MIXING SECRETS: A MULTI-TRACK DATASET FOR INSTRUMENT RECOGNITION IN POLYPHONIC MUSIC

Siddharth Gururani, Alexander Lerch
Georgia Institute of Technology, Center for Music Technology
{siddgururani, alexander.lerch}@gatech.edu

ABSTRACT

Instrument recognition as a task in MIR is largely data driven. This drives a need for large datasets that cater to the need of these algorithms. Several datasets exist for the task of instrument recognition in monophonic signals. For polyphonic music, creating a finely labeled dataset for instrument recognition is a difficult task and using multi-track data eases that process. We present 250+ multi-tracks that have been labeled for instrument recognition and release the annotations to be used in the community. The process of data acquisition, cleaning and labeling has been detailed in this paper.

1. INTRODUCTION

The task of instrument recognition in music is a widely researched task in Music Information Retrieval. It is usually performed for three different scenarios: (i) Instrument recognition in monophonic music, (ii) Predominant instrument recognition in polyphonic music, (iii) Instrument identification in polyphonic music. As most approaches to the task are data-driven, there is need for publicly available datasets for evaluating and comparing methods. We introduce a new multi-track dataset that has been labeled for the task of instrument identification in polyphonic music.

2. DATASETS FOR INSTRUMENT RECOGNITION

As musical instrument recognition is a well established task in the MIR community, multiple public datasets have been published and used over the past two decades.

2.1 Monophonic Instrument Recognition

The following datasets are most commonly used in literature for monophonic instrument recognition.

- RWC Database [6]: A part of the RWC Database is the RWC Musical Instrument Sound Database which covers the full tonal range, playing techniques and dynamics of 50 musical instruments.
- The University of Iowa Musical Instrument Samples (MIS) Dataset [5]: This dataset consists of single notes for 10 woodwind, 5 brass, 4 strings, 9 percussion instruments and a piano.

2.2 Predominant Instrument Recognition

IRMAS Dataset [3]: This dataset contains 11 instruments. It is split into training (6705 audio files, 3 seconds long) and testing (2874 audio files, 5 – 20 seconds long) sets.

2.3 Instrument recognition in Polyphonic Music

While there are not any datasets, to the best of our knowledge, that have been released specifically for this task, multi-track datasets are easily utilizable for polyphonic instrument recognition by processing raw or stem tracks for instruments to determine their activation in a mix.

- MedleyDB [1] [2]: This multi-track dataset consists of 250 multi-tracks. It contains automatically generated annotations for instrument activations which make it highly suitable for the task.
- QUASI Database: This dataset consists of 11 multi-tracks. It is primarily a source separation dataset.

3. MIXING SECRETS DATASET

The dataset presented in this paper is sourced from the website: Mixing Secrets for The Small Studio. [1] The website hosts around 300 multi-tracks free of charge for academic purposes, making it perfect for use in the MIR community. The tracks are professionally recorded. These multi-tracks are also listed on The Open MultiTrack Testbed [4]. The multi-tracks span several genres, and while a detailed quantitative analysis has not been carried out yet, it is definitely a promising source of data for tasks such as source separation and instrument recognition.

3.1 Data Format

Each archive obtained from the website contains only the raw tracks that may be used by mixing engineers to generate a mix of the song. Most raw tracks are named based on the instrumentation of the audio. Each archive is associated with a preview mix mp3 encoded file which is not a part of the archive and has to be downloaded separately.

[1] http://www.cambridge-mt.com/ms-mtk.htm (last accessed: 18-October-2017)

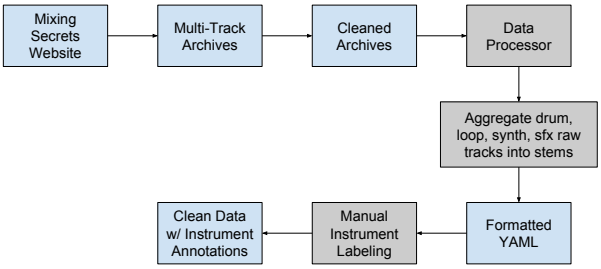

Figure 1. Flowchart of Dataset Creation Process

3.2 Data Acquisition

The data was acquired using a Python script that parses the webpage for the urls of all full multi-track zip archives and mix mp3s. After acquiring the URLs, a download manager was used to download all the data.

