On-line Music Beat Tracking with Kalman Filtering and Probability Data Association (KF-PDA)

Yu Shiu and C.-C. Jay Kuo

Ming Hsieh Department of Electrical Engineering and Signal and Image Processing Institute
University of Southern California, Los Angeles, CA 90089, USA
E-mails: yshiu@usc.edu, cckuo@sipi.usc.edu

Abstract- **An on-line music beat tracking algorithm with Kalman filtering (KF) and probability data association (PDA) is proposed. The KF-PDA method not only estimates the music tempo and beat pulse positions in real-time but also provides a temporal metrical unit for human-perception-like processing. The superior performance of the proposed KF-PDA beat tracking system is demonstrated using the MIREX beat tracking competition's practice data set.**

I. INTRODUCTION

When listening to music, most people can follow it by foot-tapping, head-shaking or hand-clapping along with beats. Usually, after listening to several seconds of music, one can catch the speed of music and start to follow it by tapping with beats. However, it is challenging for computers to comprehend tempo and melody of digital music well. Automatic music beat tracking is one of the important tasks in low level music signal analysis.

Automatic music beat tracking finds numerous applications in consumer electronics, *e.g.* automatic music accompaniment and music-driven graphic animation. Human performers can play solo along with computer's accompanying music for the former while graphic characters can have different movements based on music's beat and rhythm for the latter. A good beat tracking algorithm enhances the interaction between the music and animation.

Although music beat tracking techniques have been widely studied in the past, only a few of them apply to real-time (or causal) audio processing, *e.g.*, [1], [2], [3]. Scheirer [1] used a comb filter to estimate the tempo and the beat location with an open-loop approach. That is, new estimates do not take prediction errors made before into account. The beat tracking methods proposed in [2] and [3] adopt the particle filtering technique. The particle filter is more general than the Kalman filter since it makes no assumption on the linearity of the tracking system and the Gaussianarity of underlying signals. However, its complexity is significantly higher and it does not address the problem of incorrect measurements, which will be discussed in this paper.

Beat tracking with Kalman filtering was studied in [5] and [6]. However, the tracking algorithm is not robust for some special type of instruments. In this work, we proposed an improved Kalman filtering technique by incorporating probability data association (PDA), which works well for a much wider range of music types. The superior performance of the proposed KF-PDA method is demonstrated using the MIREX beat tracking competition's practice data set.

II. PROPOSED KF-PDA BEAT TRACKING ALGORITHM

The proposed on-line music beat tracking system is depicted in Fig. 1. As shown in the figure, we first perform onset detection and estimate its period from the input musical signal. Then, the proposed beat tracking algorithm work on these data for beat tracking in the time domain.

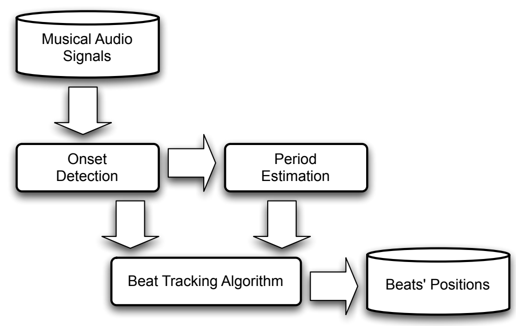

Fig. 1. An overview of the proposed beat tracking system.

A. Onset Detection and Period Estimation

The onset indicates certain magnitude and/or phase changes of musical signals along time. It can be classified into two types: 1) instantaneous pulse-like changes caused by percussion instruments; and 2) changes of music pitches and/or harmonics due to the arrival of new notes.

There are many onset detection methods developed. Here, we consider a simple scheme; namely, to measure the the spectral content change between two adjacent shifting windows with 50% overlap of 20-msec long. Mathematically, we compute the following distance measure:

$$d_c^2 = \sum_{m=1}^{L} (c_m - c'_m)^2, \qquad (1)$$

where c_m and c'_m are mel-scale frequency cepstral coefficients (MFCC) of two consecutive frames, and compare it with a threshold value. If it is above the selected threshold value, we claim that an onset is detected. The output of the onset detection module, denoted by $x(i)$, serves as the input to the period estimation module and the beat tracking module as shown in Fig. .

The tempo, which is the inverse of the period of musical onset signals, is assumed to be nearly constant when the beat tracking algorithm is applied. The auto-correlation function (ACF) is computed to estimate the period of musical onsets and thus tempo.

