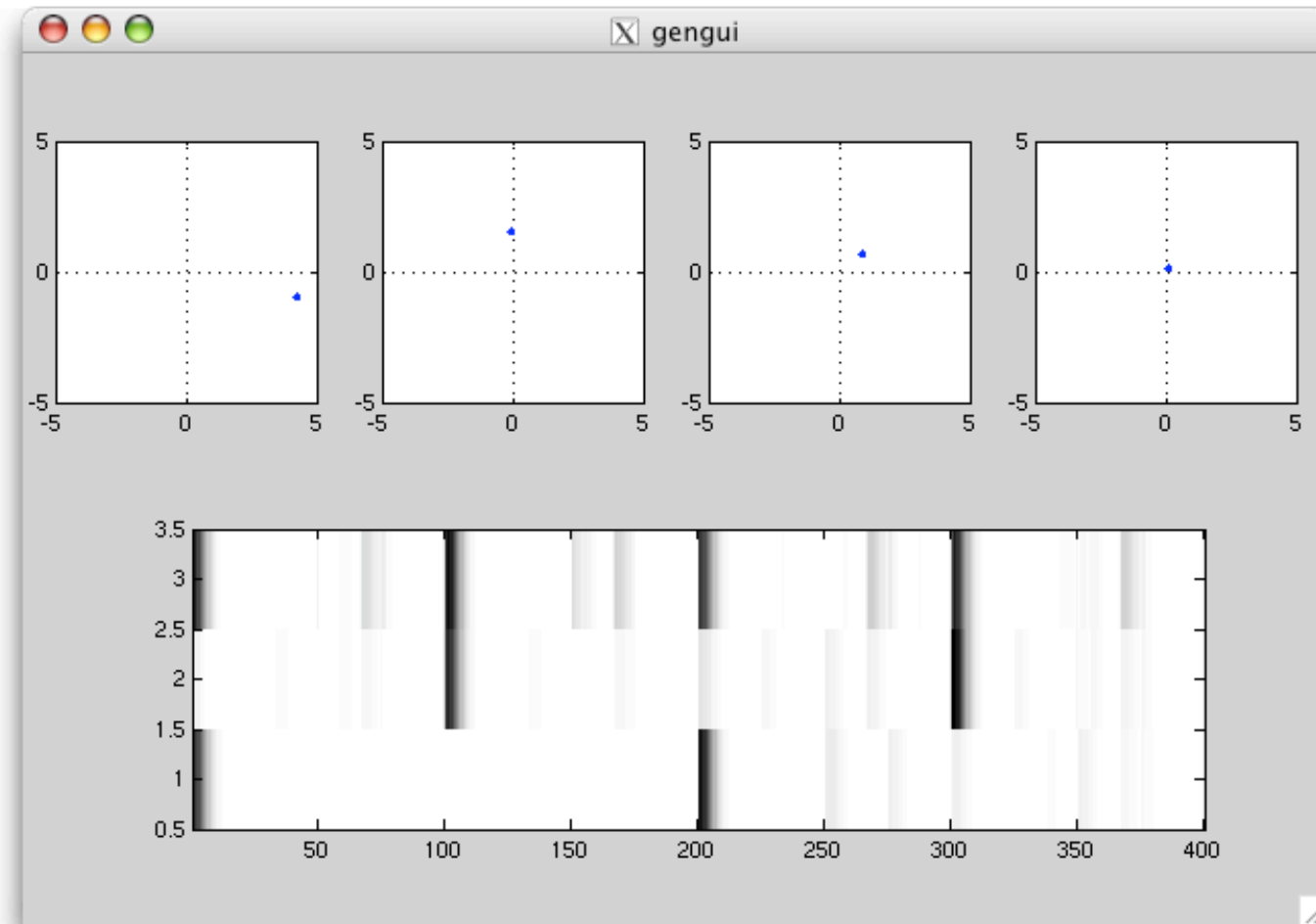
3. Melodic-Harmonic Fragments

- How similar are two pieces?
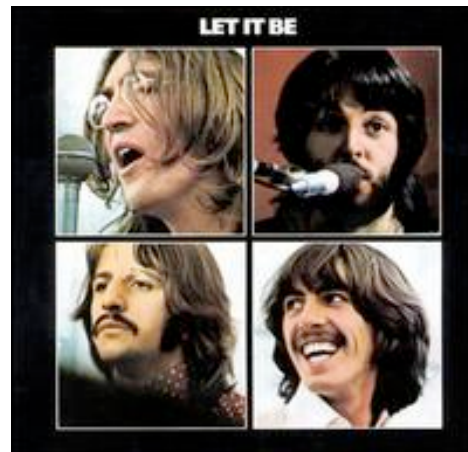
- Can we find all the pop-music clichés?
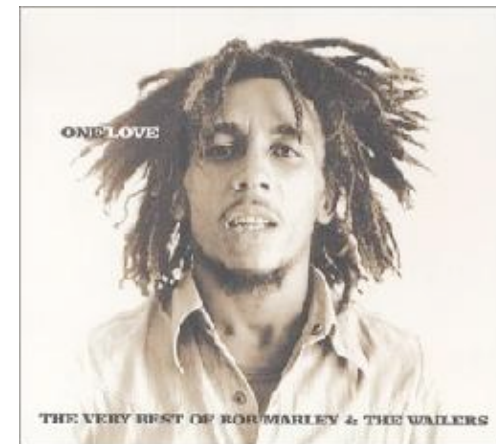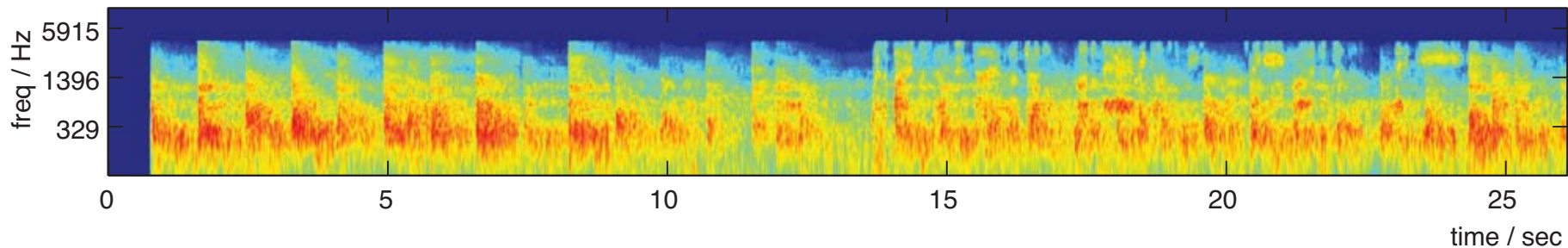
MFCC Features

