A TEMPOGRAM-BASED PROBABILISTIC DYNAMIC MODEL OF BEAT TRACKING FOR AUDIO MUSIC

Fu-Hai Frank Wu
Department of Computer Science
National Tsing Hua University
Hsinchu, Taiwan
frankwu@mirlab.org

ABSTRACT

Automatic beat tracking and tempo estimation are challenging tasks, especially for audio music with time-varying tempo. This paper proposes tempogram-based probabilistic dynamic model to deal with beat tracking with time-varying tempo. In particular, the tempogram is the base to obtain correct beat positions. The probabilistic model to estimate the beat positions include tempogram strength model and relative position model. Combining both model to form the state and transition probabilities. Experimental results demonstrate satisfactory performance for music with significant tempo variations. The beat tracking algorithm could obtain compatible score values for mirex2006 tempo training dataset with 20 excerpts in length of 30 seconds.

Index Terms – Tempogram, Time-varying tempo, Dynamic programming, Probabilistic model

1. INTRODUCTION

Tempo and beat are two essential elements in music. However, automatic beat tracking and tempo estimation are still challenging tasks when the music has time-varying tempos.

Conventional beat tracking schemes [1] rely on certain assumptions about music contents such as stable tempo. Under these assumptions, most approaches of beat tracking are accomplished by two phases. In the first phase, the onset strength of music along time, called *novelty curve*, is estimated to indicate the possible positions of note onsets. In the second phase, the quasi-periodic patterns in novelty curve are analyzed to discover the possible tempo value and the corresponding beat positions. Here, tempo is assumed to be stable throughout the whole piece of music. However, the above-mentioned assumptions do not hold true universally, especially for music of classical and jazz. Music of these genres often has significant tempo variations, making it difficult to detect the periodical patterns. The tempo information is embedded in tempogram. We can then apply dynamic programming (DP) to the tempogram to derive the so-called *tempo curve*, which represents the most likely tempo at each time frame.

There are several important previous studies that attempted to deal with time-varying tempos. Klapuri et al. [6] used the bandwise time-frequency method to obtain accentuation information, then used comb filter resonators and probabilistic models to estimate pulse width and phase of different music meters, including tatum, tactus, and measurement. Davies and Plumbley [7] proposed the use of complex spectral difference onset function to obtain middle level representation. Their algorithm employs two-state switching model, including general state and context-dependent state, to obtain final beat positions. Groshe and Muller [4] used the novelty curve to generate predominant local pulse (PLP) for estimating time-varying tempos.

There are probabilistic framework studies. G. Peeters [2,3] and H. Papadopoulos [3] propose a probabilistic framework for estimation of beat and down-beat simultaneously given information of tempo and meter. The probabilistic model is based on HMM (Hidden Markov Model) which has beat-times and their associated beat-position-inside-a-bar (bpib) as the hidden states. The model is based on non-casual signal observations of the local bar the beat is located in. This provides the work with an inherent local optimization of the probabilities (an adaptation to the local properties of the signal).

In this study, we follow the three-phase framework [8] of beat tracking and attempt to remove the stable-tempo restriction by developing a two-fold DP approach for robust beat tracking with time-varying tempos. To this end, the first DP estimates the time-varying tempo curve from the tempogram (which is obtained from the novelty curve). Then the second DP uses the time-varying tempo curve to identify the optimum beat positions on the novelty curve. In the second DP, the weighting of the tempogram strength and relative position to previous beats are obtained by learning process. The weighting of the tempogram strength is based on spectral distribution and strength. The relative position to previous beats which are in backward or forward direction is compared to tempo at that time.

The remainder of this paper is organized as follows. Section 2 describes the details of the proposed framework with a figure. The phase of beat tracking is the key element to be illustrated.

2. SYSTEM DESCRIPTION

The proposed beat tracking system is shown in Figure 1.

