

MIR Task: Automatic Drum Transcription

Pedro Ricardo

February 6, 2019

Introduction and Task Proposal

According to MIREX2018 [1], drum transcription is defined as the task of detecting the positions in time and labeling the drum class of drum instrument onsets in polyphonic music. The article *A Review of Automatic Drum Transcription* [2] offers an in-debt explanation about automatic drum transcription as well as an extremely up to date state of the art.

The present task proposes to test and evaluate an automatic drum transcription algorithm, implemented using ADTLib [3] [4] and Essentia, in a beatbox (vocal percussion) dataset [5].

ADTLib

Automatic Drum Transcription Library (ADTLib) is a library that contains open source ADT algorithms to aid other researchers in areas of music information retrieval (MIR). The algorithms return both a .txt file of kick drum, snare drum, and hi-hat onsets and an automatically generated drum tablature. Further information about the library and how it was implemented can be found on *Automatic drum transcription for polyphonic recordings using soft attention mechanisms and convolutional neural networks* [7]. A web-based implementation of ADTLib is also available online for free - ADTWEB [8].

After some research on the topic, there wasn't found any type of algorithm implemented directly in Essentia that deals with this particular problem. *Exploration of Techniques for Automatic Labeling of Audio Drum Tracks' Instruments* [9] by Perfecto Herrera and Fabien Gouyon, produced in the MTG in 2001, is used as a reference for the implementation of the algorithm StrongDecay [10] in Essentia. Although this algorithm is related to the recognition of percussive instruments, it does not tackle this problem directly.

Beatboxset1

The beatboxset1: beatboxing audio data set available in the Audio Content Analysis [6] contains beatboxing recordings from various contributors, who recorded the clips themselves in various conditions. There is a spreadsheet file "beatboxset1.csv" accompanying the dataset that provides metadata for

the recordings. Further annotations of the recordings are also included: these mark the positions of onsets as well as categorizing the events into a handful of standard classes.

Evaluation Method

The evaluation method focus in understanding how well the implemented algorithm performs by comparing the hand-made annotations, available in the data set, to the annotations produced by the algorithm. To do this, a F-measure value will be computed for each different sound: kick, snare and hi-hat. This method is based on the evaluation procedure of the MIREX2016 contest related to Audio onset detection [11]. After an F-measure is computed for each individual sound, a value combining these 3 different F-measure values will be calculated.

Some Notes

Stephane Dupont worked on this problem [12] in 2015. By using various pitch tracking and onset detections methods he tried to achieve automatic classification of human beat box sounds (Kick, hi-hat and snare), his study focused on identifying single note events instead of events on a polyphonic recording.

The following sub-tasks, used to organize the research, were created having in consideration some of the MIR Course Content [13] already in existence.

1. Understand the goal and overview of audio drum transcription
 - (a) Understand the input and output
 - (b) Understand the evaluation metrics
2. Search for suitable datasets for the particular task
3. Understand the State-of-the-Art of audio drum transcription
4. Implementation of a drum transcription algorithm using ADTLIB and Essentia on Jupiter notebooks.
5. Understanding and evaluating the effectiveness of the particular algorithm when applied to the particular dataset
6. Conclusions

References

- [1] https://www.music-ir.org/mirex/wiki/2018:Drum_Transcription
- [2] Wu, C. W., Dittmar, C., Southall, C., Vogl, R., Widmer, G., Hockman, J., ... Lerch, A. (2018). A Review of Automatic Drum Transcription. *IEEE/ACM Transactions on Audio Speech and Language Processing*, 26(9), 1457–1483. <https://doi.org/10.1109/TASLP.2018.2830113>

- [3] <https://github.com/CarlSouthall/ADTLib>
- [4] Southall, C., Jillings, N., Stables, R., Hockman, J., & Kingdom, U. (2017). Adtweb : an Open Source Browser Based Automatic Drum Transcription System. In Ismir-2017.
- [5] <https://archive.org/details/beatboxset1>
- [6] <https://www.audiocontentanalysis.org/data-sets/>
- [7] Southall, C., Stables, R., & Hockman, J. (2017). Automatic drum transcription for polyphonic recordings using soft attention mechanisms and convolutional neural networks. Proc. of the International Society for Music Information Retrieval Conference (ISMIR), 606–612. [https://doi.org/10.1016/S0016-7878\(78\)80013-3](https://doi.org/10.1016/S0016-7878(78)80013-3)
- [8] <http://dmtlab.bcu.ac.uk/ADT/>
- [9] Herrera, P., Gouyon, F. (2002) Exploration of Techniques for Automatic Labeling of Audio Drum Tracks' Instruments
- [10] https://essentia.upf.edu/documentation/reference/std_StrongDecay.html
- [11] https://www.music-ir.org/mirex/wiki/2016:Audio_Onset_Detection
- [12] Dupond, S. (2015) Analysis and Automatic Recognition of Human Beatbox Sounds: a Comparative Study Article
- [13] <https://musicinformationretrieval.wordpress.com/2017/01/17/audio-onset-detection-week-1/>