- Used in speech recognition
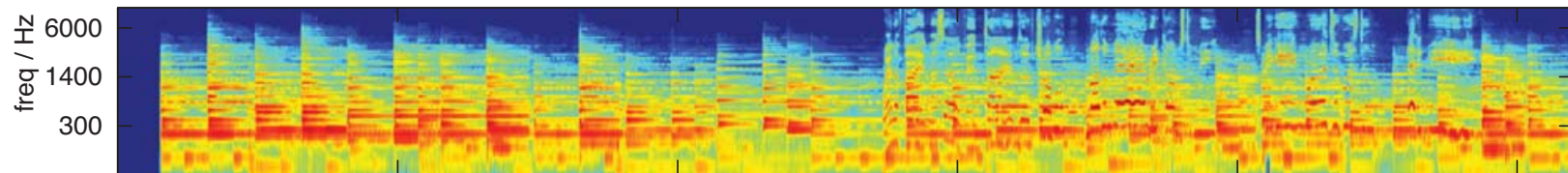
Let It Be (LIB-1) - log-freq specgram
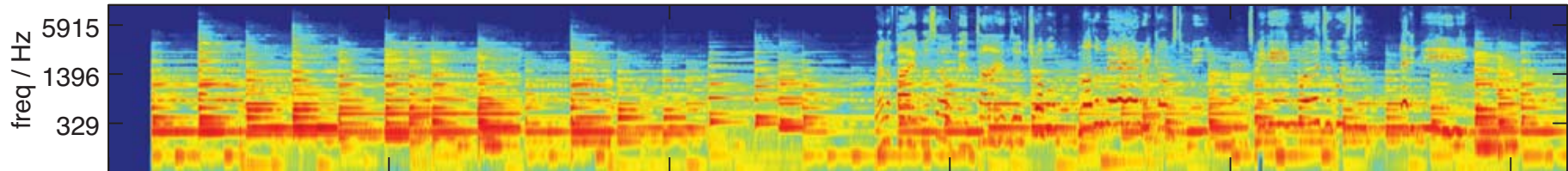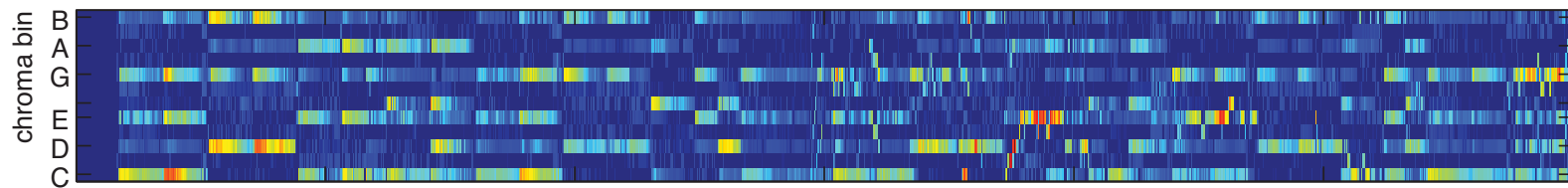

MFCCs
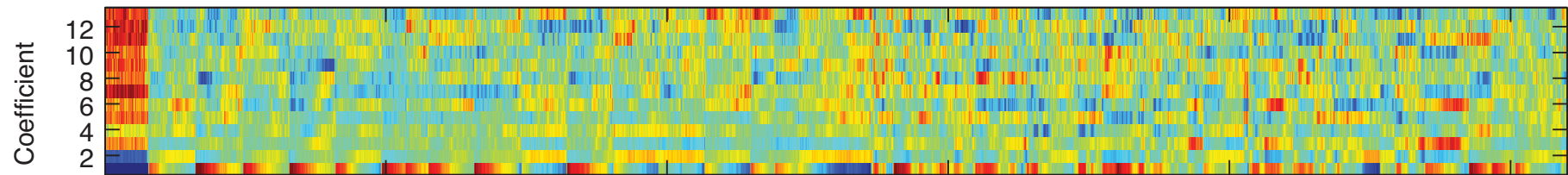
Noise excited MFCC resynthesis (LIB-2)

time / sec

Lab ROSA
Laboratory for the Recognition and
Organization of Speech and Audio

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Chroma Features

- To capture "musical" content

Let It Be - log-freq specgram (LIB-1)

