

MIREX 2017 DRUM TRANSCRIPTION SUBMISSIONS

Chih-Wei Wu

Center for Music Technology, Georgia Institute of Technology
cwu307@gatech.edu

ABSTRACT

This extended abstract provides information on the 2017 MIREX submission for drum transcription task. The submission includes one Non-negative Matrix Factorization (NMF) based approach and its two adaptive variants.

1. SUBMITTED ALGORITHMS

Three submitted systems, namely the **CW1**, **CW2**, and **CW3**, were originally introduced in [2]; the same methods have been applied to related tasks such as music genre classification [1] and drum playing technique detection [3]. The implementation of these algorithms is publicly available from the online repository.¹ All of the parameters, including the initial drum dictionary, peak picking threshold, and harmonic rank, were set to default. No specific hyperparameter tuning was conducted on the public set for this submission. More information about the submitted algorithms are given as follows:

CW1 is based on Partially-Fixed Non-negative Matrix Factorization (PFNMF). The basic concept is to initialize a NMF process with predefined drum dictionary as well as undefined entries for absorbing the harmonic content in the signals. For this MIREX submission, the harmonic rank R_H is set to 50.

CW2 is a variant of PFNMF with additional iterative steps for template adaption (also referred to as *AMI* in [2]). This adaptation process seeks to refine the initial drum dictionary by combining the learned harmonic templates based on their cross-correlation with the target drum instruments. The ρ_{thres} is set to 0.5.

CW3 is another variant of PFNMF (referred to as *AM2* in [2]) with a slightly different template adaptation process. This process starts by alternatively fixing the drum dictionary and drum activation matrix while allowing the other matrices to update and converge to better solutions.

¹ <https://github.com/cwu307/NmfDrumToolbox>, Last accessed: 10/30/2017

Table 1. The resulting averaged F-measure on the evaluation (test) set.

Dataset	CW1	CW2	CW3
IDMT	0.37	0.41	0.42
KT	0.48	0.52	0.48
RBMA	0.50	0.54	0.54
MDB	0.62	0.62	0.59
GEN	0.60	0.65	0.63
Overall	0.51	0.55	0.53

2. RESULTS

The evaluation is done on the evaluation set for MIREX drum transcription task 2017. The results are shown in Table 1.

Potential improvements could be achieved by 1) extracting a new initial drum dictionary from the public set. 2) selecting the suitable R_H based on the complexity of the audio signal (e.g., $R_H = 0$ for drum-only signals). 3) selecting the optimal peak picking threshold via grid search on the public set.

3. REFERENCES

- [1] Athanasios Lykartsis, Chih-Wei Wu, and Alexander Lerch. Beat histogram features from nmf-based novelty functions for music classification. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, 2015.
- [2] Chih-Wei Wu and Alexander Lerch. Drum transcription using partially fixed non-negative matrix factorization with template adaptation. In *Proceedings of International Society for Music Information Retrieval Conference (ISMIR)*, 2015.
- [3] Chih-Wei Wu and Alexander Lerch. On drum playing technique detection in polyphonic mixtures. In *Proceedings of International Society for Music Information Retrieval Conference (ISMIR)*, 2016.

