Real-Time Percussive Beat Tracking

Andrew Robertson[1], Matthew Davies[2], and Adam Stark[1]

[1] *Centre for Digital Music, Queen Mary University of London*

[2] *INESC TEC, Porto*

Correspondence should be addressed to Andrew Robertson (`andrew.robertson@eecs.qmul.ac.uk`)

ABSTRACT

We present a real-time percussive beat tracking algorithm for synchronisation within live music. A percussive detection function that represents the percussive component of the audio input is created by an efficient method for median filtering. Dynamic programming techniques are used to predict the beat locations and update a cumulative beat function. It is possible to use the percussive component of the spectrogram to create functions which correlate with kick and snare events, thereby generating a prediction of drum pattern events.

1. INTRODUCTION

Beat tracking algorithms aim to replicate the human ability to tap along in time with a song. Whilst this task appears to be simple for humans, it has proved difficult to perform this task automatically, with little recent improvement in results across the standard databases on [1]. On rock and pop material, where the beat has a relatively steady tempo and there are salient percussive events, algorithms have tended to be more successful.

An early system, BTS, by Goto and Muraka [2] exploited the regularity of kick and snare patterns on alternating strong and weak beats. Template matching was used to learn the characteristic bass drum and snare drum signal for current song. Multiple agents interpret information according to their individual hypotheses and the system makes beat predictions according to those made by the strongest agent. Subsequent work tended to look at continuous features such as the onset detection function [3], a form of novelty representation, or the power envelopes of a set of sub-bands [4]. These features act as input to methods such as comb filter resonators [4] [5] and autocorrelation [6] which detect regularity in the signal.

Fitzgerald [7] proposed median filtering as a method to distinguish percussive events, characterised by vertical lines in the spectrogram, from harmonic events, which can be characterised by horizontal lines. Median filtering takes place across frequencies and temporal bins of the spectrogram to create a percussive and harmonic component respectively. Weiner filtering is then used to create a percussive mask and the percussive component of the spectrogram is created by point-wise multiplication of the original spectrogram with the mask. In this paper, we describe a beat tracker utilising this method, which is designed for music with percussive content.

2. PERCUSSIVE BEAT TRACKING

Figure 1 shows an overview of how the percussive component spectrogram is created. To perform real-time median filtering, we exploit the computation that has been carried out at the previous time step [8]. The percussive detection function at each temporal bin is created by summing the percussive spectrogram across all frequency bins. Functions which correlate to kick and snare events can be created by summing selected bands of the percussive spectrogram, using the bins up to 120Hz for kick and those between 200 and 500 Hz for snare.

An estimation of tempo is obtained by using the autocorrelation method described by Davies and Plumbley [6]. To prevent sudden leaps, when tracking tempo we limit changes in tempo to adjacent bins and choose the peak value. To predict beat locations, we make use of the cumulative detection function [9], based on the dynamic programming approach by Ellis [10], which quantifies the extent to which periodic peaks in the percussive signal occur at the current tempo hypothesis. Beat predictions are made iteratively on the eighth note between expected beat times by comparing the predicted beat time with the cumulative detection function. We adapt the parameters which control the systems behaviour according to how well the observed data fits our prediction.

In our implementation, the median filtering takes place

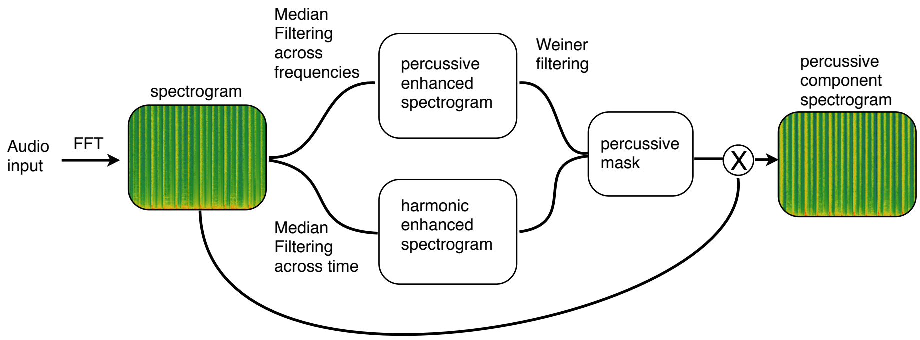

Fig. 1: Overview of the alogrithm.

over 17 frames of the FFT, with a framesize of 2048 samples and a hopsize of 512 samples. This introduces a fixed latency of 8 FFT frames (93 msec) before the percussive detection function is calculated for any given frame. We then look for peaks in this signal which correspond to percussive events played on the beat. By recursively partitioning the audio into successively smaller frames and finding the greatest change in energy between neighbouring frames, we can specify an exact location for each onset. In live performance, this procedure helps to synchronise as close as possible to the beat.

3. ACKNOWLEDGMENTS

The authors would like to thank the Royal Academy of Engineering and the EPSRC for supporting this research. This research was partially funded by the FCT post-doctoral grant (SFRH/BPD/88722/2012).

4. REFERENCES

[1] André Holzapfel, Matthew E. P. Davies, José R. Zapata, and Jo ao Lobato Oliveira, "Selective sampling for beat tracking evaluation," in *IEEE Transactions on Audio, Speech and Language Processing*, 2012, vol. 20, pp. 2539–2548.

[2] Masataka Goto and Yoichi Muraoka, "A beat tracking system for acoustic signals of music," in *Proceedings of the Second ACM International Conference on Multimedia*, 1994, pp. 365–372.

[3] J. P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. Sandler, "A tutorial on onset detection in music signals," *IEEE Transactions on Speech and Audio In Processing*, vol. 13, no. 5, Part 2, pp. 1035–1047, 2005.

[4] Eric D. Scheirer, "Tempo and beat analysis of acoustic musical signals," *The Journal of the Acoustical Society of America*, vol. 103, no. 1, pp. 588–60, 1998.

[5] Anssi P. Klapuri, Antti J. Eronen, and Jaakko T. Astola, "Analysis of the meter of acoustic musical signals," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, no. 1, pp. 342–355, 2006.

[6] M. E. P. Davies and M. D. Plumbley., "Context-dependent beat tracking of musical audio.," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 15, no. 3, pp. 1009–1020, 2007.

[7] Derry Fitzgerald, "Harmonic/percussive separation using median filtering," in *13th Int. Conference on Digital Audio Effects (DAFx-10), Graz, Austria*, 2010.

[8] Andrew Robertson, Matthew E. P. Davies, and Adam M. Stark, "Percussive beat tracking using real-time median filtering," in *Proc. 6th International Workshop on Music Machine Learning, Prague*, 2013, pp. 71–74.

[9] Adam M. Stark, Matthew E. P. Davies, and Mark D. Plumbley, "Real-time beat-synchronous analysis of musical audio," in *Proceedings of the 12th International Conference on Digital Audio Effects (DaFx-09), Italy*, 2009, pp. 299–304.

[10] Daniel P. W. Ellis, "Beat tracking by dynamic programming," *Journal of New Music Research*, vol. 36, no. 1, pp. 51–60, 2007.