Chroma features

Shepard tone resynthesis of chroma (LIB-3)
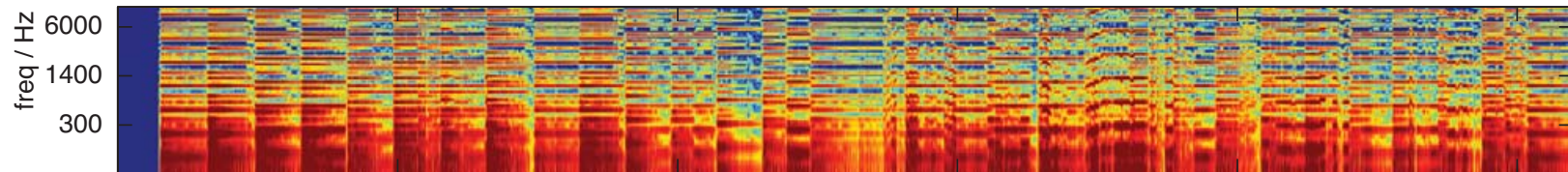
MFCC-filtered shepard tones (LIB-4)
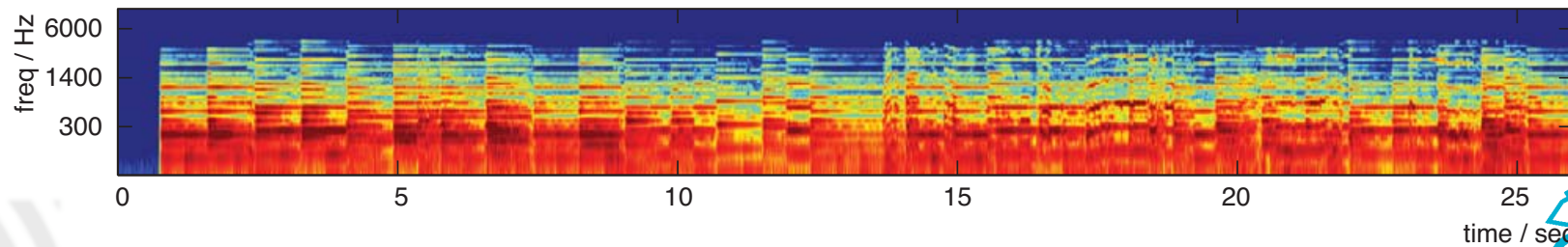
Beat-Synchronous Chroma

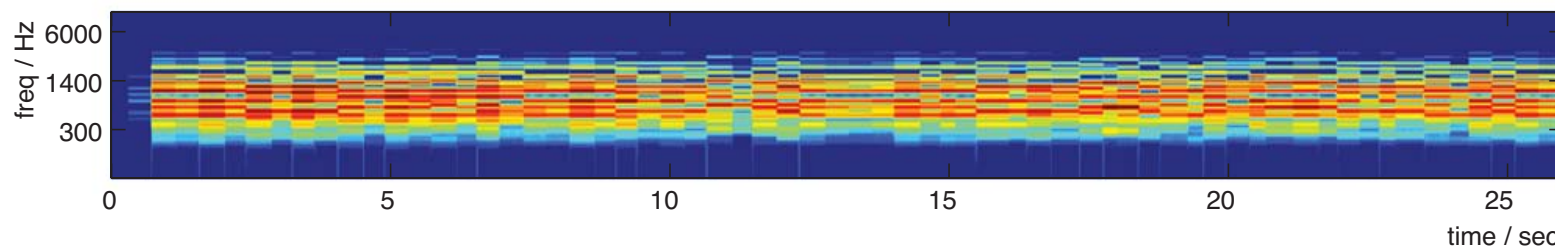- **Compact** representation of harmonies

