

Automatic Drum Transcription

E6820 Project Proposal

Ron Weiss

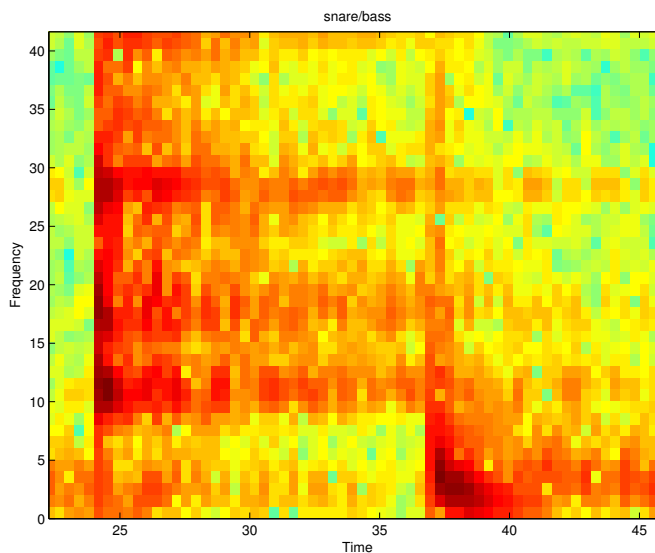
`ronw@ee.columbia.edu`

Motivation

- What
 - Detect drum events in polyphonic music signal and assign class label
