MIREX 2015: MULTI FEATURE BEAT TRACKER (INF AND REG) + SOURCE SEPARATION

José R. Zapata
Universidad Pontificia Bolivariana, Medellin, Colombia
joser.zapata@upb.edu.co

ABSTRACT

The Multi-feature Beat tracker [16] 2015 uses six different onset detection functions to estimate the beats of a musical audio signal, this signal is processed with the REPET [17] algorithm in order to extract the repeating parts of the audio, the beat tracker algorithm proccess equally the original audio andd the processed audio signalusing only one beat tracker algorithm, finally the beat tracker output is selected using a committee technique presented in previous works. This is a new version of the algorithm ZDG1 and ZD2, which uses five onset detection function, submitted to MIREX 2012 audio beat tracking task.

1. INTRODUCTION

Based in the beat tracking selection from a committee of state of the art algorithms presented in [16] [15] [11], This Beat tracker uses six different onset detection functions as input signal to the Degara Beat Tracker [6]. Each output is considered as a committee member and in each case the beat tracker output chosen to represent the committee is selected automatically as the one which most agrees with the remainder of the committee (Maximum Mutual Agreement, MaxMA).

2. BEAT TRACKING SYSTEM

The Beat tracking estimation is computed six times, each time with a different onset detection function. First over the original signal and finally over the processed audio signal with the REPET algorithm, giving us 12 different beat estimation resutls. the output is selected using the Maximum Mutual Agreement method. The original Beat tracker was presented in [16] and implemented in the essentia framework, http://essentia.upf.edu/ This implementation is a variation using the REPET algorithm as preproccess step.

2.1 Repeating Pattern Extraction Technique (REPET):

REPET [1] is a method for separating the repeating background from the non-repeating foreground in a excerpt audio mixture. The approach assumes that musical pieces are often characterized by an underlying repeating structure over which varying elements are superimposed. The system identifies the repeating elements in the audio, compares them to repeating models derived from them, and extracts the repeating patterns via time-frequency masking. A detailed description of the algorithm is in [17].

2.2 Onset Detection Function

The onset detections used are:

- Complex Spectral Difference [7]
- Energy Flux [12]
- Harmonic Function [9]
- Sub Bands weight [5]
- Phase Slope Function [13]
- Spectral Flux Log Filtered [1]

2.3 Period Estimation

The beat tracking system estimates the beat period and phases independently. To extract the sequence of periods from the beat period salience observation signal. The beat period is estimated by passing the autocorrelation function of the onset detection function through a shift invariant comb filterbank matrix. Then the information of the beat period is used to recover the beat alignment by passing the onset detection function through a comb filter matrix. The system assumes the beat period like a slowly varying process and the transition probabilities are modeled using a Gaussian distribution of fixed standard deviation. For a complete description of the beat period estimation method see [2, 6].

2.4 Beat Tracking

The Degara Beat tracking probabilistic model [6] takes as input parameters the phase observation signal and the beat period estimation, returning the set of beat time estimates.

[1] music.cs.northwestern.edu

2.5 Output Selection with Maximum Mutual Agreement (MaxMA)

The Mean Mutual Agreement (MMA) follows the Query by Committee concept [14] which selects the most informative set of samples from a database based on the mutual (dis-)agreement between a designated committee of learners. Given a committee of beat trackers, the low MMA between their estimated beat sequences (see Figure 1) on a music database was shown to indicate difficult samples for beat tracking, by being strongly correlated with low performance against the ground truth of this data [11].

As depicted in Figure 1 and proposed in [11], the MMA of a sample is computed by using the beat estimations of N beat trackers on a musical piece, measuring the mutual agreement (MA) among their estimated beat sequences, and retrieving the mean of all $N(N-1)/2$ mutual agreements:

$$MMA = \frac{1}{N(N-1)/2} * \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} MA_{i,j}. \quad (1)$$

The Maximum Mutual Agreement (MaxMA) refer to the MA of the beat sequence that highly agrees with the others:

$$MaxMA = max_i \left(\sum_{j=1, j \neq i}^{N} MA_{i,j} \right), i = 1, \ldots, N. \quad (2)$$

In order to measure the MA between each pair of estimated beat sequences a beat tracking evaluation criteria was selected. A combination of Information Gain measure [4] for Multi feature beat tracker information gain (JZ1) and Regularity Function [10] for Multi feature beat tracker information gain (JZ2). The output selection is done using the Maximum mutual agreement presented in [15] [11].

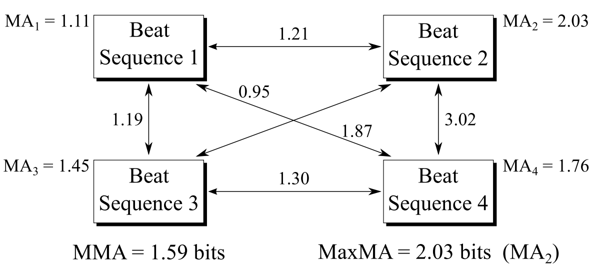

Figure 1. Example calculation of the MMA and MaxMA for a musical piece with the beat sequences estimated from a committee of four beat trackers.

