Lecture 6

Rhythm Analysis

(some slides are adapted from Zafar Rafii and some figures are from Meinard Mueller)
Definitions for Rhythm Analysis

- **Rhythm**: “movement marked by the regulated succession of strong and weak elements, or of opposite or different conditions.”
  
  ---- Oxford English Dictionary

- **Beat**: basic unit of time in music

- **Tempo**: speed or pace of a given piece, typically measured in beats per minute (BPM)
More Definitions

- **Onset**: single instant marking the beginning of transient
  - Onsets often occur on beats.
- **Attack**: sharp increase of energy
- **Transient**: a short duration with high amplitude within which signal evolves quickly

Waveform of one piano note
More Definitions

- **Measure (or bar):** segment of time defined by a given number of beats

A 4-beat measure drum pattern.

[http://en.wikipedia.org/wiki/Metre_(music)]
More Definitions

- **Meter**: Organization of music into regularly recurring measures of stressed and unstressed beats

Hypermeter: 4-beat measure and 4-measure hypermeasure. Hyperbeats in red. [http://en.wikipedia.org/wiki/Metre_(music)]
Rhythm Analysis Tasks

• Onset Detection
• Beat Tracking
• Tempo Estimation
• Higher-level Structure Analysis
Why is it important?

• Intellectual merit
  – Important component of music understanding
  – Music cognition research

• Broad applications
  – Identify/classify/retrieve by rhythmic similarity
  – Music segmentation/summarization
  – Audio/video synchronization
  – Source separation
Onset Detection

• Signal processing: define a detection function
  – Energy-based
  – Spectral-based
  – Phase-based

• Machine Learning: learn patterns from labeled data
  – Probabilistic models
  – Neural networks
Energy-based Onset Detection

Waveform

Signal Envelope (energy)

\[ E_w^x(n) := \sum_{m=-M}^{M} |x(n+m)w(m)|^2 \]

Envelope Derivative (half-wave rectified)

\[ \Delta_{\text{Energy}}(n) := |E_w^x(n+1) - E_w^x(n)|_{\geq 0} \]

Thresholding → Onsets
Energy-based Onset Detection

• Pros and Cons
  – Simple
  – Works well for percussive sounds
  – Soft onsets by string/wind instruments are hard to detect
  – Tremolo/vibrato can cause false detections

• How to improve
  – Use logarithmic-energy to replace linear energy
  – Perform analysis in different frequency bands, then summarize
Spectral-based Onset Detection

• STFT to get magnitude spectrogram $|\mathcal{X}|$

• (optional) compression
  \[ \mathcal{Y} := F_\gamma(|\mathcal{X}|) = \log(1 + \gamma \cdot |\mathcal{X}|) \]

• Spectral flux:
  - Take derivative w.r.t. time (half-wave rectified)
  \[ \Delta_{\text{Spectral}}(n) := \sum_{k=0}^{K} |\mathcal{Y}(n + 1, k) - \mathcal{Y}(n, k)|_{\geq 0} \]
Spectral-based Onset Detection

• Pros and Cons
  – More complex than energy-based
  – Can weigh different frequencies differently
  – Works better for soft onsets (e.g., legato notes) and polyphonic music
  – Still doesn’t work very well for vibrato
Tempo Estimation

- Tempo = beats / minutes
- Beat tracking is sufficient but not necessary condition for tempo estimation
- How to estimate tempo without tracking beats?

- Idea: look at the regularity of onsets
- Assumptions
  - Onsets mostly occur on beats
  - Tempo is constant within a period of time
Tempo Estimation

- Take the onset strength curve and analyze its periodicity
  - Autocorrelation
  - STFT

Onset strength curve

Onsets

Tempogram
Beat Tracking

• Identify the beat times, i.e., the times to which we tap our feet
  – Detected onsets provide useful but noisy information, since not all onsets are on beats.
  – Estimated tempo tells us the space between two beats, but not the exact locations (i.e., phase).