- Why
 - Characterize rhythm of particular piece of music
 - Classify/search based on rhythmic similarity
 - Genre classification?

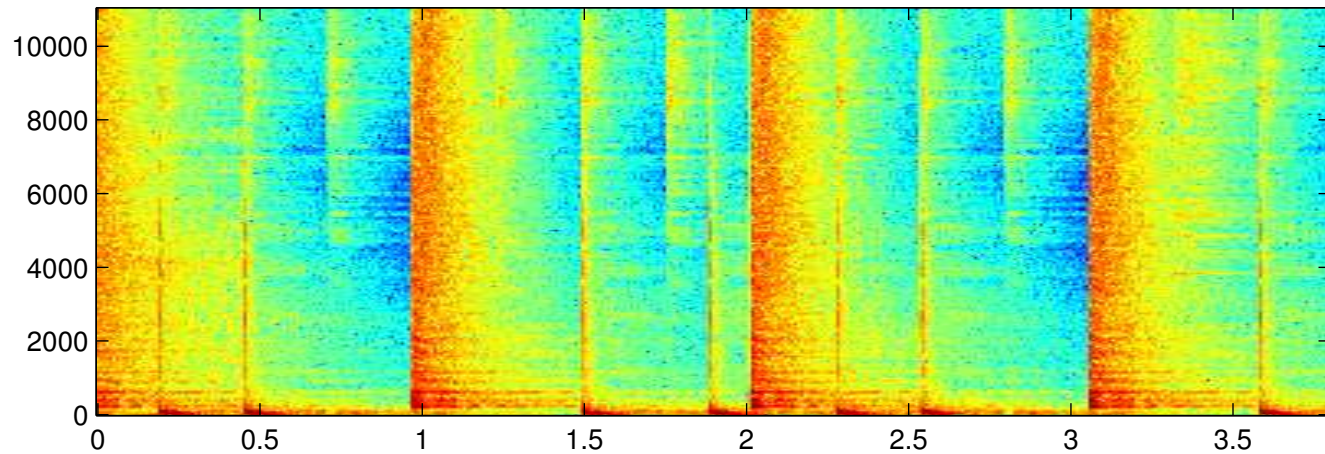
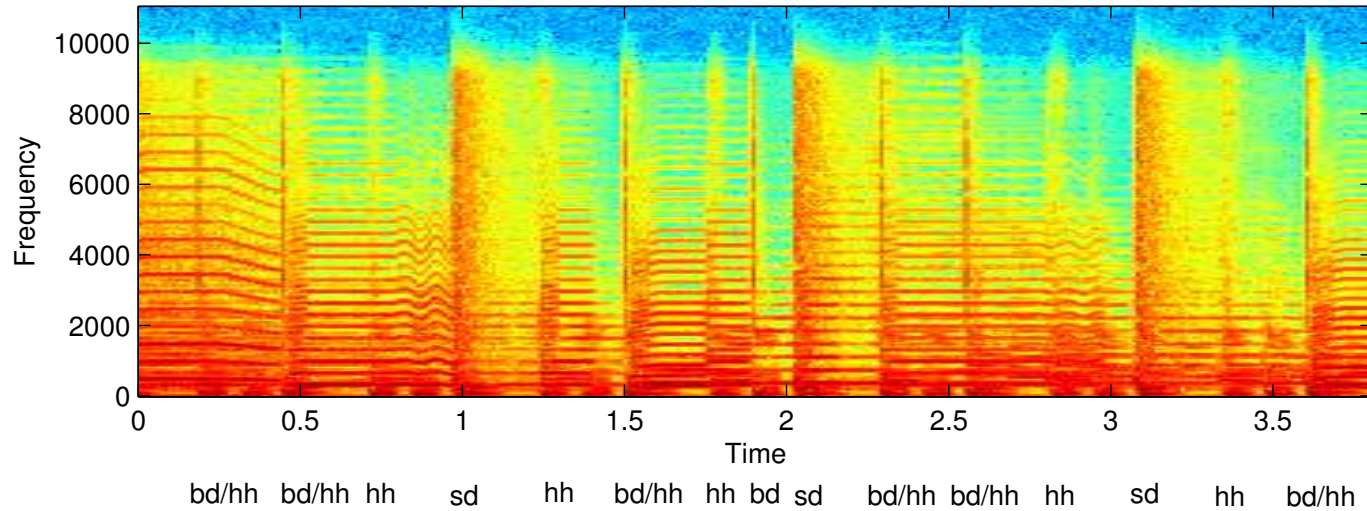
Challenges

