EXPLORE THE FACTORS TO IMPROVE ACCURACY OF DYNAMIC-PROGRAMMING-BASED BEAT TRACKER

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

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

ABSTRACT

For our dynamic programming (DP) based [2] beat tracker, we try to identify the factors to improve accuracy. For the specified algorithm, it has merit to handle time-varying-tempo excerpts. Contrast to the effectiveness, the algorithm is inappropriate for musical audio with stable tempo. One of the reason to cause poor accuracy is due to inaccurate tempo estimation. This study adopts the research result of our tempo estimation to improve beat tracking accuracy. Besides, we also attempt to approve features, the modified tempo shape vector (TSV), to discriminate time-varying-tempo and stable-tempo excerpts in order to combine a universal beat tracker for those excerpts.

1. INTRODUCTION

The purpose to join the evaluation of MIREX 2014 audio beat tracking (abt) is to conduct self-improvement, but not compete with others. Although it is more challenge to compete with peer algorithms, to confirm major improvement has the similar importance. Comparing with the results of MRIREX 2014 and 2013 abt, we could find we meet the target set up for the MIREX 2014 evaluation.

According to the results shown in table, we have improve in MAZ (time-varying tempo) and SMC (hard to exercise beat tracking) dataset for universal beat tracker. We have three submissions for MIREX 2013 and MIREX 2014, respectively. The submissions in 2013 and 2014 are all derived from our previous work [3] suitable to time-varying-tempo excerpts. We devise a stable-tempo beat tracker with multi-paths and a selector to decide only one beat sequence. This kind of beat tracker are targeted for stable-tempo excerpts and are close to state-of-art beat tracking algorithms. The last category, named as universal beat tracker, contains the arbitration mechanism to select the best between time-varying-tempo beat tracker and stable-tempo beat tracker for the specified excerpt.

In this evaluation, we have approved the usefulness that the adoption of tempo-pair value from the tempo estimation algorithm. Actually, we have tested it in our experiment on "Hainsworth" dataset with F-Measure around 68%, to our best knowledge it is close to the top score in the repository of beat trackers of some senior peer researcher in this area. The implementation is as follows: the two tempi are the guidance to derive tempo-curve, which are the reference of the cost function in dynamic programming beat tracker; at the same time, we grow paths when the tempo has abruptly change, so the paths are growing exponentially when tempo changes; the selector is simply voter based on similarity of beat sequence and cost of DP.

The evaluation also demonstrates the success of discriminating excerpts of time-varying-tempo and stable-tempo by TSV. While the improvement is not much as Table 1 shown, we could observe that universal beat tracker has better accuracy than other two algorithms for SMC dataset. Because there is big score difference between time-varying-tempo beat tracker and stable-tempo beat tracker for dataset MCK and MAZ, the universal beat tracker is only superior to that of low-score beat tracker.

Next section describes the major factors of the universal beat tracker. More specifically, we explain the notion of tempo-guided tempo curve and the TSV feature in following section.

Table 1 MIREX evaluation result of audio beat tracking for years of 2013 and 2014

Beat tracker type	MIREX 2013				MIREX 2014			
	Submission code	F-MEASURE			Submission code	F-MEASURE		
		DAV(SMC)	MAZ	MCK		SMC	MAZ	MCK
Time-varying-tempo	FW1	35.1852	66.9948	44.0698	FW3	35.4572	67.1690	44.2559
Stable-tempo	FW2	31.3006	41.1565	50.1784	FW4	36.9533	53.2305	51.2939
Universal	FW4	30.9980	42.2357	49.8823	FW5	37.5585	54.4881	50.9856

