AN ADAPTIVE MECHANISM FOR BEAT TRACKER

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

Jyh-Shing Roger Jang
Department of Computer Science
National Taiwan University
Taipei, Taiwan
jang@mirlab.org

ABSTRACT

In the current state-of-art beat tracker, it is common that a specific algorithm do the best on particular kind of audio. For example, the algorithm based on two-fold dynamic programing have sequential 1st place on time-varying-tempo MAZ dataset of MIREX audio beat tracking. This study proposed a tempogram-sensing-vector (TSV) method to be adaptive to time-varying-tempo and stable-tempo excerpts.

***Index Terms* –** Time-varying Tempo, Stable Tempo, Tempogram

1. INTRODUCTION

Rhythmic information is the essential element in music. The prominent features of rhythm are beat position and tempo which comprise the basic characteristic of music. Although the sense of beat sometimes is obvious for human being, the exact estimation is still challenging task when especially the music has time-varying tempo.

Conventional beat tracking schemes [1] handle certain music contents with stable tempo well. Under the related stable-tempo assumptions, most approaches of beat tracking are accomplished by two phases. In the first phase, the onset detection of music along time, called onset detection function, onset strength and *novelty curve*, is obtained to indicate the possible positions of note onsets. In the following phase, the quasi-periodic patterns in novelty curve are analyzed to discover the possible tempo value and the corresponding beat positions. Usually in the deduction process, 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 unreasonable to make the assumption of stable tempo. In our work, we break the assumption of stable tempo. Therefore, we generate the tempogram from the novelty curve, which the tempo information is embedded. Then we apply dynamic programming (DP) to the tempogram to derive the so-called *tempo curve*, which represents the most likely tempo at each time frame which is time-varying.

There are several important previous studies that attempted to deal with time-varying meters. Klapuri et al. [2] 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 [3] 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. G. Peeters and H. Papadopoulos [5] propose a probabilistic framework for estimation of beat and downbeat 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 which 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 explore the method to take advantage of the specified algorithms [6] good at time-varying-tempo excerpts and stable-tempo excerpts individually. The frame based or beat-synchronous analysis for audio [7] is suitable for the study. At the current stage, we focus on the frame-based features because estimated beat sequence is generated with prediction. The following section describes the block diagram of the system and illustrates tempogram-sensing vectors (TSV).

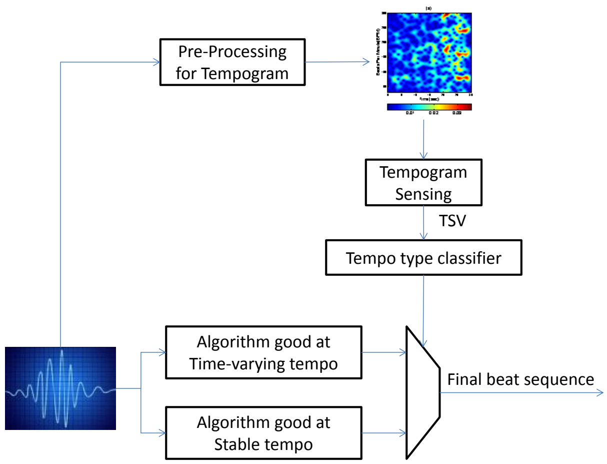

Figure 1 Block diagram of the adaptive mechanism

2. THE ADAPTIVE MECHANISM

The adaptive mechanism to time-varying and stable excerpts is shown in Figure 1. The audio is pre-processed to obtain tempogram; then, the tempogram is analyzed to produce TSV, which is input feature of a classifier. Finally the beat sequence is selected by the results of the classifier from the time-varying beat sequence and stable beat sequence.

2.1 Tempogram Sensing Vector (TSV)

The TSV contains the features computed from tempogram, where the features are tempogram mean (μ_T), tempogram standard deviation (σ_T), tempogram coefficient of variation (cv_T), tempogram skewness (γ_T) and tempogram kurtosis(κ_T), which are calculated from the whole and strips of tempograms and the delta of the statistics to generate low-level features.

3. REFERENCES

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

[2] 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

[3] 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

[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] 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

[6] 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

[7] *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)*