Mining for the Meaning of Music

Dan Ellis

Laboratory for Recognition and Organization of Speech and Audio
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
LabROSA Overview

Information Extraction

Music

Environment

Recognition

Separation

Retrieval

Signal Processing

Speech

Machine Learning
I. 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?
• What are the most popular chord progressions in pop music?
Potential Applications

- Compression
- Judgments / 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

\[ \text{data} \quad X \quad = \quad \text{weights} \quad W \quad \times \quad \text{patterns} \quad H \]
Different basis projections

- Principal Component Analysis (PCA)
- Independent Component Analysis (ICA)
- Linear Discriminant Analysis (LDA)
- Nonnegative Matrix Factorization (NMF)
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?
  - I - V - VI - IV
MFCC Features

- Used in speech recognition

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

MFCCs

Noise excited MFCC resynthesis (LIB-2)
Chroma Features

- To capture "musical" content

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

Shepard tone resynthesis of chroma (LIB-3)

MFCC-filtered shepard tones (LIB-4)
Beat-Synchronous Chroma

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

• .. but appears in beat-chroma

Beat-sync chroma features

Beat-sync chroma features
Finding Cover Songs

- Match via global cross-correlation
Finding Common Fragments

• Data mining in beat-chroma database

• .. a large set of excerpts (patches)

![Beat tracking diagram]

Music audio → Chroma features → Key normalization → Landmark identification → Locality Sensitive Hash Table
Landmarks

• Extract patches from starts of phrases

Come Together - Spectrogram, Beat-sync chromogram, and top 10 segment points
Locality Sensitive Hash

- Nearby items go into same hash bucket via multiple random projections

from Slaney & Casey ‘08
Common Fragments

- Cluster beat-synchronous chroma patches

```
#32273 - 13 instances

#65512 - 10 instances

#10929 - 7 instances

#55881 - 5 instances
```

```
#51917 - 13 instances

#9667 - 9 instances

#61202 - 6 instances

#68445 - 5 instances
```
Clustered Fragments

- ...for a dictionary of common themes?
NMF Decomposition

Beat-chroma fragments: 30 Basis NMF decomposition
NMF Decomposition

• Originals and 30-basis NMF reconstructions

• Conclusion?
4. Example Applications: Music Discovery

- Connecting listeners to musicians

![Playola](image)
Playlist Generation

- **Incremental learning of listeners’ preferences**

  Mandel, Poliner, Ellis ‘06
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.
MEAPsoft

- **Music Engineering Art Projects**
  - collaboration between EE and Computer Music Center

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

• Lots of **data**
  + noisy transcription
  + weak clustering
  \[ \Rightarrow \text{musical insights?} \]