Let It Be - log-freq specgram (LIB-1)
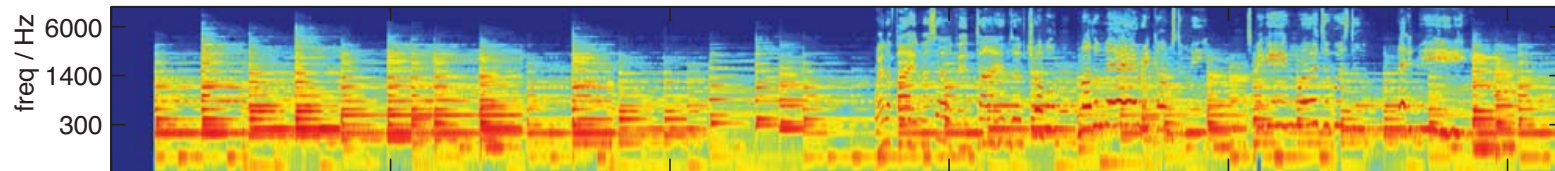
Onset envelope + beat times
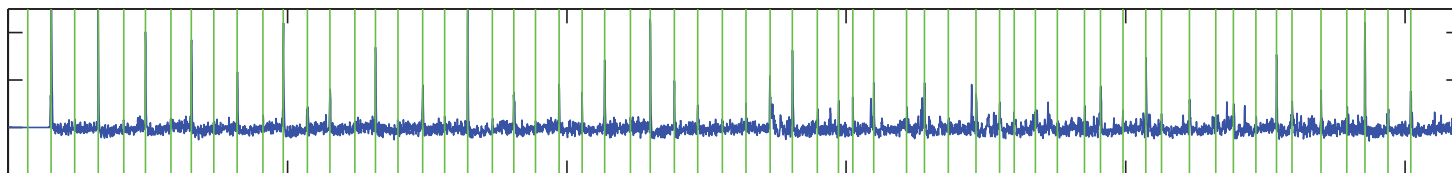
Beat-synchronous chroma
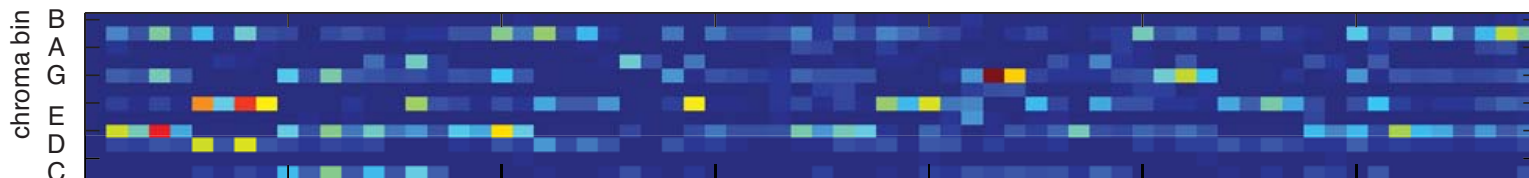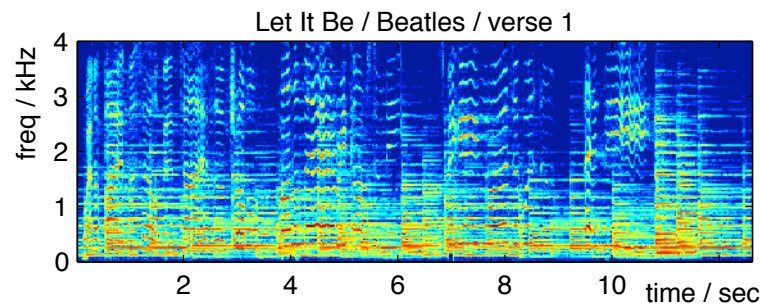
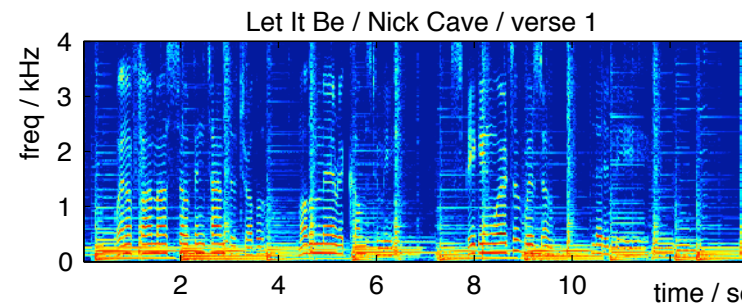Beat-synchronous chroma + Shepard resynthesis (LIB-6)
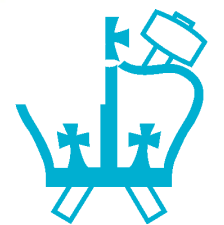
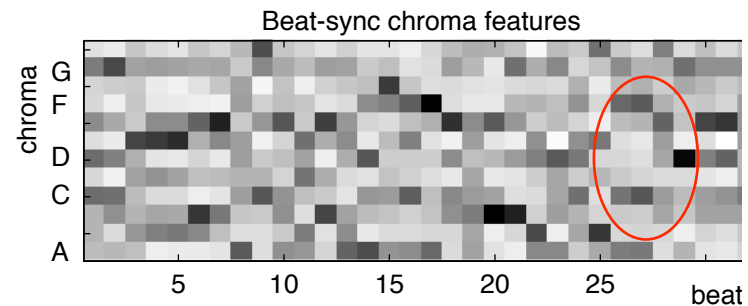# Finding 'Cover Songs'

- ## Little similarity in surface audio...

Let It Be - The Beatles
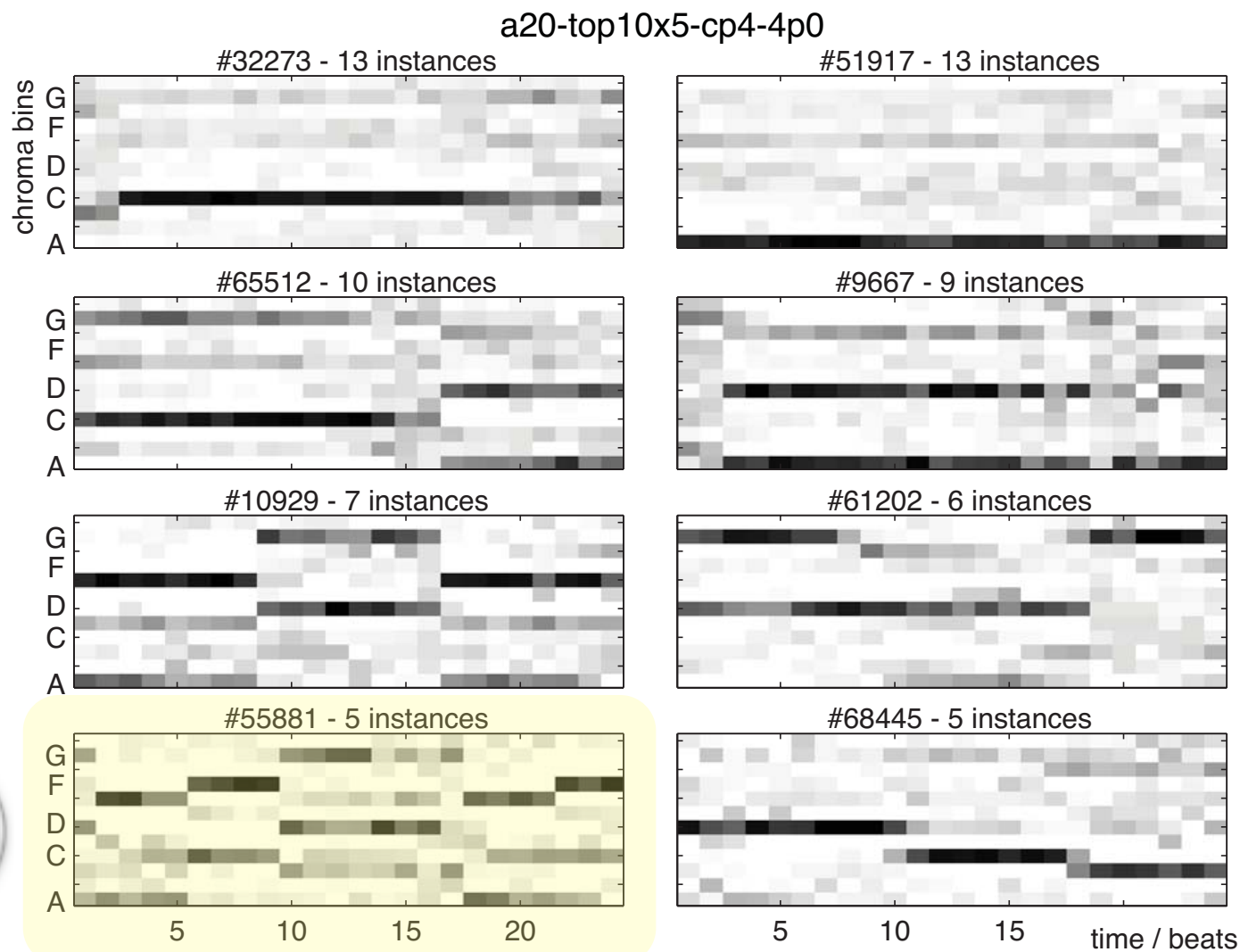
Let It Be - Nick Cave

