Beat Tracking

Corey Kereliuk
McGill University

March 8, 2007

Introduction

⇒ Consider beat tracking versus chess

- Human cognition point of view
  - Beat tracking is relatively easy—Any non-trained musician can tap their foot, clap or dance to a beat
  - Conversly, to become skilled at chess is considerably more difficult
- Artificial intelligence point of view:
  - Chess programs have been designed that can compete against the world's best players
  - Thus far no beat tracking system has been designed which can rival the performance of a good musician

⇒ Clearly the task of beat tracking is not trivial

⇒ Robust tracking across multiple genres is a difficult task

Motivation

⇒ Automatic beat tracking is important to the field of MIR

⇒ Discerning the tempo and beat of musical data is useful for a number of different applications:

- Performance analysis
- Audio content analysis
- Symbolic metadata generation
- Musical transcription systems
- Audio segmentation
- Rhythm alignment
- Cut and paste operations in non-linear audio editing application
- etc...

BeatRoot Beat Tracker

⇒ There are many different approaches and incarnations of beat trackers

⇒ Describing them all would be far too time consuming

⇒ Instead I will focus on one state-of-the-art beat tracker

⇒ The system *BeatRoot* was the best performing beat tracker at the 2006 MIREX (Music Information Retrieval Evaluation eXchange) competition

BeatRoot System Schematic

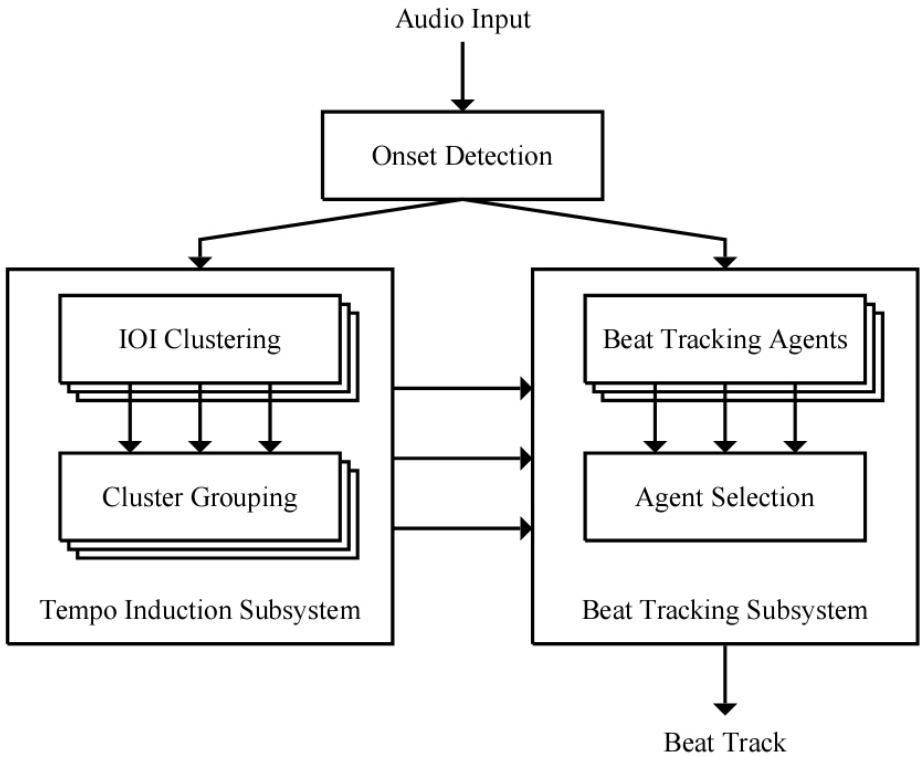

Figure: BeatRoot System Schematic (Dixon 2006a)

Onset Detection

⇒ Onset detection is accomplished by finding peaks in the spectral flux (Dixon 2006b). Spectral flux is defined as the time derivative of the Short Time Fourier Transform $X[n, k]$:

$$\frac{\partial X[n,k]}{\partial n} = SF[n] = \sum_{k=0}^{N} H(|X[n,k]| - |X[n-1,k]|)$$

Where

$$H(x) = \frac{x + |x|}{2}$$

In other words only positive derivatives are saved (half-wave rectification).

