MIREX 2017: COVER SONG IDENTIFICATION USING SIMILARITY FUSION OF HPCPS, MFCCS, AND MFCC SSMS

Christopher J. Tralie
Duke University Department of
Electrical and Computer Engineering
ctralie@alumni.princeton.edu

1. INTRODUCTION

This submission uses the recently developed technique in [7], which fuses beat-synchronous blocked HPCP, MFCC, and MFCC SSM features at the cross-similarity level before local alignment between every pair of songs, in addition to fusing features after all pairwise scores have been computed with each individual feature set. Please refer to [7] for more details.

2. IMPLEMENTATION DETAILS

Since the technique in [7] relies on synchronizing features to the beat, it must address the tendency of beat trackers to return arbitrary metrical levels which, in the case of 4/4, for instance, may return beat intervals which are doubles or halves of each other between two cover songs (the so-called "octave problem"). To mitigate this, [7] compares features computed with a dynamic programming beat tracker [3] biased at 3 different levels, which blows up the computation by a factor of 9, which is infeasible on larger datasets. To make this technique fast enough to meet the time constraints of MIREX, we instead use a state of the art beat tracker from [6] implemented with the Madmom library [1] to return a single beat level for each song. This beat tracker is based on trained recurrent neural networks [2] to create beat activation functions, followed by a dynamic Bayes network to extract beat positions [5], the combination of which is specifically designed to deal with the octave problem inherent in many beat trackers.

3. RESULTS ON BENCHMARK DATASETS

Since we are only using one set of beats for each song instead of taking the best match over many different beats, we expect to take a performance hit over [7]. However, since [6] is a high quality beat tracker, the performance is still reasonable in our experiments. In particular, Table 1 shows the results on the covers 80 dataset [4], and Table 2 shows the results on the covers 1000 dataset [7]. For reference, the best scores (late fusion) from [7] using all beat levels are shown. Fortunately, the late fusion is still quite good on both datasets, even with the performance hit due to using only one beat level.

Table 1. Results on the Covers 80 dataset

	MR	**MRR**	**Top1**	**Top10**	**../80**
MFCCs	31.5125	0.531269	79	93	41/80
SSMs	24.15	0.578772	87	102	45/80
Chromas	24.9	0.616607	93	107	46/80
SNF	18.3625	0.745443	115	130	59/80
Late	16.1938	0.762246	117	129	61/80
[7] Late	7.59	0.873	136	144	69/80

Table 2. Results on the Covers 1000 dataset

	MR	**MRR**	**Top1**	**Top10**
MFCCs	112.241	0.549329	523	633
SSMs	92.68	0.572648	533	694
Chromas	64.195	0.686407	653	793
SNF	54.588	0.751717	720	835
Late	38.262	0.82227	803	878
[7] Late	14	0.904	884	950

4. ACKNOWLEDGEMENTS

Christopher Tralie was partially supported by an NSF Graduate Fellowship NSF under grant DGF-1106401 and an NSF big data grant DKA-1447491.

5. REFERENCES

[1] Sebastian Böck, Filip Korzeniowski, Jan Schlüter, Florian Krebs, and Gerhard Widmer. madmom: a new python audio and music signal processing library. In *Proceedings of the 2016 ACM on Multimedia Conference*, pages 1174–1178. ACM, 2016.

[2] Sebastian Böck and Markus Schedl. Enhanced beat tracking with context-aware neural networks. In *Proc. Int. Conf. Digital Audio Effects*, pages 135–139, 2011.

[3] Daniel PW Ellis. Beat tracking by dynamic programming. *Journal of New Music Research*, 36(1):51–60, 2007.

[4] Daniel PW Ellis. The "covers80" cover song data set. *URL: http://labrosa. ee. columbia. edu/projects/coversongs/covers80*, 2007.

[5] Florian Krebs, Sebastian Böck, and Gerhard Widmer. Rhythmic pattern modeling for beat and downbeat

tracking in musical audio. In *ISMIR*, pages 227–232, 2013.

[6] Florian Krebs, Sebastian Böck, and Gerhard Widmer. An efficient state-space model for joint tempo and meter tracking. In *ISMIR*, pages 72–78, 2015.

[7] Christopher J Tralie. Mfcc and hpcp fusion for robust cover song identification. In *18th International Society for Music Information Retrieval (ISMIR)*, 2017.