3. RESULTS

Year	Name	AMLc	AMLt	ODF	Method
2015	JZ2	**53.1466**	**68.2513**	6	Reg + Repet
2015	JZ1	50.9666	67.2172	6	Inf + Repet
2014	JZ2	52.8680	67.0925	7	Reg
2014	JZ1	49.8374	65.5170	7	Inf
2013	ZDG2	52.0420	66.8829	6	Reg
2013	ZDBG1	49.8669	65.9591	6	Inf
2013	ZDG1	49.5117	65.3080	5	Inf
2012	ZDG2	51,759	66,6591	5	Reg
2012	ZDG1	49,4494	65,0939	5	Inf

4. ACKNOWLEDGEMENTS

We would like to thank all the authors (N.Degara, D.Ellis, A. Holzapfel, S. Böck, F. Krebs, M. Schedl and Z. Rafii) of the available matlab code for their support. Universidad Pontificia Bolivariana (Colombia) and Colciencias.

5. REFERENCES

[1] S. Böck, F. Krebs, and M. Schedl, "Evaluating the online capabilities of onset detection methods," *13th International Society for Music Information Retrieval Conference (ISMIR 2012)*, Porto, pp. 49–54, 2012.

[2] 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.

[3] M. E. P. Davies, N. Degara, and M. D. Plumbley, "Evaluation methods for musical audio beat tracking algorithms," Queen Mary University of London, Centre for Digital Music, Tech. Rep. C4DM-TR-09-06, 2009.

[4] M. E. P. Davies, N. Degara and M. D. Plumbley, "Measuring the performance of beat tracking algorithms using a beat error histogram," *IEEE Signal Processing Letters*, vol. 18, no. 3, pp. 157–160, 2011.

[5] M. E. P. Davies, M. D. Plumbley and D. Eck, "Towards a Musical Beat Emphasis Function," In *Proceedings of the 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2009)*, New Paltz, NY, October 18-21, pp. 61-64, 2009

[6] N. Degara, E. Argones, A. Pena, S. Torres-Guijarro, M. E. P. Davies and M. D. Plumbley, "Reliability-Informed Beat Tracking of Musical Signals," *IEEE Transactions on Audio, Speech and Language Processing,* Vol. 20, pp. 290–301, 2012.

[7] S. Dixon, "Onset detection revisited," in *6th International Conference on Digital Audio Effects (DAFx-06)*, Montreal, Canada, Sept. 1820, pp. 133137,2006.

[8] Ellis, D. "Beat Tracking by Dynamic Programming". In *Journal of New Music Research*, vol. 36, no. 1, pp. 51–60, 2007.

[9] S. Hainsworth and M. Macleod. “Onset detection in musical audio signals”. In *Proceedings of International Computer Music Conference (ICMC)*, pp. 136–166, Singapore. 2003.

[10] M. Marchini and H. Purwins. “Unsupervised Analysis and Generation of Audio Percussion Sequences,” *Lecture Notes in Computer Science*, vol. 6684, pp. 205–218, Springer Berlin 2011.

[11] A. Holzapfel, M. E. P. Davies, J.R. Zapata, J.L. Oliveira and F. Gouyon, “Selective sampling for beat tracking evaluation,” *IEEE Transactions on Audio, Speech and Language Processing,* Vol. 20, No. 9, P. 2539–2548, Nov 2012.

[12] J. Laroche, “Efficient tempo and beat tracking in audio recordings,” *Journal of the Audio Engineering Society,* vol. 51, no. 4, pp. 226–233, 2003.

[13] A. Holzapfel and Y. Stylianou, “Beat tracking using group delay based onset detection,” in *Proc. of ISMIR - International Conference on Music Information Retrieval*, Philadelphia, EEUU, pp. 653658, 2008.

[14] H. S. Seung, M. Opper and H. Sompolinsky “Query by committee,” in *Proc. of the 5th Annual Workshop on Computational learning theory,* pp. 287–294, 1992.

[15] J.R. Zapata, A. Holzapfel, M. E. P. Davies, J.L. Oliveira and F. Gouyon, “Assigning a Confidence Threshold on Automatic Beat Annotation in Large Datasets,” In *Proceedings of 13th International Society for Music Information Retrieval Conference (ISMIR 2012),* Porto. 2012.

[16] Zapata, Jose R and Davies, Matthew E P and Gómez, Emilia, “Multi-Feature Beat Tracking,” In *Audio, Speech, and Language Processing, IEEE/ACM Transactions on,* Vol. 22, No. 4, P. 816–825, 2014.

[17] Rafii, Z. and Pardo, B., “REpeating Pattern Extraction Technique (REPET): A Simple Method for Music/Voice Separation,” In *Audio, Speech, and Language Processing, IEEE Transactions on,* Vol. 21, No. 1, P. 71–82, 2013.