Beat Tracking
Mumt 611 Summary

Corey Kereliuk

March 1, 2007

1 Introduction

Automatic beat tracking is an important pursuit in the field of MIR. Discerning the tempo and beat of musical data is useful for a number of different applications including (but not limited to):

- 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...

The task of beat tracking is relatively simple and intuitive, when examined from the perspect of human cognition. Any non-trained musician can tap their foot, or clap their hands to the beat of a song with reasonable accuracy. Compare this to the task of playing chess, which is a much more complex cognitive task. However, when we examine these tasks from an artifical intelligence perspective, we note that:

- There are chess programs capable of beating some of the worlds best players
- Thus far, no beat tracking system has been produced which can rival the performance of a good musician

Clearly, the task of beat tracking is not a trivial pursuit. Robust performance across multiple genres is a difficult task. There are many different approaches to beat tracking. Describing all of them would take far too much time. Instead I will focus on one state-of-the-art approach and describe its key characteristics. In Simon de Leon's 2005 presentation on beat tracking he described the winning algorithm from the 2004 MIREX (Music Information Retrieval Evaluation eXchange) competition (Alonso 2004). For this presentation I will describe the winning algorithm from the 2006 MIREX competition, a sytem called BeatRoot (Dixon 2006a).

2 BeatRoot Beat Tracking System

The first step in the BeatRoot system is to extract salient information from an audio signal. In terms of beat tracking the onset of a note is one of the most salient features.

2.1 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).

2.2 Tempo Induction

Onset times are clustered into sets of inter-onset intervals (IOIs). An IOI is defined as the time between two onsets. In most music, the IOIs will be an integer multiple of the tempo. Variance in timing/tempo (common for human performers) means that this integer correspondence may not be precise. A clustering algorithm can group IOIs into similar categories representing half notes, quarter notes, etc. Then the information about clusters is combined into a weighted list of *tempo hypothesis* which are passed to the Beat tracking system.

2.3 Beat Tracking

The beat tracking system intializes a number of beat tracking agents—one for each tempo hypothesis generated above. Starting from the beginning of the

piece, each agent is incremented by its IOI. The system then looks for alignment between the agent and an onset (peak in the spectral flux). Two different windows (one large and one small) are used to account for variance in the tempo. If the agent aligns with an onset in the small window, the position is taken to be a beat. If the agent aligns with an onset in the large window the position *may be* a beat. If no alignment is found the agent will continue incrementing. If one of the next beats is aligned, the missing beat will be interpolated. If no beats are found after several increments the agent will be terminated (it has lost track of the beat).Once the agents have gone through a piece, they are subjected to an evaluation process. The agent with the highest score will be the final result. Scoring is accomplished based on a number of criteria including, how evenly beat times are spaced, how many predicted beats correspond to actual events, and the relative strength of the spectral flux at aligned events.

2.4 Results

Out of the 5 systems submitted to the 2006 MIREX competition BeatRoot had the highest score. 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).

3 Conclusion

In this summary I described the basic elements of a beat tracking system. I described several important applications of beat tracking, and illustrated some of the difficulties related to this pursuit. I then described the major functional blocks of the BeatRoot beat tracking system. This system is a state-of-the-art beat tracker capable of a high degree of accuracy across a multiple musical genres.

4 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