MIREX 2012: MULTI FEATURE BEAT TRACKER (ZDG1 AND ZDG2)

José R. Zapata[1], Matthew E.P. Davies[2] and Emilia Gómez[1]
[1]Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain
[2]Sound and Music Computing Group, INESC TEC, Porto, Portugal
joser.zapata@upf.edu, mdavies@inescporto.pt, emilia.gomez@upf.edu

ABSTRACT

The Multi-feature Beat tracker uses 5 different onsets detection function to estimates the beats of a musical audio signal using only one beat tracker algorithm, finally the beat tracker output is selected using a committee technique presented in previous works. The algorithm ZDG2 get the higher value in five of the ten measures in the Mckinney Dataset in 2012 and the higher AMLt and AMLc values in all the years in the Beat tracking task in the same dataset.

1. INTRODUCTION

Based in the beat tracking selection from a committee of state of the art algorithms presented in [12] [9], we use five different onset detection functions as input signal to the Degara Beat Tracker [5], 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 five times, each time with a different onset detection function. the output is selected using the Maximum Mutual Agreement.

2.1 Onset Detection Function

- Complex Spectral Difference [6]
- Mel Auditory Features [7]
- Spectral Flux [6]
- Sub Bands harmonic [4]
- Sub Bands weight [8]

2.2 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 system uses an off-line version of the hidden Markov model method presented in [1] that assumes the beat period like a slowly varying process and the transition probabilities are modeled using a Gaussian distribution of fixed standard deviation.

2.3 Beat Tracking

The Degara Beat tracking probabilistic model [5] takes as input parameters the phase observation signal and the beat period estimation, returning the set of beat time estimates.

2.4 Output Selection with Maximum Mutual Agreement (MaxMA)

The Mean Mutual Agreement (MMA) follows the Query by Committee concept [11] 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 groundtruth of these data [9].

As depicted in Figure 1 and proposed in [9], 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:

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

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 [3] and AMLt [2] for (ZDG1) and Regularity Function [10] for (ZDG2).

Algorithm	F-Measure	Cemgil	Goto	P-score	CMLc	CMLt	AMLc	AMLt	D (bits)	Dg (bits)
ZDG2	53.3908	40.6227	**22.4347**	58.2377	25.0103	33.3751	**51.759**	**66.6591**	**1.8055**	**0.3133**
GP3	50.3246	37.2708	21.1824	56.5626	23.9642	33.691	49.272	66.4488	1.7829	0.2522
ZDG1	51.6075	38.8155	20.9114	57.3818	23.7155	32.3427	49.4494	65.0939	1.7953	0.2639
GP2	50.0944	37.0009	20.2158	56.1765	23.2642	32.3007	48.5807	64.8903	1.783	0.2411
GKC2	50.1021	37.8267	19.0269	55.1619	25.8119	32.9415	51.0431	64.2324	1.686	0.2729
ODGR1	50.5011	38.2091	17.786	55.5026	21.5578	29.9927	49.3793	64.1496	1.6592	0.2568
FK1	56.7275	42.6967	21.389	61.1646	22.2548	35.076	41.4778	63.2658	1.6594	0.3127
ODGR2	50.3833	38.1724	18.7814	55.4386	22.3619	30.3893	47.036	62.7007	1.6095	0.2668
KB1	53.5053	39.584	17.4612	57.712	17.5112	29.9126	35.8856	60.2136	1.6216	0.2286
ODGR3	49.7496	37.718	16.0461	55.0283	21.8343	29.7376	44.2349	59.7396	1.5388	0.2585
FW4	52.1262	39.5043	21.6458	57.6836	23.684	34.5203	42.4399	59.1434	1,643	0.2586
KFRO1	51.1306	38.9734	20.6787	56.0343	25.0057	32.0236	47.0872	58.8376	1.6635	0.2918
ODGR4	47.804	36.1936	14.9805	53.751	19.9762	28.3435	41.3848	58.1546	1.4727	0.2278
SB6	52.946	40.2607	18.8169	56.8074	20.3858	29.3413	40.8115	57.152	1.6018	0.2533
FW3	51.9314	39.226	20.0735	57.8761	22.5415	34.0889	39.1905	56.9639	1,600	0.2486
SB3	52.6909	39.9243	19.7332	57.0825	20.8264	29.9532	37.4468	53.6397	1.5724	0.2044
GP4	49.6137	36.7119	12,6218	55.6275	19.5719	30.3776	35.1656	52.4657	1.5115	0.2228
SB7	52.7201	40.0796	6,6035	55.7986	16.4673	26.4197	27.6062	44.2446	1.4314	0.2521
SB4	51.3173	38.9062	8,9260	55.0298	14.211	24.0246	24.3687	42.1385	1.2514	0.1518
FW5	43.2851	31.6835	3,7358	49.9666	9,3728	18.8435	17.0242	34.7984	1.1285	0.0883

Table 1. 2012 Mirex Resuls of Beat tracking task in the Mckinney Dataset sorted by AMLt.

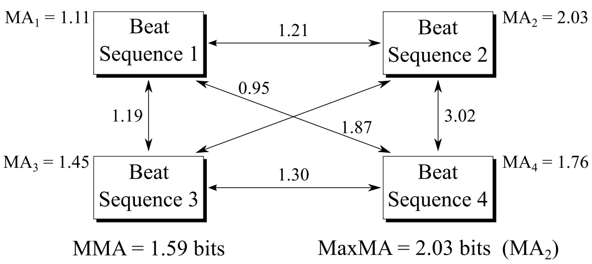

Figure 1. Example calculation of the MMA and MaxMA for a musical piece with the beat sequences estimated from a committee of four beat trackers.

The output selection is done using the Maximum mutual agreement presented in [12] [9].

3. RESULTS

In Table 1 the main results in the beat tracking task shows that the algorithm ZDG2 get the higher results in five of the ten measures (Amlt, AMlc, D, Dg and Goto) in the Mckinney dataset (MCK), and get the second result in two measures (Cemgil and P-score), More information about the measures are presented in [2].

In Table 2 and Table 3 shows that the ZDG2 algorithm has the higher results in the AMLt and AMLc measures comparing each year when the Beat tracking task was presented (Excluding 2006, because the only measure used is the p-score).

Year	Algorithm	AMLt
2012	**ZDG2**	**66.6591**
2011	GP5	66.451
2010	GP6	63.5963
2009	GP1	66.6

Table 2. AMLt best Mirex Resuls of Beat tracking task per year in the Mckinney Dataset.

Year	Algorithm	AMLc
2012	**ZDG2**	**51.759**
2011	GKC2	51.0453
2010	BES3	51.0786
2009	DRP4	50.8

Table 3. AMLc best Mirex Resuls of Beat tracking task per year in the Mckinney Dataset.

4. ACKNOWLEDGEMENTS

We would like to thank all the authors (Norberto Degara and Daniel Ellis) of the available matlab code for their support. Universidad Pontificia Bolivariana (Colombia) and Colciencias.

5. REFERENCES

[1] A. M. Stark, M. E. P. Davies, and M. D. Plumbley, "Real-time beatsynchronous analysis of musical audio," in *12th International Conference on Digital Audio Effects (DAFx-09),* Como, Italy, Sep. 2009.

[2] 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.

[3] 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.

[4] 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

[5] 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.

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

[7] D. P. W. Ellis, "Beat tracking by dynamic programming," *Journal of New Music Research*, vol. 36, no. 1, pp. 51–60, 2007.

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

[9] 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.

[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] 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.

[12] 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.