• How to identify beats?
• To simply the problem, we assume
  – Onsets, especially strong ones, are mostly on beats.
  – Tempo is constant.
Beat Tracking

• A 2-step approach
  – Step 1: Tempo estimation
  – Step 2: Identify beats from onsets using the tempo
    • Create an impulse train (i.e., “comb”) with the tempo
    • Cross-correlate the “comb” with the onset strength curve.
    • The lag that gives us the highest cross-correlation value tells us the beat phase.
Beat Tracking

- A 2-step approach, illustration

Onset strength curve

Combs with the same tempo but different phases

- Problem: **too rigid** about beat spacing
Beat Tracking by Dynamic Programming

- Beat tracking: finding a sequence of beat locations such that they
  - 1) are well aligned with strong onsets
  - 2) mostly regularly spaced

\[ S(B) := \sum_{\ell=1}^{L} \Delta(b_\ell) + \lambda \sum_{\ell=2}^{L} P_\delta(b_\ell - b_{\ell-1}) \]

- Find \( B = (b_1, b_2, \cdots, b_L) \) that maximizes \( S(B) \)

\[ P_\delta(\delta) := -\left( \log_2(\delta / \hat{\delta}) \right)^2 \]

\[ [\text{Ellis, 2007}] \]

Rough estimate of beat spacing
Suppose beat locations are precise to audio frames, and suppose there are \( N \) frames, then how many possible sequences?

\[ 2^N \] (although many are bad ones!)

Can’t enumerate all!

Key idea: reuse calculations by recursion!
Beat Tracking by Dynamic Programming

- Consider a beat sequence $B_n = (b_1, b_2, \ldots, b_L)$ where $b_L = n$.
- Let $D(n)$ be the maximal score over all such sequences ending at $n$.
- Then

  $\begin{align*}
  D(n) &= \Delta(n) + \lambda \hat{P}_{\delta}(n - b_{L-1}) + D(b_{L-1}) \\
  &\text{if } L > 1 \\
  D(n) &= \Delta(n) \\
  &\text{if } L = 1
  \end{align*}$

  recursion
Beat Tracking by Dynamic Programming

• Considering the two cases, we have

\[ D(n) = \Delta(n) + \max \left\{ 0, \max_{m \in [1:n-1]} \left\{ D(m) + \lambda P_\delta(n-m) \right\} \right\} \]

– We can calculate \( D(n) \) from \( D(1) = \Delta(1) \).

• Record the preceding beat

\[ P(n) := \arg\max_{m \in [1:n-1]} \left\{ D(m) + \lambda P_\delta(n-m) \right\} \]

• Best score

\[ S(B^*) = \max_{n \in [0:N]} D(n) \]

• Trace back from \( b_L = n^* \) to get the best sequence
Rhythmic Structure

Beginning of *Another one bites the dust* by Queen.

- One approach: detect onsets; analyze tempo and beats at different levels.
- Another approach: analyze repetition of spectral content
  - Beat spectrum
• Definition
  – Using the autocorrelation function, we can derive the beat spectrum [Foote et al., 2001]

Beginning of *Another one bites the dust* by Queen.
Beat Spectrum

- Use
  - The beat spectrum reveals the hierarchically periodically repeating structure of the audio

Periodicity at the measure level
Sub-periodicity at the kick level
Sub-periodicity at the beat level

Beginning of *Another one bites the dust* by Queen.

Beat Spectrum.
Beat Spectrum

- Calculation
  - Compute the power spectrogram from the audio using the STFT (square of magnitude spectrogram)
beat spectrum

- Calculation
  - Compute the autocorrelation of the rows of the spectrogram
Beat Spectrum

- **Calculation**
  - Compute the mean of the autocorrelations (of the rows)
• Notes
  – The first highest peak in the beat spectrum does not always correspond to the repeating period!
  – The beat spectrum does not indicate where the beats are or when a measure starts!

Beat Spectrum

This is how you find the period...

This is not the period

Beat Spectrum.
Some interesting links

- Dannenberg’s articles on beat tracking: http://www.cs.cmu.edu/~rbd/bib-beattrack.html
- Goto’s work on beat tracking: http://staff.aist.go.jp/m.goto/PROJ/bts.html
- Ellis’ Matlab codes for tempo estimation and beat tracking: http://labrosa.ee.columbia.edu/projects/beattrack/
- MIREX’s annual evaluation campaign for Music Information Retrieval (MIR) algorithms, including tasks such as onset detection, tempo extraction, and beat tracking: http://www.music-ir.org/mirex/wiki/MIREX_HOME