Spectral Flux

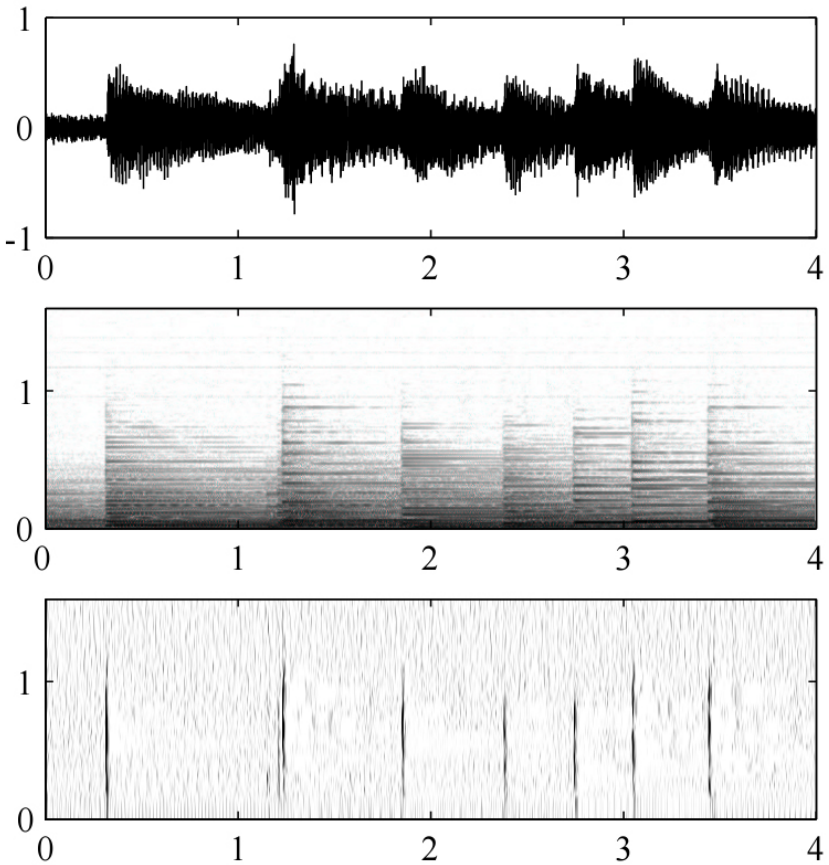

Figure: Spectral Flux (Alonsp 2004)

Tempo Induction

⇒ Onsets are characterized by the peaks of the spectral flux function shown above

⇒ Inter-onset intervals (IOI) are defined as the interval between peaks (not necessarily successive peaks)

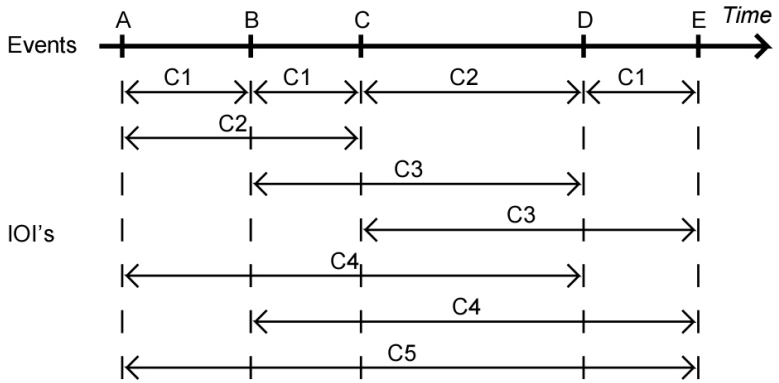

Figure: Tempo Hypotheses (Dixon 2006a)

⇒ These IOIs can be clustered into a set of *tempo hypotheses*

Beat Tracking

⇒ The beat tracking system initializes beat tracking *agents* — one for each tempo hypothesis

⇒ Starting from the first onset each agent is incremented by its IOI interval

⇒ The system then tests to see if the agent has aligned with an onset (one large window and one small window is used to account for variances in the tempo)

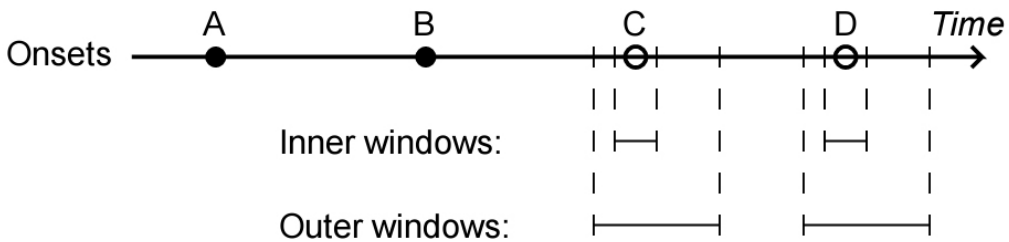

Figure: Window sizes (Dixon 2006a)

Beat Tracking

- ⇒ If an agent aligns with an onset within the range of the small window, it is considered as a beat
- ⇒ If an agent aligns with an onset within the range of the large window, it *may be* a beat
- ⇒ If no beat is found the agent will increment itself. If a beat is found in the next increment, the missing beat will be interpolated. If after a number of frames no onset is found the agent will be terminated (it has lost track of the beat).

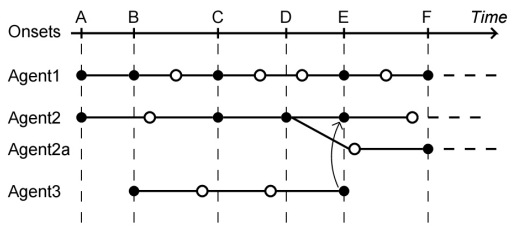

Figure: Agents (Dixon 2006a)

Beat Tracking

⇒ After the agents have traversed the entire audio file, they are subjected to an evaluation process

⇒ The agent that receives the highest score will be returned as the final results

⇒ Scoring is based on several criteria including:

- How evenly beats are spaced in time
- How many predicted beats correspond to actual beats
- The relative strenght of spectral flux at aligned events

Results

- ⇒ Out of the 5 systems submitted to the 2006 MIREX competition BeatRoot had the best performance.
- ⇒ The competition consisted of 140 files from a wide range of musical genres (with hand marked ground truth data).
- ⇒ Additionaly, on a set of 13 complete piano sonatas, and a small set of popular, jazz, and latin songs the system found on average of over 90% of the beats (Dixon 2006b).

Demo of BeatRoot

Go to demo!

Bibliography

Alonso, M., D. Bertrand and R. Gael. 2004. Tempo and beat estimation of musical signals. *Proceedings of the 5th International Conference on Music Information Retrieval.* 158–64

Dixon, S. 2006a. MIREX 2006 Audio beat tracking evaluation: BeatRoot. [cited. Available from http://www.ofai.at/ simon.dixon/beatroot/index.html. Last accessed March 1, 2007]

Dixon, S. 2006b. Onset detection revisited. *Proceedings of th 9th International Conference on Digital Audio Effects.* 133–7