MIREX-09 "AUDIO BEAT TRACKING" TASK: IRCAMBEAT SUBMISSION

Geoffroy Peeters
Ircam Sound Analysis-Synthesis Team - CNRS STMS

ABSTRACT

This extended abstract details a submission to the Music Information Retrieval Evaluation eXchange (MIREX) 2009 for the "Audio Beat Tracking" task. The system named ircambeat performs time-variable tempo and meter estimation, beat and downbeat marking. Detailed description of the two parts are given in [1] and [2]. We briefly summarized them below.

1. IRCAMBEAT IMPLEMENTATION

Ircambeat is a C++ software and library running under Linux, Windows-XP and Mac-OS-X which performs time-variable tempo and meter estimation, beat and downbeat marking.

2. IRCAMBEAT ALGORITHM DESCRIPTION

The flowchart of ircambeat is represented in Figure 1.

2.1 Tempo and meter estimation

The tempo and meter estimation algorithm works in three stages.

First, an onset-energy-function $f(t)$ is extracted from the audio signal by computing a reassigned spectral-energy-flux (using time and frequency reassignement for better precision).

Second, the dominant periodicities of $f(t)$ over time are estimated using a combination of Discrete Fourier Transform and Frequency-Mapped Auto-Correlation-Function. The combination of both allows to better emphasizing the periodicities due to the meter, the beat and the tatum periodicities in $f(t)$. We note $p(t)$ the resulting function.

Finally, a Viterbi decoding algorithm is used to decode simultaneously the tempo and the meter. For this, we define states of a hidden Markov model as all the combinations of possible tempi and meter (among 22: binary grouping of beat/ binary subdivision of beat, 23: binary/ ternary and 32: ternary/ binary). Given $p(t)$, we compute the observation probabilities of the states over time. The decoding then produces the best estimates of tempo and meter over time.

More details about the algorithm can be found in [1].

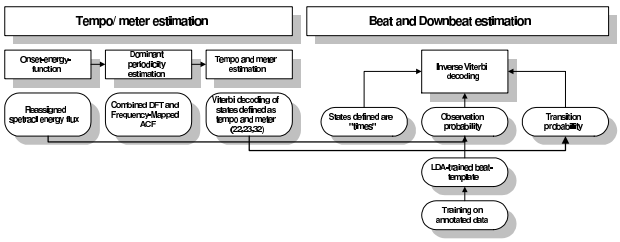

Figure 1. Flowchart of ircambeat

2.2 Beat and downbeat tracking

Beat and downbeat positions are estimated simultaneously using an inverse Viterbi formulation. In this formulation, a state is defined as a specific time. Observation probabilities of states (times) are obtained using a LDA-trained beat-template. This beat-template is obtained by considering the function $f(t)$ inside a measure as a N-dimensional feature vector. A two-class (beat/ non-beat) problem is then solved using LDA and a training set. The resulting LDA-axe is then used as the best beat-template in order to perform discrimination between beat and non-beat positions.

More details about the beat estimation algorithm can be found in [2]. Details about the simultaneous estimation of beat and downbeat are not available since the corresponding paper was rejected to ISMIR09.

3. MIREX09 RESULTS AND DISCUSSIONS

3.1 Experiment

Two test-sets were used for MIREX-09 evaluation:

McKinney Collection: A collection of 160 musical excerpts; the same collection as the one used for the 2006 Audio Tempo Extraction and Beat tasks. Each recording has been annotated by 40 different listeners (39 in a few cases).

Sapp's Mazurka Collection: 322 files drawn from the Mazurka.org dataset put together by Craig Sapp from CHARM / Royal Holloway, University of London. Craig Sapp was also responsible for creating the high-quality ground-truth files.