- Drums masked by other instruments
- Need to detect simultaneous drum events
- How to characterize different drum sounds?



- Bass/snare drum can be characterized by narrowband spectral peaks at onset
- Hi-hat/cymbals pretty much noise

Challenges



How to deal with interference from other instruments?

Previous Work

- Template matching
 - Begin with seed template sound for each drum class
 - Detect onset times in signal - finds both note attacks and percussion events
 - Median filter to adapt template to actual drum sounds in music
 - Search narrowband STFT at each onset for matches with STFT of template. Compare top few spectral peaks with those of template - mostly ignores other instruments
 - Won't work for noisy drums (hi-hat), works well for bass, snare
 - Template sometimes adapts to non drum sound (e.g. bass guitar note)
 - I already have a working version (sound)

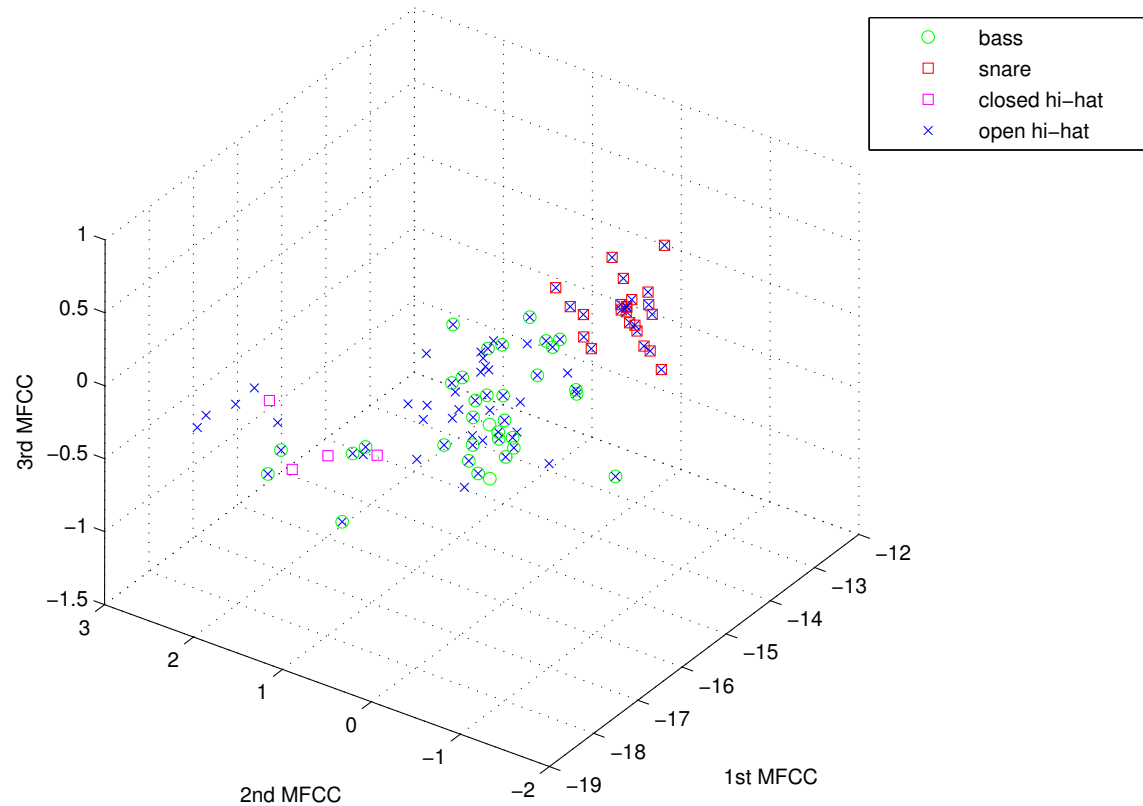
Previous Work

- Sinusoidal modeling
 - Remove sustained notes using sin+noise model. Noise residual contains drums and attack transients (sound)
 - Extract features corresponding to general shape of spectrum at each onset
 - Match general shape of spectrum at each onset
 - But spectral peaks are removed too...

Preliminary Results

- 30 second clip of synthesized MIDI
- Sin+noise model, detect onsets in residual signal
- MFCCs of 100ms window around each onset

Preliminary Results



- First 3 MFCCs show promise for clustering snare drums
- Hi-hats almost always occur with other drums
- Spectral peaks probably needed to better detect bass drum

Goals

- Combine the two methods to transcribe bass drum, snare drum, hi-hats.
- Use features from both domains since some drum sounds are better characterized by the general shape of spectrum vs. narrowband spectral peaks.
- Machine learning to discriminate between drum classes
- Need to investigate features that are good at discriminators
- Train on audio synthesized from MIDI - ground truth labels

References

- [1] K. Yoshii, M. Goto, and H. Okuno. Drum sound description for real-world music using template adaptation and matching methods. In *Proceedings of ISMIR*, 2004.
- [2] J. Sillanpaa, A. Klapuri, J. Seppanen, and T. Virtanen. Recognition of acoustic noise mixtures by combined bottom-up and top-down processing. In *Proceedings of European Signal Processing Conference*, 2000.
- [3] A. Zils, F. Pachet, O. Delerue, and F. Gouyon. Automatic extraction of drum tracks from polyphonic music signals. In *Proceedings of WEDELMUSIC*, December 2002.
- [4] P. Herrera, A. Yeterian, and F. Gouyon. Automatic classification of drum sounds: a comparison of feature selection methods and classification techniques. In *Proceedings of the 2nd International Conference on Music and Artificial Intelligence*, 2002.
- [5] M. Gruhne, C. Uhle, C. Dittmar, and M. Cremer. Extraction of drum patterns and their description within the MPEG-7 high-level-framework. In *Proceedings of ISMIR*, 2004.