B. Kalman Filtering

We first define the state vector as [5],[6]

$$x_k = [\tau_k, \Delta_k]^T, \tag{2}$$

where τ_k is the temporal location of the kth beat and the beat interval

$$\Delta_k = \tau_{k+1} - \tau_k. \tag{3}$$

The time of the next beat and the next beat interval can be predicted by

$$\begin{aligned} \tau_{k+1} &= \tau_k + \Delta_k, & (4) \\ \Delta_{k+1} &= \Delta_k. & (5) \end{aligned}$$

Then, we can set up the following dynamic system:

$$\begin{aligned} x_{k+1} &= \phi(k+1|k)x_k + \mu_k = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} x_k + \mu_k, & (6) \\ y_k &= M(k)x_k + \upsilon_k = \begin{pmatrix} 1 & 0 \end{pmatrix} x_k + \upsilon_k, & (7) \end{aligned}$$

where y_k is the measurement (or observation) at time k. Eq. (6) gives the system dynamic model while Eq. (7) is the observation equation. With the above formulation and proper initializations, one can conduct the standard Kalman filter accordingly.

C. Probability Data Association (PDA)

As shown in Eq. (2) and (7), y_k is the beat position τ_k. When applying the Kalman filter to musical beat tracking, measurement y_k is needed at all k to adaptively adjust the estimate of state vector x_k. However, it is difficult to estimate τ_k accurately on the fly. Conventionally, one may use a fixed window around $\hat{\tau}_k$, which is the estimate of τ_k, and choose the local maximum of music onsets as the measurement. Mathematically, we have

$$y_k = \underset{|m-\hat{\tau}_k|<w}{\arg\max}\ d(m), \tag{8}$$

where $d(m)$ is the onset signal obtained from the onset detection module. The error associated with Eq. (8) is often high, which deteriorates the tracking performance significantly.

Here, a mechanism called the probability data association (PDA) [7] is used to enhance the tracking ability of the standard Kalman filter. PDA does not select a single measurement as done in Eq. (8). Instead, it uses a weighted average of several candidates to estimate the current state vector $\hat{x}(k|k)$ as

$$\hat{x}(k|k) = E(x(k)|Z^k) = \sum_{i=0}^{m_k} \hat{x}_i(k|k)\beta_i(k). \tag{9}$$

A probabilistic method is used in PDA to calculate the weight $\beta_i(k)$ for each candidate $\hat{x}_i(k|k)$.

III. EXPERIMENTAL RESULTS AND DISCUSSION

MIREX 2006 beat tracking competition practice data are used in the experiments. It consists of twenty 30-sec music clips of diverse genres. The first segement of 5 seconds is used as the training data in estimating the period and the initial beat. The remaining segment of 25 seconds is used for beat tracking performance evaluation.

To measure the beat tracking performance, the ratio of the longest continuous correctly tracked segment over the total length(*i.e.* 25 sec) for each music clip is applied [4]. Their average over all the music clips is shown in Table 1. We compare the performance of Kalman filtering with the local maximum as given in Eq. (8) and Kalman filtering with PDA given in Eq. (9). Kalman filter with LM and PDA can achieve performances of 86.28% and 88.94%, respectively.

Table 1. Bit tracking performance comparison with local maximum selection (LM) and probability data association (PDA).

	KF-LM	KF-PDA
Correct tracking ratio	86.28%	88.94%

The performance comparison with respect to each individual test clip is shown in a scatter plot as shown in Fig. 2, where each square indicates the result of a music clip and its x-coordinate and y-coordinate represent the performance of KF-LM and KF-PDA, respectively. Most squares lie above the 45 degree line, which means that KF-PDA performs better than KF-LM in most test music clips.

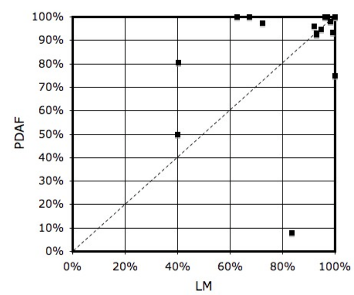

Fig. 2. Performance comparison with the scatter plot, where the x- and the y-coordinates of a sequare denote the performance of KF-LM and KF-PDA for the same music clip, respectively.

IV. REFERENCES

[1] E. Scheirer, "Tempo and beat analysis of acoustic musical signals", *Journal of Acoustic Society America*, vol. 103, pp. 588-601, 1998.

[2] W. A. Sethares, R. D. Morris and J. C. Sethares, "Beat tracking of musical performances using low-level audio features", *IEEE Trans. on Speech and Audio Proccessing*, vol. 13, No. 2, pp. 275-285, 2005.

[3] S. Hainsworth and M. Macleod, "Beat tracking with particle filtering algorithms", *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 91-94, 2003.

[4] A. P. Klapuri and A.J. Eronen and J.T. Astola, "Analysis of the Meter of Acoustic Musical Signals", *IEEE Trans. on Speech and Audio Proccessing*, vol. 14, No. 1, pp. 342-355, 2006.

[5] A. T. Cemgil and B. Kappen and P. Desain and H. Honing, "On tempo tracking: tempogram representation and Kalman filtering", *Journal of New Music Research*, 2001.

[6] Yu Shiu and C.-C. Jay Kuo, "A modified Kalman filtering approach to on-line musical beat tracking," *IEEE International Conference on Acoustics, Speech and Signal Processing*, Honolulu, Hawaii, USA, April 16-20, 2007.

[7] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*, Academic Press, 1988.