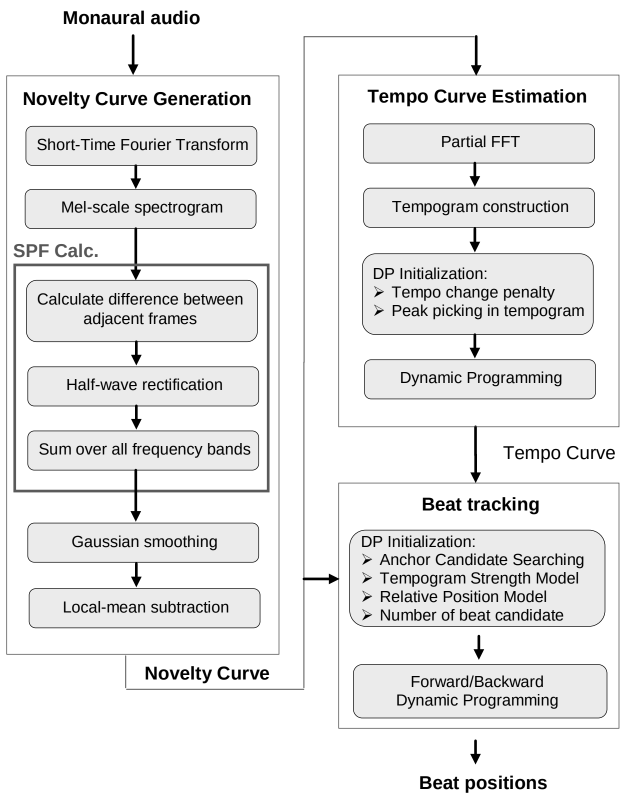

Figure 1. Flowchart of the proposed beat tracking system

The first block computes the novelty curve, while the second block generates the tempogram and estimates the tempo curve from the novelty curve [8]. The third block of beat tracking setup a probabilistic model to count the strength of tempogram and relative position to previous beat position. The detail is shown in the following subsection.

2.1 Beat Tracking

This block utilizes both the tempo curve and the novelty curve to find a sequence of beat positions that fits the tempo curve and the novelty strengths as much as possible. To achieve this task, we apply another DP-based method in a probabilistic framework (just like Viterbi search in speech recognition) to perform forward and backward beat tracking, starting from the anchor beat position (the position of the most prominent peak) of the novelty curve. We proposed the tempogram strength model and relative position model to form the state and transition probabilities.

Here we use Figure 2 to explain the probability-based DP method for beat position identification. First of all, we find the maximum of the novelty curve as the first beat position, which is referred to as the anchor candidate. Starting from the anchor candidate, we search on both sides, one side at a time, to obtain all beat positions.

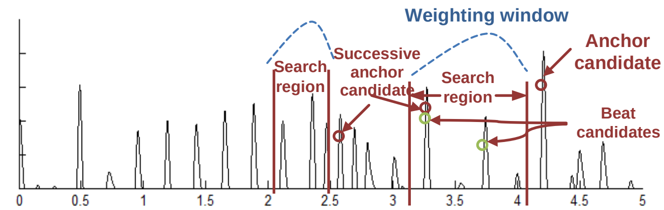

Figure 2. Backward beat search with N =2

Once the state and transition probabilities are defined, we can apply DP just like Viterbi search for the optimum beat positions. The search is performed twice for both forward and backward directions from the anchor candidate, and the results of them are merged to obtain the complete beat positions. In our experiment, we set N to 2. Figure 2 shows a typical result with θ = 0.01 and N=2.

3. REFERENCES

[1] D.P.W. Ellis, "Beat Tracking by Dynamic Programming," *Journal of New Music Research*, Vol. 36(1), 51–60, 2007.

[2] G. Peeters, "Beat-marker Location Using a Probabilistic Framework and Linear Discriminant Analysis," in *Proc. DAFX*, Como, Italy, 2009.

[3] G. Peeters and H. Papadopoulos, "Simultaneous Beat and Downbeat-tracking Using a Probabilistic Framework: Theory and Large-scale Evaluation," submitted to *IEEE. Trans. on Audio, Speech and Language Processing*, Vol. 19, No. 6, August 2011.

[4] Peter Grosche, Meinard M¨uller, "Computing Predominant Local Periodicity Information in Music Recordings" *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2009

[5] *Adam M. Stark, Matthew E. P. Davies, Mark D. Plumbley*, "Real-Time Beat-Synchronous Analysis of Musical Audio" *in Proc. the 12th Int. Conference on Digital Audio Effects (DAFx-09)*

[6] Anssi P. Klapuri, Antti J. Eronen, Jaakko T. Astola, "Analysis of the Meter of Acoustic Musical Signals" *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 14, No. 1, January 2006

[7] Matthew E. P. Davies, Mark D. Plumbley, "Context-Dependent Beat Tracking of Musical Audio" *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 15, No. 3, March 2007

[8] Fu-Hai Frank Wu, Tsung-Chi Lee, Jyh-Shing Roger Jang, Kaichun K. Chang, Chun Hung Lu, Wen Nan Wang, "A Two-Fold Dynamic Programming Approach to Beat Tracking For Audio Music with Time-Varying Tempo" in *Proc. ISMIR*, Florida, USA, 2011, Accepted.