Let It Be / Beatles / verse 1

Let It Be / Nick Cave / verse 1

- ## .. but appears in beat-chroma

Beat-sync chroma features
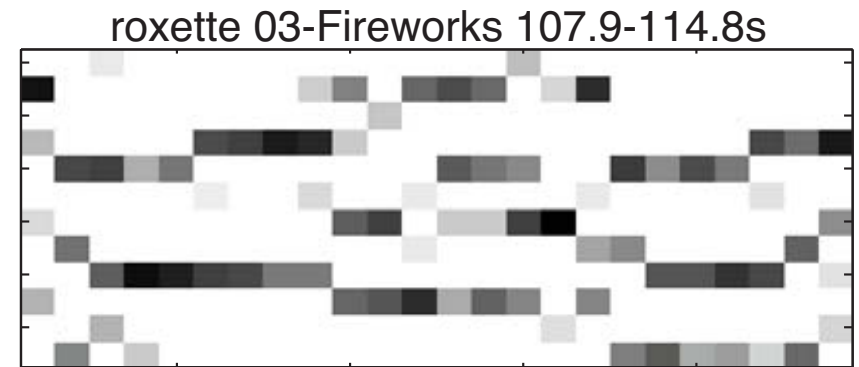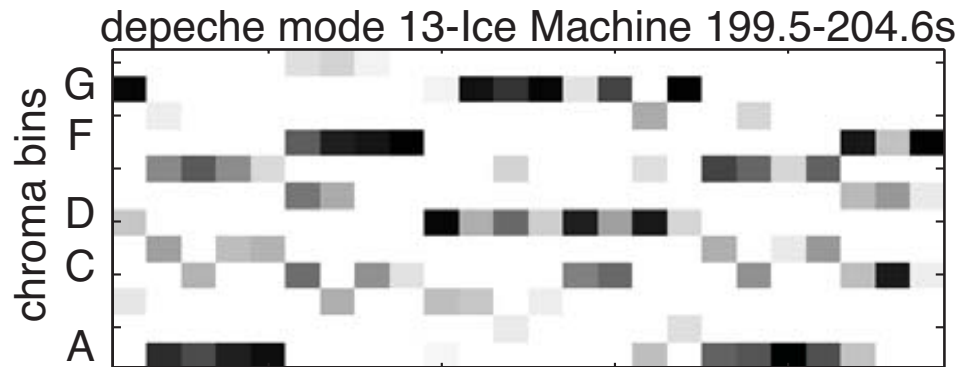

Beat-sync chroma features
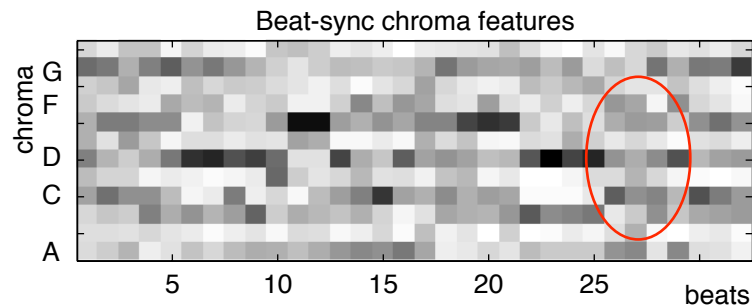
Finding Common Fragments