3.3 Data Cleaning

There were broken archives in the list which were ignored. Some of the tracks' naming convention was inconsistent with the majority and hence those multi-tracks were also ignored. More details on how these were identified and a list of all remaining multi-tracks is available in the repository. The cleaned set consists of 258 multi-tracks.

3.4 Data Labeling

Since MedleyDB has defined an instrument taxonomy as well as a metadata standard for multi-track data, this dataset was labeled to be as consistent with MedleyDB as possible. A YAML format similar to MedleyDB is generated using a Python Script. [2] In addition, all drum tracks, excluding room mics, are automatically mixed by averaging across tracks. Synth, Loops and Sound FX are similarly grouped into one stem and labeled appropriately.

A major inconsistency between this data and MedleyDB, especially in the context of instrument recognition, is that electric guitars are all grouped into one category, while in MedleyDB, there is a separate label for clean and distorted guitars. Similarly, vocals aren't separated into male or female singers. To address these and possibly other inconsistencies, one annotator passed through the entire dataset to label the remaining raw tracks' instruments. After labeling, instrument activation annotations are generated using code from MedleyDB's annotation module.

4. CONCLUSION

This paper presents a dataset for instrument recognition in polyphonic music in the form of a multi-track dataset consisting of 258 songs. The dataset presented addresses the lack of large amounts of data for the task and augments other datasets currently in use, such as MedleyDB.

[2] https://github.com/SiddGururani/mixing_secrets

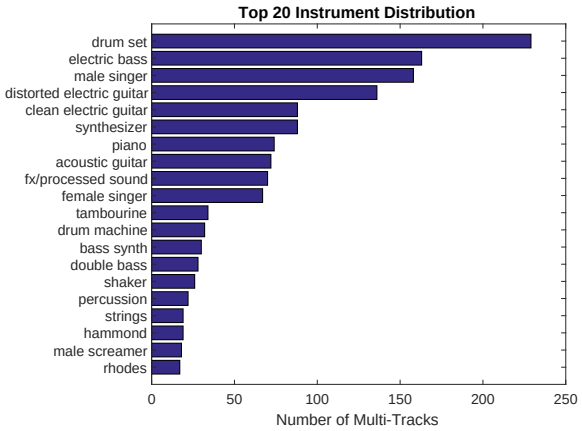

Figure 2. Top 20 Instruments Distribution

5. ACKNOWLEDGEMENTS

The work was funded by Gracenote. We would like to thank them for their support.

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. of the International Society for Music Information Retrieval Conference (ISMIR)*, volume 14, pages 155–160, 2014.

[2] Rachel M Bittner, Julia Wilkins, Hanna Yip, and Juan P Bello. Medleydb 2.0: New data and a system for sustainable data collection. *ISMIR Late Breaking and Demo Papers*, 2016.

[3] Juan J Bosch, Jordi Janer, Ferdinand Fuhrmann, and Perfecto Herrera. A comparison of sound segregation techniques for predominant instrument recognition in musical audio signals. In *Proc. of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 559–564, 2012.

[4] Brecht De Man, Mariano Mora-Mcginity, György Fazekas, and Joshua D Reiss. The open multitrack testbed. In *Proc. of the Audio Engineering Society Convention 137*. Audio Engineering Society, 2014.

[5] Lawrence Fritts. University of iowa musical instrument samples. `http://theremin.music.uiowa.edu/MIS.html`, 1997. Online; accessed: 18-October-2017.

[6] Masataka Goto, Hiroki Hashiguchi, Takuichi Nishimura, and Ryuichi Oka. Rwc music database: Music genre database and musical instrument sound database. In *Proc. of the International Conference on Music Information Retrieval*, pages 229–230. Johns Hopkins University, 2003.