Aaron Rourk
Computational Musicology
Fall 2018
Assignment 5

How does beat tracking differ from onset detection?

Beat tracking describes the effort to derive, from an audio signal representing music, a "sequence of beat instants that might correspond" to the timepoints that a human listener would indicate by tapping their feet to *the beat* of the music. The *beats* selected should, "generally correspond to moments in the audio where a beat is indicated, for instance by the onset of a note played by one of the instruments," and should also "reflect a locally-constant inter-beat-interval." This sort of "regular spacing between beat times" defines rhythm in the musical context.

Onset detection, on the other hand, refers to the process of detecting the moments when musical sounds are initiated, such as when a vibraphone is struck, or a voice sings a pitch.

Beat tracking *utilizes* onset detection as a part of its process.

What is onset strength? What is the onset strength envelope?

The onset envelope is calculated by resampling the input signal to 8kHz, the STFT of which is then calculated with a 32ms window and 4ms frame size. This data is then mapped to 40 Mel bands "via a weighted summing" of the values derived from the STFT. "The Mel spectrogram is converted to dB, and the first-order difference-along time is calculated" for each frequency band, derived using a perceptually-weighted auditory frequency scale.

Points in this envelope "correspond to times when there are significant energy onsets across multiple bands in the signal."

What is a global tempo estimate?
Calculating the global tempo estimate involves utilizing the onset strength envelope which is autocorrelated and then windowed according to a

perceptually weighted model (given the fact that humans perceive tempo with a bias to 120bpm).