MDB DRUMS – AN ANNOTATED SUBSET OF MEDLEYDB FOR AUTOMATIC DRUM TRANSCRIPTION

Carl Southall[1], **Chih-Wei Wu**[2], **Alexander Lerch**[2], **Jason Hockman**[1]

[1] DMT Lab, Birmingham City University
[2] Center for Music Technology, Georgia Institute of Technology
{carl.southall, jason.hockman}@bcu.ac.uk, {cwu307, alexander.lerch}@gatech.edu

ABSTRACT

In this paper we present *MDB Drums*, a new dataset for automatic drum transcription (ADT) tasks. This dataset is built on top of the MusicDelta subset of the MedleyDB dataset, taking advantage of real-world recordings in multi-track format. The dataset is comprised of a variety of genres, providing a balanced pool for developing and evaluating ADT models with respect to various musical styles. To reduce the cost of the labor-intensive process of manual annotation, a semi-automatic process was utilised in both the annotation and quality control processes. The presented dataset consists of 23 tracks with a total of 7994 onsets. These onsets are divided into 6 classes based on drum instruments or 21 subclasses based on playing techniques. Every track consists of a drum-only track as well as multiple accompanied tracks, enabling audio files containing different combinations of instruments to be used in the ADT evaluation process.

1. INTRODUCTION

Automatic drum transcription (ADT) is a task that concerns the extraction and detection of drum events from a stream of audio signals. Like many research fields that rely on data-driven approaches, the advance of ADT research is linked to the quality of the existing annotated datasets. To date, there are a few publicly available datasets such as ENST drums [1] and IDMT-SMT drums [2], which are commonly used in recent ADT related studies [3, 4]. However, as discussed in [4], most of the existing datasets only contain annotations of basic techniques (e.g., strikes, cross sticks, rim shots), and more detailed annotations on techniques such as flam, drag, and buzz rolls are missing. Additionally, the existing datasets are limited in different aspects such as variety of musical styles and complexity of drum patterns. In this paper we present *MDB Drums*, a newly annotated drum dataset to address these limitations.

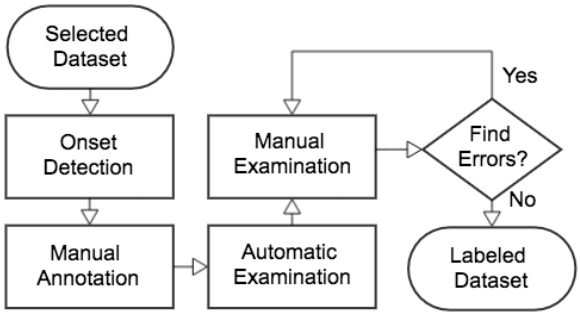

Figure 1. Overview of the creation process of *MDB Drums* dataset.

2. DATASET CREATION

2.1 Overview

The creation process of *MDB Drums* is shown in Fig.1. Since the objective of this project was to create a realistic drum dataset without introducing the extra cost of recording and post-processing, we chose to annotate an existing dataset. The MedleyDB dataset [1] was chosen as it contains multi-track recordings of the drums, allowing for easier annotation. As the original purpose of the dataset was not for ADT tasks, the recording quality of drum tracks varies widely across the dataset, which introduced a layer of inconsistency. As a result, the MusicDelta subset was selected as it contains recordings of better quality with reasonable durations. More details are presented in the following sections.

2.2 Annotation Stage

Dataset creation began with the annotation stage, which consisted of two steps: 1) onset detection and 2) manual annotation. To facilitate the process, we used the OnsetDetector algorithm from the madmom [2] library on each of the audio files. The use of the drum-only tracks in MedleyDB allowed the onset detector to consistently achieve reliable results. Next, the extracted onset times were imported into Sonic Visualizer [3] for refinement by human an-

[1] http://perso.telecom-paristech.fr/∼grichard/ENST-drums, last access: 02/10/17

[2] https://www.idmt.fraunhofer.de/en/business_units/m2d/smt/drums.html, last access: 02/10/17

[3] http://www.sonicvisualiser.org, last access: 02/10/17

Table 1. Onset classes in *MDB Drums* dataset.