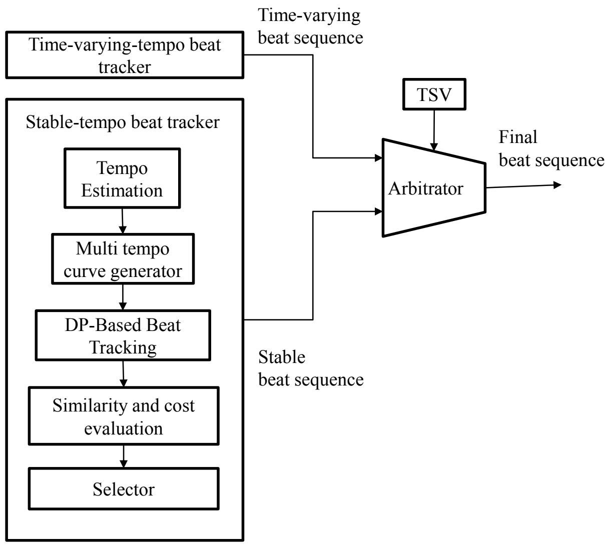

Figure 1. Flowchart of proposed method

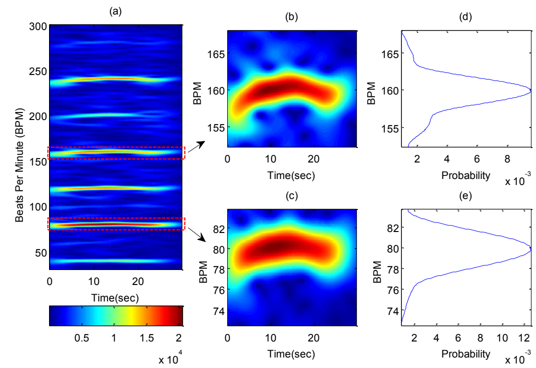

Figure 2 Typical example of LTP derivation: (a) Tempogram; (b) tempogram strip around tempo2 (c) tempogram strip around tempo1 (d) LTP of tempo2 tempogram strip (e) LTP of tempo1 tempogram strip.

2. IMPORTANT FACTORS TO APPROACH UNIVERSAL BEAT TRACKER

In order to build a beat tracker for public usage, the beat tracker should be suitable for all kinds of music. For our algorithm, the stable-tempo-tracking ability seems to have some room to improve. So we provide the tempo-guided multipath beat tracker. Another important factor is how to combine the presumed best beat tracking algorithms, the first step is to discriminate which beat tracker is the best fit to the specified type of tempo. TSV is borrowed from our previous work to approach this. In short, **Figure 1** is shown to illustrate the role of these two major factors. The audio is processed by stable-tempo beat tracker and time-varying-tempo beat tracker simultaneously; then, the tempogram is analyzed to produce TSV, which is input feature of a classifier. Finally, the last beat sequence is selected by the results of the arbitrator, which is the classifier to discriminate the time-varying beat sequence and stable beat sequence.

2.1 Stable-Tempo Beat Tracker with Tempo-Guided Multipath

From theoretical and empirical perspective, more accurate tempo could improve our DP-based beat tracker. So we put constrains on the tempogram with the weighting function centered at the estimated tempo pair to lead the derivation of tempo curve. Then the beats could be likely to constitute the tempi as the result of tempo estimation. After we acquire the tempo curve, we create more reference tempo curves. Therefore for each tempo curve, we create beat sequence for it. Finally, the decision of final beat sequence depend on the factors: the similarity of those beat sequences, the cost of beat sequence, and the frequency of beat sequence.

2.2 Tempogram Shape 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). The statistical quantities are calculated from the whole and stripes of tempogram to generate low-level features.

In term of the statistics of tempogram stripe, the long-term-periodicity (LTP), as Figure 2 shown, is regarded as probability mass function and the statistical quantities are calculated as the components of the TSV. The detail derivation of the vector could be referenced in the work [4].

3. REFERENCES

[1] Downie, J. Stephen, et al, “The music information retrieval evaluation exchange: Some observations and insights,” *Advances in music information retrieval*. Springer Berlin Heidelberg, 2010. 93-115.

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

[3] 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

[4] Fu-Hai Frank Wu, Jyh-Shing Roger Jang, “A Supervised Learning Method into Tempo Estimation of Musical Audio”, Control and Automation (MED), 2014 Mediterranean conference on, IEEE published.