Rhythm Analysis in Music

EECS 352: Machine Perception of Music & Audio
Some Definitions

• **Rhythm**
  – “movement marked by the regulated succession of strong and weak elements, or of opposite or different conditions.” [OED]
Some Definitions

• **Beat**
  – Basic unit of time in music
Some Definitions

• **Tempo**
  – Speed or pace of a given piece, typically measured in beats per minute (BPM)
Some Definitions

• **Bar (or measure)**
  – Segment of time defined by a given number of beats


A 4-beat measure drum pattern.
Some Definitions

• **Meter (or metre)**
  – 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)]
Some Applications

– Onset detection
– Tempo estimation
– Beat tracking
– Higher-level structures
Practical Interest

– Identify/classify/retrieve by rhythmic similarity
– Music segmentation/summarization
– Audio/video synchronization
– And... source separation!
Intellectual Interest

– “Music understanding” [Dannenberg, 1987]
– Music perception
– Music cognition
– And... Fun!
Onset Detection (what?)

– Identify the starting times of musical elements
– E.g., notes, drum sounds, or any sudden change
– See *novelty curve* [Foote, 2000]

Beginning of *Another one bites the dust* by Queen.
Onset Detection (how?)

– Analyze amplitude (drums have high energy!)
– Analyze other cues (e.g., spectrum, pitch, phase)
– Analyze self-similarity (see similarity matrix)

Beginning of Another one bites the dust by Queen.
Tempo Estimation (what?)

– Identify periodic or quasi-periodic patterns
– Identify some period of repetition
– See beat spectrum [Foote et al., 2001]

Beginning of Another one bites the dust by Queen.
Tempo Estimation (how?)

– Analyze periodicities using the *autocorrelation*
– Compare the onsets with a bank of comb filters
– Use the Short-Time Fourier Transform (STFT)

Beginning of *Another one bites the dust* by Queen.
Beat Tracking (what?)

– Identify the beat times
– Identify the times to which we tap our feet
– See (also) beat spectrum [Foote et al., 2001]

Beginning of Another one bites the dust by Queen.
Beat Tracking (how?)

– Find optimal beat times given onsets and tempo
– Use Dynamic Programming [Ellis, 2007]
– Use Multi-Agent System [Goto, 2001]

Beats at the kick-snare level

Beginning of Another one bites the dust by Queen.
Higher-level Structures (what?)

– Rhythm, meter, etc.
– “Music understanding”
– See (again) *beat spectrum* and *similarity matrix*

Beginning of *Another one bites the dust* by Queen.
Higher-level Structures (how?)

– Extract onsets, tempo, beat
– Use/assume additional knowledge
– E.g., how many beats per measure? Etc.

Beginning of *Another one bites the dust* by Queen.
State-of-the-Art

• 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
  – 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
The Autocorrelation Function

• Definition
  – Cross-correlation of a signal with itself = measure of self-similarity as a function of the time lag

![Autocorrelation plot](image)
The Autocorrelation Function

• Application
  – Identify repeating patterns
  – Identify periodicities

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

Periodicity of about 4 s
The Autocorrelation Function

• Application
  – Identify repeating patterns
  – Identify periodicities

Periodic signal + random signal.

Autocorrelation plot.

Zafar Rafii, Winter 2014
The Autocorrelation Function

• Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i + j) \]
The Autocorrelation Function

• Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i + j) \]

\[ x(i) = \begin{bmatrix} 1 & 3 & 1 & 3 \end{bmatrix} \]

\[ x(i + 0) = \begin{bmatrix} 1 & 3 & 1 & 3 \end{bmatrix} \]

\[ a(j) = \begin{bmatrix} \_ & \_ & \_ & \_ \end{bmatrix} \]
The Autocorrelation Function

• Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i + j) \]

\[
\begin{array}{cccc}
1 & 3 & 1 & 3 \\
1 & 3 & 1 & 3 \\
\end{array}
\]

\[ a(j = 0) = \frac{1 + 9 + 1 + 9}{4} = 5 \]

\[
\begin{array}{cccc}
5 & & & \\
0 & 1 & 2 & 3 \\
\end{array}
\]

Zafar Rafii, Winter 2014
The Autocorrelation Function

• Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i + j) \]

\[ x(i) = \begin{bmatrix} 1 & 3 & 1 & 3 \end{bmatrix} \]

\[ x(i + 1) = \begin{bmatrix} 1 & 3 & 1 & 3 \end{bmatrix} \]

\[ a(j) = \begin{bmatrix} 5 \end{bmatrix} \]
The Autocorrelation Function

• Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i+j) \]

\[
\begin{array}{c}
\text{x(i)} = 1 & 3 & 1 & 3 \\
\text{x(i+1)} = 1 & 3 & 1 & 3 \\
\end{array}
\]

\[ a(j = 1) = \frac{3 + 3 + 3}{3} = 3 \]

\[
\begin{array}{c}
\text{a(j)} = 5 & 3 & 0 & 2 & 3 \\
\end{array}
\]
The Autocorrelation Function

• Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i + j) \]

\[ x(i) = \begin{bmatrix} 1 & 3 & 1 & 3 \end{bmatrix} \]

\[ x(i + 2) = \begin{bmatrix} 1 & 3 & 1 & 3 \end{bmatrix} \]

\[ a(j) = \begin{bmatrix} 5 & 3 & \_ & \_ \end{bmatrix} \]
The Autocorrelation Function

- Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i+j) \]

\[
x(i) = \begin{bmatrix} 1 & 3 & 1 & 3 \end{bmatrix}
\]

\[
x(i+2) = \begin{bmatrix} 1 & 3 & 1 & 3 \end{bmatrix}
\]

\[ a(j = 2) = \frac{1 + 9}{2} = 5 \]

\[
a(j) = \begin{bmatrix} 5 & 3 & 5 \end{bmatrix}
\]

samples

lags
The Autocorrelation Function

- Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i + j) \]

\[
x(i) = \begin{bmatrix}
1 & 3 & 1 & 3 \\
\end{bmatrix}
\]

\[
x(i + 3) = \begin{bmatrix}
1 & 3 & 1 & 3 \\
\end{bmatrix}
\]

\[
a(j) = \begin{bmatrix}
5 & 3 & 5 & \\
0 & 1 & 2 & 3 \\
\end{bmatrix}
\]
The Autocorrelation Function

• Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i+j) \]

\[ x(i) = \begin{bmatrix} 1 & 3 & 1 & 3 \end{bmatrix} \]

\[ x(i + 3) = \begin{bmatrix} 1 & 3 & 1 & 3 \end{bmatrix} \]

\[ a(j = 3) = \frac{3}{1} = 3 \]

\[ a(j) = \begin{bmatrix} 5 & 3 & 5 & 3 \end{bmatrix} \]
The Autocorrelation Function

• Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i + j) \]

\[ x(i) = \begin{array}{cccc}
1 & 3 & 1 & 3 \\
\end{array} \]

\[ a(j) = \begin{array}{cccc}
5 & 3 & 5 & 3 \\
\end{array} \]
The Autocorrelation Function

• Calculation

\[ a(j) = \frac{1}{n-j} \sum_{i=1}^{n-j} x(i)x(i + j) \]

Periodic sequence of 2 samples

Lag 0 = similarity with itself

Period of 2 samples
The Autocorrelation Function

- Notes
  - The autocorrelation generally starts at lag 0 = similarity of the signal with itself
  - Wiener-Khinchin Theorem: Power Spectral Density = Fourier Transform of autocorrelation
Foote’s 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.
Foote’s Beat Spectrum

• Application
  – The beat spectrum reveals the hierarchically periodically repeating structure

Periodicity at the measure level
Periodicity at the kick level
Periodicity at the beat level

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

• Calculation
  – Compute the power spectrogram from the audio using the STFT (square of magnitude spectrogram)
Foote’s Beat Spectrum

• Calculation
  – Compute the autocorrelation of the rows (i.e., the frequency channels) of the spectrogram
Foote’s Beat Spectrum

• Calculation
  – Compute the mean of the autocorrelations (of the rows)
Foote’s Beat Spectrum

• 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!
Foote’s Beat Spectrum

• Notes
  – The beat spectrum can also be calculated using the *similarity matrix* [Foote et al., 2001]
  – A *beat spectrogram* can also be calculated using successive beat spectra [Foote et al., 2001]
Question

– Can we use the beat spectrum for source separation?...

– To be continued...
References

The Similarity Matrix

• Definition
  – Matrix where each point measures the similarity between any two elements of a given sequence.
The Similarity Matrix

• Application
  – Visualize time structure [Foote, 1999]
  – Identify repeating/similar patterns

Similarity matrix.

Region of high self-similarity around 3 s.

Region around 3 s repeating around 8 s, 12 s, and 17 s.

Similarity between the times at 2 s and 10 s.
The Similarity Matrix

• Calculation
  – The similarity matrix $S$ of $X$ is basically the matrix multiplication between transposed $X$ and $X$, after (generally) normalization of the columns of $X$

$$S(j_1, j_2) = \frac{\sum_{k=1}^{n} X(k, j_1)X(k, j_2)}{\sqrt{\sum_{k=1}^{n} X(k, j_1)^2} \sqrt{\sum_{k=1}^{n} X(k, j_2)^2}}$$
The Similarity Matrix

• Calculation
  – Compute the magnitude spectrogram from the audio using the STFT
The Similarity Matrix

• Calculation
  – Normalize the columns of the spectrogram by dividing them by their Euclidean norm

\[ \hat{X}_j(i) = \frac{X_j(i)}{\sqrt{\sum_{k=1}^{n} X_j(k)^2}} \]
The Similarity Matrix

• Calculation
  – Compute the dot product between any two pairs of columns and save them in the similarity matrix

\[
S(j_1, j_2) = \sum_{k=1}^{n} \hat{X}_{j_1}(k) \hat{X}_{j_2}(k)
\]
The Similarity Matrix

• Notes
  – The similarity matrix can also be built from other features (e.g., MFCCs, chromagram, pitch contour)
  – The similarity matrix can also be built using other measures (e.g., Euclidean distance)