Class	Subclass	Description	Onsets
KD	KD	kick drum	1539
SD	SD	snare drum	1510
	SDB	snare drum: brush	332
	SDD	snare drum: drag	2
	SDF	snare drum: flam	11
	SDG	snare drum: ghost note	790
	SDNS	snare drum: no snare	9
HH	CHH	closed hi-hat	1847
	OHH	open hi-hat	269
	PHH	pedal hi-hat	523
TT	HIT	high tom	4
	MHT	high-mid tom	26
	HFT	high floor tom	14
	LFT	low floor tom	46
CY	RDC	ride cymbal	835
	RDB	ride cymbal bell	16
	CRC	crash cymbal	126
	CHC	china cymbal	15
	SPC	splash cymbal	10
OT	SST	side stick	38
	TMB	tambourine	32

notators. Each onset was annotated with its corresponding class, and missing onsets were added.

2.3 Examination Stage

To ensure the consistency and accuracy of the resulting annotations, all files were assessed automatically and verified manually in the examination stage. For the automatic examination, three checks were implemented: a check for invalid class labels (i.e., labels with incorrect names), a check for duplicate labels within a 50ms window, and a check for three or more different labels occurring within the 50ms window. These rules were determined heuristically based on the annotators' domain knowledge. For the manual examination, a cross-check was conducted on all of the tracks. This required the two annotators to examine each other's assignments. Additionally, the dataset was examined by an external reviewer who was not involved in the annotation stage.

3. DATASET DETAILS

The *MDB Drums* dataset consists of a total of 23 tracks with an average length of 54 seconds. However, as the dataset contains multi-track files (i.e., all the isolated instrumental tracks are included), a variety of combinations can be easily generated (e.g., drum + guitar, drum + bass guitar). In total there are 7994 onsets divided across 6 classes, namely the kick drum (KD), snare drum (SD), hi-hat (HH), toms (TT), cymbals (CY), and other percussion (OT). Table 1 highlights additional categorisation of the onsets into the 21 subclasses based on playing technique. The full dataset is available for download from the *MDB Drums* Github repository. [4]

[4] https://github.com/CarlSouthall/MDBDrums, last access: 02/10/17

3.1 MIREX Split

Prior to public release, the *MDB Drums* dataset was used in the MIREX 2017 drum transcription evaluation. Both the drum-only and full-mix tracks were used for the evaluation resulting in 46 tracks in total. For comparison against the achieved results, the training and tests splits used can also be found in the Github repository.

4. CONCLUSION

The contributions of the presented dataset can be summarized as follows: first, this dataset covers a diverse range of genres (e.g., such as Rock, Country, Disco, Reggae, and Jazz) which provides a more representative sample pool of the real-world music than the existing datasets. Second, all tracks are real-world recordings, which include drums played by musicians. As these are real drum tracks—as opposed to synthesized or sample-based audio—they are capable of capturing the musical expression of percussionists and reflect the difficulty of ADT in real-world recordings. Third, this dataset includes the most commonly used drum classes (e.g., kick drum, snare drum and hi-hat) as well as additional drum classes (e.g., toms and cymbals) in conjunction with detailed categorisation for each class (e.g., flam, drag, and roll for snare drum), encouraging further ADT studies such as the detection of playing technique. Finally, all tracks contain multi-track files (e.g., drum-only tracks and multiple accompanied tracks) allowing for a more diverse range of instrument combinations to be used in the evaluation of ADT systems.

5. ACKNOWLEDGEMENTS

The authors would like to thank Rachel Bittner and the MedleyDB team for the creation of the MedleyDB dataset and also Maciek Tomczak for helping to review the annotations.

6. REFERENCES

[1] Rachel M Bittner, Justin Salamon, Mike Tierney, Matthias Mauch, Chris Cannam, and Juan Pablo Bello. MedleyDB: A multitrack dataset for annotation-intensive MIR research. In *Proc. Intl. Society for Music Information Retrieval Conf. (ISMIR)*, 2014.

[2] Sebastian Böck, Filip Korzeniowski, Jan Schlüter, Florian Krebs, and Gerhard Widmer. madmom: A new Python audio and music signal processing library. In *Proc. of the ACM Intl. Conf. on Multimedia*, Amsterdam, The Netherlands, 2016.

[3] Carl Southall, Ryan Stables, and Jason Hockman. Automatic drum transcription using bi-directional recurrent neural networks. In *Proc. Intl. Society for Music Information Retrieval Conf. (ISMIR)*, 2016.

[4] Chih-Wei Wu and Alexander Lerch. On drum playing technique detection in polyphonic mixtures. In *Proc. Intl. Society for Music Information Retrieval Conf. (ISMIR)*, 2016.