- Cluster beat-synchronous chroma patches

a20-top10x5-cp4-4p0

Clustered Fragments
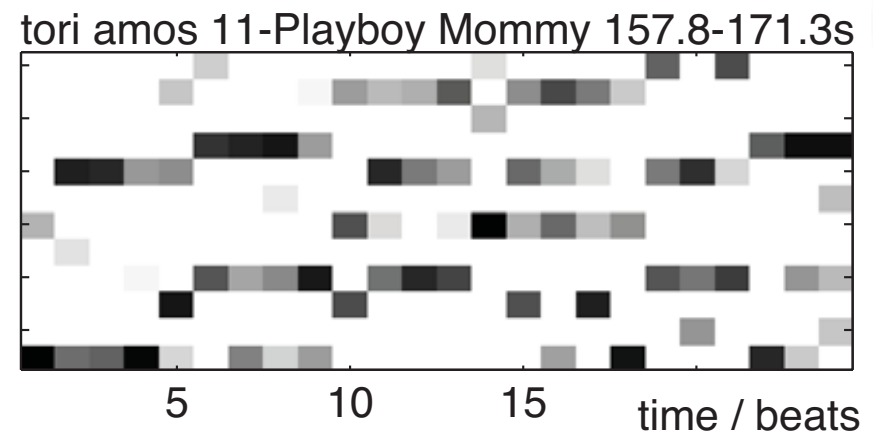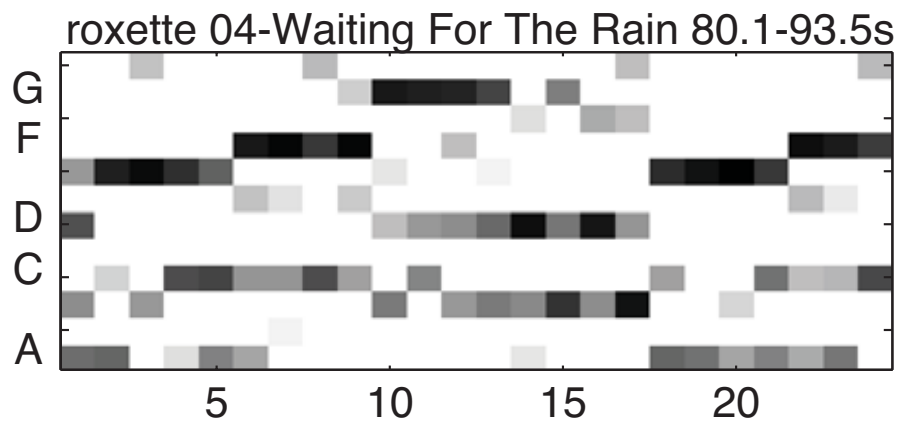
- ...for a dictionary of common themes?

Example Applications: Music Discovery

Berenzweig & Ellis '03

- **Connecting listeners to musicians**

Playlist Generation

- **Incremental** learning of listeners' preferences
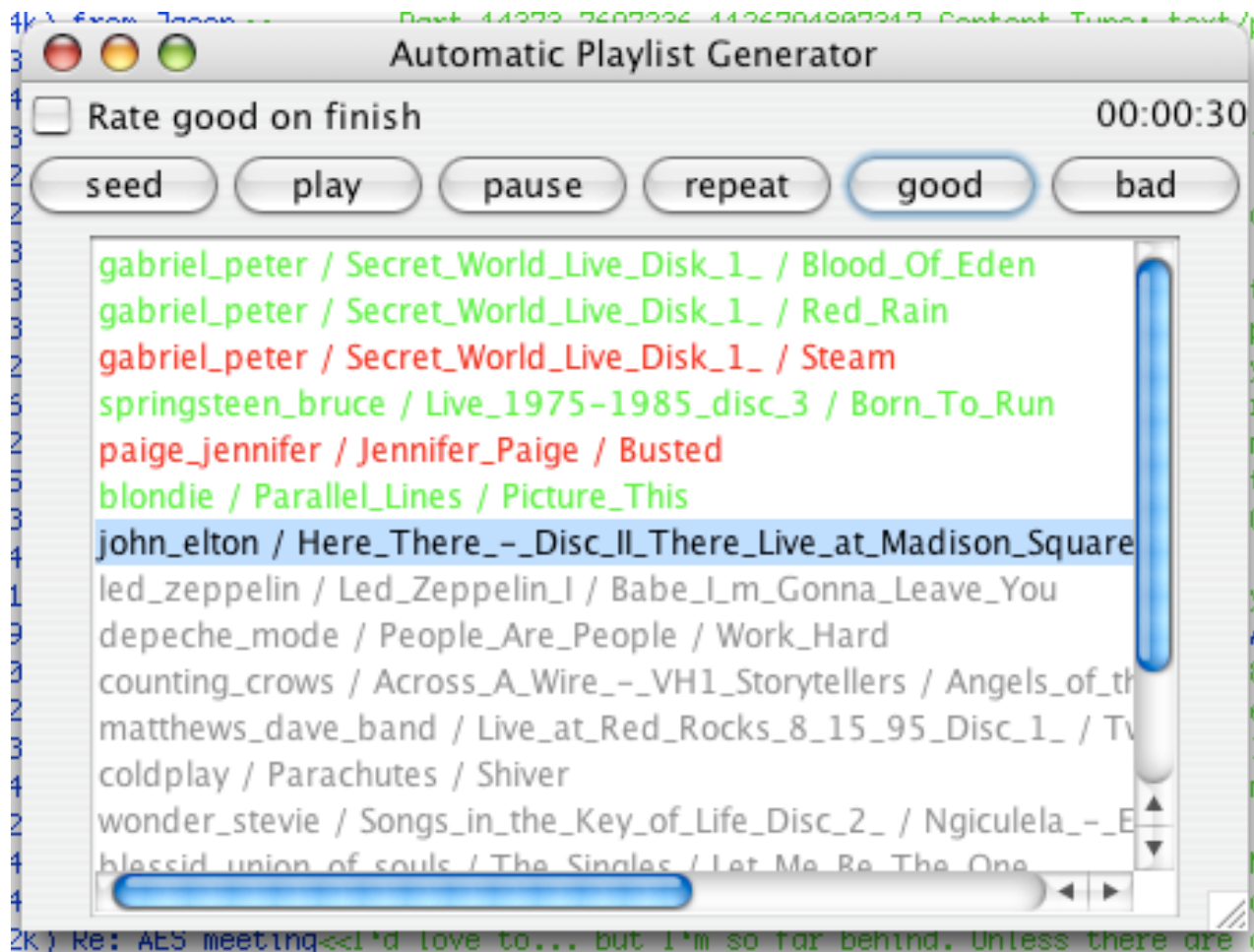
MajorMiner: Music Tagging

