MIREX 2013 AUDIO BEAT TRACKING EVALUATION: FK1

Florian Krebs
Department of Computational Perception
Johannes Kepler University, Linz, Austria

ABSTRACT

In this paper, we present a Hidden Markov Model (HMM) based beat tracking system that simultaneously extracts downbeats, beat times, tempo, meter and rhythmic patterns. Our model builds upon the basic structure proposed by Whiteley et. al [7], which we further modified by introducing a new observation model: rhythmic patterns are learned directly from data, which makes the model adaptable to the rhythmical structure of any kind of music. The MIREX beat tracking evaluation - 30 results using ten measures and three datasets - placed our algorithm among the top ten performing algorithms 18 times.

1. MODEL DESCRIPTION

The model structure is identical to the one published in [6].

2. PARAMETERS

For this submission we use the following parameters:

- $M = 1920$
- $N = 35$
- $R = 2$
- $p_n = 0.02$
- $p_r = p_t = 0$
- $r \in \{3/4, 4/4\}$
- framelength = 20 ms
- $n_{min} = \{10, 8\}$
- $n_{max} = \{30, 35\}$

3. DATASETS

3.1 Training data

Our training set consists of 1016 audio excerpts: 25 training and bonus files from the MIREX 2006 beat tracking contest, 6 musical pieces from [1], 66 pieces from [2], 697 from the ballroom dataset [6], and 222 from the hainsworth dataset [4].

3.2 Test data

Currently, three evaluation datasets are used in the yearly Music Information Retrieval Evaluation eXchange (MIREX) for audio beat tracking. They are briefly described in this section:

3.2.1 MCK dataset

The MCK dataset contains 160 30-second audio excerpts and was created by the MIREX team in 2006. The recordings are characterized by a stable tempo and a wide variety of instrumentations and musical styles. About 20% of the files have non-binary meters.

3.2.2 MAZ dataset

The MAZ dataset contains piano recordings of 322 Chopin Mazurkas, which also include tempo changes. It was contributed by Craig Sapp in 2009.

3.2.3 SMC dataset

The third collection was contributed by Holzapfel et al [5] in 2012. It consists of 217 excerpts around 40 s each, of which the majority is difficult to track (e.g., because of changes in meter and tempo, bad sound quality, expressive timing). It includes romantic music, film soundtracks, blues, chanson, and solo guitar.

4. EVALUATION

4.1 Evaluation measures

The evaluation measures are specified in [3].

4.2 Results and discussion

The table 1 shows the results of the submitted system $FK1$ together with the maximum score and the rank of $FK1$ within all submissions from the years 2006 and 2009-2013. This leads to a total number of 78 submissions for the MCK dataset, 73 submissions for the MAZ dataset, and 37 submissions for the SMC dataset.

More details of the task results can be found at `www.music-ir.org/mirex/wiki/2013:MIREX2013_Results`.

5. CONCLUSION AND FUTURE WORK

We have introduced a HMM beat tracking system that was trained with real-world music data. Compared to all submissions to MIREX from 2006 to 2013, for all three datasets and all ten performance measures it is ranked 18 (out of 30) times among the top ten performing algorithms. As it was

Dataset		F-Measure	Cemgil	Goto	P-Score	CMLc	CMLt	AMLc	AMLt	D (bits)	Dg (bits)
MCK	Results FK1	57.7	43.8	22.9	61.0	23.8	35.1	44.1	64.8	1.65	0.32
	Results best	59.0	44.9	26.2	63.0	28.8	39.1	53.1	70.1	1.87	0.39
	Rank FK1	3	3	3	4	19	8	36	18	35	8
MAZ	Results FK1	52.9	43.1	0.31	51.5	3.86	23.6	7.44	34.7	0.56	0.33
	Results best	68.5	61.5	2.5	72.2	7.8	50.9	9.7	50.9	2.93	1.95
	Rank FK1	14	14	9	19	31	34	5	11	22	15
SMC	Results FK1	41.6	32.2	10.1	50.0	14.4	21.3	24.9	37.7	1.00	0.20
	Results best	52.4	41.5	21.7	62.4	30.0	39.1	37.3	50.6	1.33	0.45
	Rank FK1	3	3	3	4	6	7	6	9	6	3

Table 1. Results of our algorithm (FK1), results of the best performing algorithm per measure and ranking of FK1 among all different submissions to MIREX 2006-2013

.

trained mainly with pop/rock recordings, it would be interesting to see if the performance on the MAZ dataset could be improved by training the system with piano music with varying tempi. We plan to add various features and rhythmic patterns. As this increases the computational complexity of the algorithm, other approximative inference methods such as particle filtering will be required.

6. ACKNOWLEDGMENTS

This research is supported by the Austrian Science Fund (FWF) under Z159 "Wittgenstein Award". We thank Ingrid Abfalter for correcting and Maarten Grachten for inspiring discussions.

7. REFERENCES

[1] J.P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M.B. Sandler. A tutorial on onset detection in music signals. *IEEE Transactions on Speech and Audio Processing*, 13(5):1035–1047, 2005.

[2] S. Böck, F. Krebs, and M. Schedl. Evaluating the online capabilities of onset detection methods. In *Proceedings of the 14th International Conference on Music Information Retrieval (ISMIR)*, Porto, 2012.

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

[4] S. Hainsworth and M. Macleod. Particle filtering applied to musical tempo tracking. *EURASIP Journal on Applied Signal Processing*, 2004:2385–2395, 2004.

[5] A. Holzapfel, M. Davies, J. Zapata, J. Oliveira, and F. Gouyon. Selective sampling for beat tracking evaluation. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(9):2539–2548, 2012.

[6] F. Krebs, S. Böck, and G. Widmer. Rhythmic pattern modeling for beat- and downbeat tracking in musical audio. In *Proc. of the 14th International Conference on Music Information Retrieval (ISMIR)*, Curitiba, 2013.

[7] N. Whiteley, A. Cemgil, and S. Godsill. Bayesian modelling of temporal structure in musical audio. In *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR)*, Victoria, 2006.