Ten performance measures have been used for the evalution:

F-measure: the standard calculation as used in onset evaluation but with a $\pm$ 70ms window

Cemgil: beat accuracy is calculated using a Gaussian error function with 40ms standard deviation

Goto: binary decision of correct or incorrect tracking based on statistical properties of a beat error sequence

P-score: McKinney's impulse train cross-correlation method as used in 2006

CMLC: the ratio of the longest continuously correctly tracked section to the length of the file, with beats at the correct metrical level

CMLT: the total number of correct beats at the correct metrical level

AMLC: the ratio of the longest continuously correctly tracked section to the length of the file, with beats at allowed metrical levels

AMLT: the total number of correct beats at allowed metrical levels

D (bits): information based criteria based on analysis of a beat error histogram

Dg (bits): same

Details about the MIREX-09 "Audio Beat Tracking" task can be found at http:// www.music-ir.org/ mirex/ 2009/ index.php/ Audio_Beat_Tracking.

ircambeat was tested in four different configurations

GP1 VF: (v)ariable-over-time tempo estimation, meter is (f)orced to be 22

GP2 VE: (v)ariable-over-time tempo estimation, meter is (e)stimated

GP3 CF: (c)onstant-over-time tempo estimation, meter is (f)orced to be 22

GP4 CE: (c)onstant-over-time tempo estimation, meter is (e)stimated

3.2 Results

For the McKinney Collection test-set (see Table 1 for details), for 8 criteria over 10 ircambeat was the best algorithm, and this whatever configuration of ircambeat (GP1, GP2, GP3, GP4). Only according to the criteria AMLC and D, was the DRP4 algorithm better. According to these two criteria, ircambeat will rank second and this whatever configuration of ircambeat (GP1, GP2, GP3, GP4).

It is interesting to compare for each algorithm, the ratio of the AMLC value to the CMLC value (or AMLT to CMLT). Since the difference between both criteria is the consideration of octave error (A) or not (C), this ratio gives information on the amount of octave errors produced by each algorithm. This ratio is around 1.8 for all ircambeat configurations, and for the DRP2 algorithm. It is larger than 2 (from 2 to 2.5) for the other algorithms. It means that the number of octave errors produced by ircambeat and DRP2 is smaller than the one of the other algorithms.

Since this test-set is the same as the one used in the MIREX-06 "Audio Beat Tracking" task, and since the P-score is available for both MIREX-06 and MIREX-09, it is interesting to compare the best performances obtained in MIREX-06 to the ones obtained in MIREX-09. The largest P-score obtained in MIREX-06 was 0.575 by Dixon. It is interesting to see that ircambeat whatever configuration of it (GP1, GP2, GP3, GP4) has a higher P-score (up to .592). (see http:// www.music-ir.org/ mirex/ 2006/ index.php/ Audio_Beat_Tracking_Results for details).

Mc Kinney Collection										
*TeamID	F-Measu	Cemgil	Goto	P-Score	CMLC	CMLT	AMLC	AMLT	D	Dg
*units ->	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(bits)	(bits)
DRP1	24,6	17,5	0,4	37,4	3,8	9,6	7,8	19,2	0,625	0,008
DRP2	35,4	26,1	3,1	46,5	7,5	16,9	13,8	31,7	1,073	0,062
DRP3	47	35,6	16,2	52,8	19,4	27,6	45,3	61,3	1,671	0,245
DRP4	48	36,3	19,5	53,9	22,2	29,1	**50,8**	64	**1,739**	0,255
GP1	**54,8**	**41**	**22,2**	59	26	35,5	49,1	**66,6**	1,68	**0,294**
GP2	53,7	40,1	20,9	57,9	25,1	33,7	47,6	63,8	1,649	0,281
GP3	54,6	40,9	22	59	26,1	**35,6**	48,8	66,3	1,695	0,287
GP4	54,5	40,8	21,6	**59,2**	**26,4**	35,5	48,2	64,7	1,685	0,289
OGM1	39,7	29,1	6,5	47,8	11,9	19,9	29,5	47,4	1,283	0,1
OGM2	41,5	30,4	5,7	49,2	11,7	19,5	28,2	46,4	1,34	0,092
TL	50	37	14,1	53,6	17,9	24,3	35,9	50,8	1,48	0,15

Table 1. MIREX-09 "Audio Beat Tracking" results for the McKinney Collection

For the Sapp's Mazurka Collection (see Table 2 for details), the best performing algorithm is DRP3 of Matthew Davies from Queen Mary University of London, whatever criteria. The best performing configuration of ircambeat is GP2 (variable-over-time tempo estimation, meter is estimated). This makes sens considering the time-variable tempo nature of Mazurka interpretations. The GP2 configuration ranked 2nd for 8 criteria over 10 (except the Goto and D criteria). The configuration GP1 (variable-over-time tempo estimation, meter forced to 22) ranked 3rd for 8 criteria over 10. Considering this, it seems that the main point of this test-set was the estimation of time-variable tempo (as used by GP1 and GP2 but not GP3 and GP4) rather than estimation of the meter.