- Describe music using words

Music Transcription

Training data and features:
- MIDI, multi-track recordings, playback piano, & resampled audio (less than 28 mins of train audio).
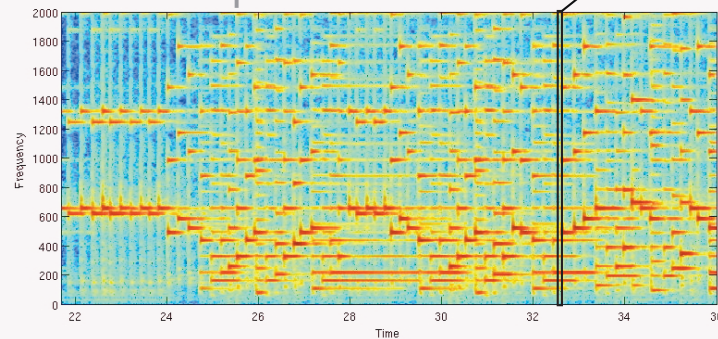- Normalized magnitude STFT.

Classification:
- N-binary SVMs (one for ea. note).
- Independent frame-level classification on 10 ms grid.
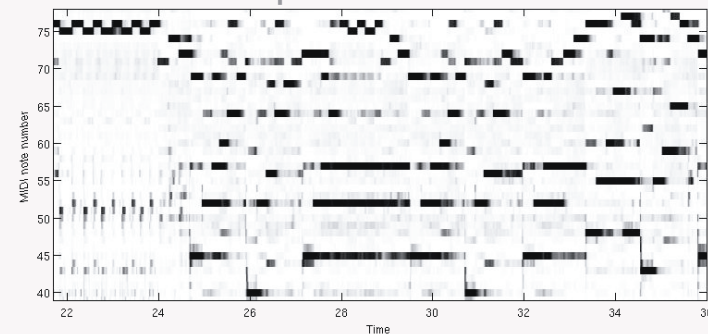- Dist. to class bndy as posterior.

Temporal Smoothing:
- Two state (on/off) independent HMM for ea. note. Parameters learned from training data.
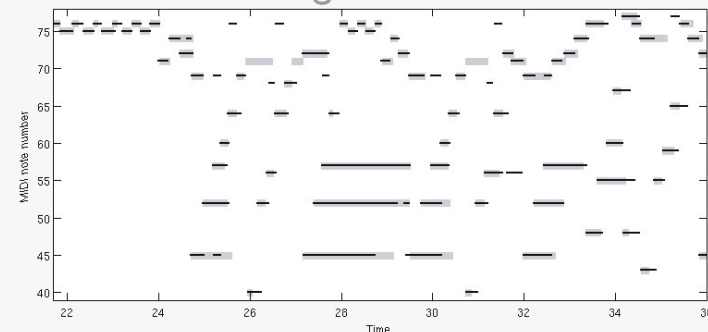- Find Viterbi sequence for ea. note.