Sapp's Mazurka Collection										
*TeamID	F-Measu	Cemgil	Goto	P-Score	CMLC	CMLT	AMLC	AMLT	D	Dg
* units ->	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(bits)	(bits)
DRP1	27,9	19,9	0	37,3	1,5	10,3	2,3	13,8	0,126	0,008
DRP2	48,4	32,5	0	55,5	2,6	26	3,1	28,1	0,683	0,183
DRP3	**67,8**	**61,5**	**2,5**	**65,3**	**7,8**	**42,1**	**9,2**	**45,6**	**1,424**	**0,993**
DRP4	45,3	37,6	0,6	50,1	3,9	24,9	5,3	28,5	0,393	0,198
GP1	54,2	42,5	0	59,5	4,4	33,2	5,6	36,6	0,447	0,252
GP2	54,7	43,1	0	59,9	4,5	33,7	5,7	37	0,467	0,269
GP3	44,7	34,8	0	51,2	3,2	24,7	4	27,9	0,262	0,111
GP4	45,4	35,6	0	51,9	3,3	25,2	4	28,4	0,284	0,124
OGM1	30,7	21,9	0	38,6	1,7	11	2,8	15,4	0,134	0,014
OGM2	32,1	23,1	0	39,7	1,7	11,5	2,7	15,6	0,134	0,016
TL	44,9	35,7	0,3	51	4,4	20,5	4,7	22,5	0,44	0,095

Table 2. MIREX-09 "Audio Beat Tracking" results for the Sapp's Mazurka Collection

Whatever configuration of ircambeat (GP1, GP2, GP3, GP4), it is also the fastest algorithm for "Audio beat Tracking" (see Table 3 for details).

Details about the results of the tasks can be found at http:// www.music-ir.org/ mirex/ 2009/ index.php/ Audio_Beat_Tracking_Results.

*Participant	Machine	Runtime (hh:mm)
DRP1	ALE	00:35
DRP2	ALE	00:42
DRP3	ALE	01:19
DRP4	ALE	01:16
GP1	ALE	**00:33**
GP2	ALE	**00:34**
GP3	ALE	**00:34**
GP4	ALE	**00:32**
OGM1	ALE	01:52
OGM2	ALE	01:51
TL	ALE	00:51

Table 3. MIREX-09 "Audio Beat Tracking" computation time

4. CONCLUSION

This extended abstract reviewed ircambeat submission for the "Audio Beat Tracking" task of MIREX-09. Ircambeat whatever configuration of it (constant tempo, time-variable tempo, fixed meter, estimated meter) performed the best on the McKinney Collection test-set. Compared to the P-score results obtained on this test-set in MIREX-06, it also provides better results than the best results obtained in 2006. Ircambeat in "time-variable tempo estimation" configuration performed second and third for the Sapp's Mazurka Collection. Finally, ircambeat was the fastest algorithm for the "Audio Beat Tracking" task.

5. ACKNOWLEDGMENTS

This work was partly supported by "Quaero" Programme, funded by OSEO, French State agency for innovation

6. REFERENCES

[1] G. Peeters. Template-based estimation of time-varying tempo. *EURASIP Journal on Advances in Signal Processing*, 2007(Special Issue on Music Information Retrieval Based on Signal Processing):Article ID 67215, 14 pages, 2007. doi:10.1155/2007/67215.

[2] G. Peeters. Beat-marker location using a probabilistic framework and linear discriminant analysis. In *Proc. of DAFX*, Como, Italy, 2009.