feature representation — feature vector

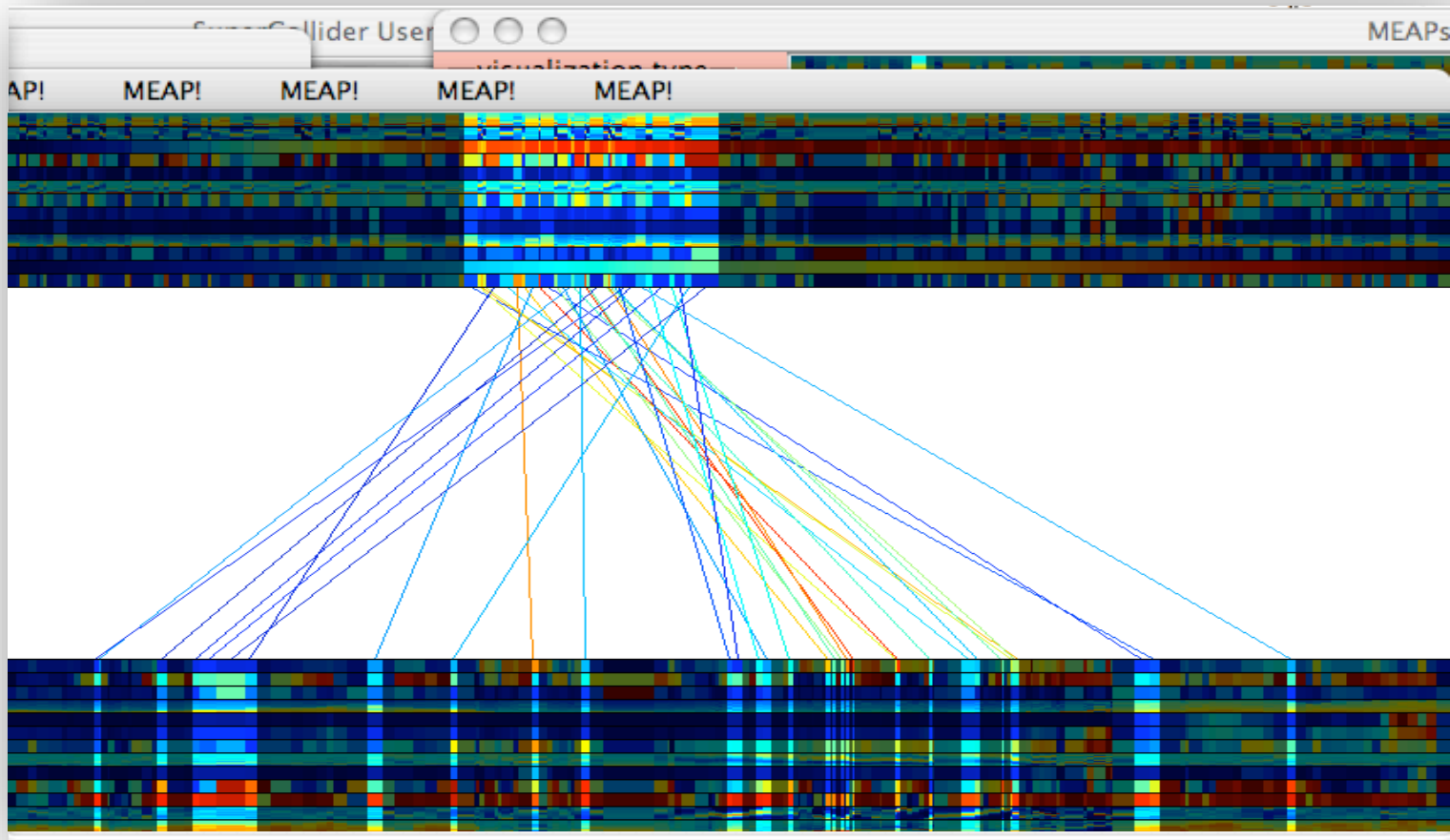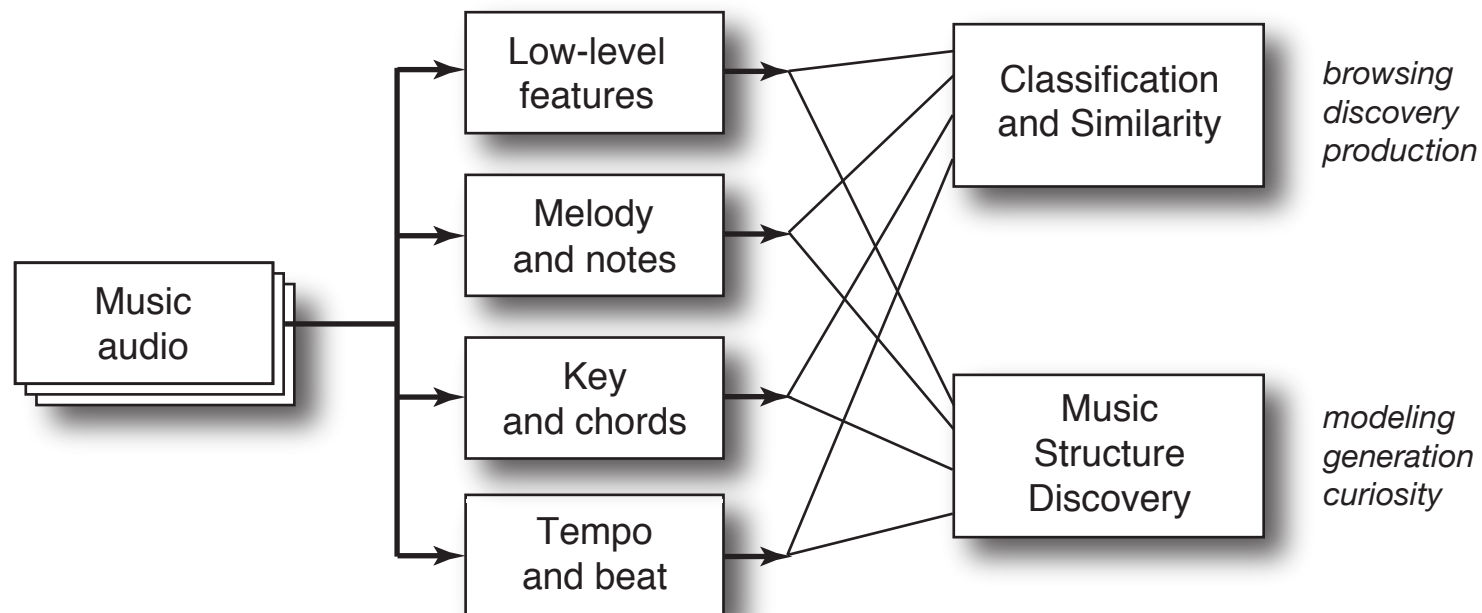classification posteriors

hmm smoothing

Lab ROSA
Laboratory for the Recognition and Organization of Speech and Audio

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

MEAPsoft

- ## Music Engineering Art Projects

  with Douglas Repetto, Ron Weiss, and the rest of the MEAP team

  - collaboration between EE and Computer Music Center

Lab
ROSA
Laboratory for the Recognition and
Organization of Speech and Audio

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Conclusions

- Lots of data
  + noisy transcription
  + weak clustering
  ⇒ musical insights?