MUSDB18 - a corpus for music separation

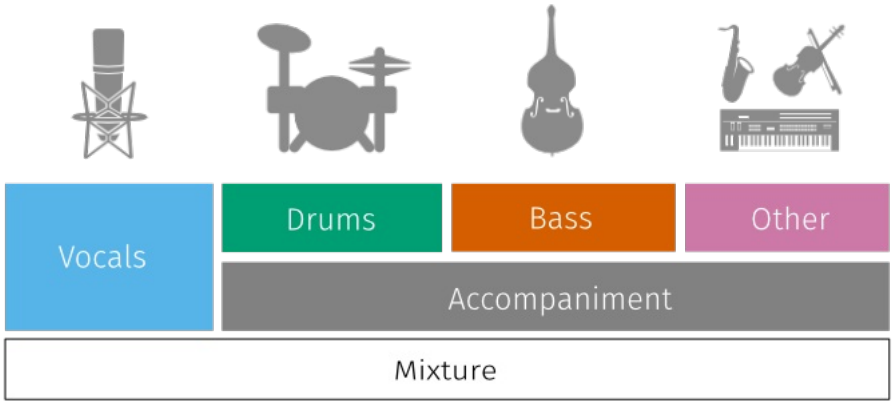
Zafar Rafii, Antoine Liutkus, Fabian-Robert Stöter, Stylianos Ioannis Mimilakis, Rachel Bittner

MUSDB18 - a corpus for music separation

DOI 10.5281/zenodo.1117372

Authors

- **Zafar Rafii**, Gracenote, Media Technology Lab, USA
- **Antoine Liutkus**, Inria, LIRMM, University of Montpellier, France
- **Fabian-Robert Stöter**, Inria, LIRMM, University of Montpellier, France
- **Stylianos Ioannis Mimilakis**, Fraunhofer Institute for Digital Media Technology, Germany
- **Rachel Bittner**, Music and Audio Research Lab, USA



Abstract

The sigsep musdb18 data set consists of a total of 150 full-track songs of different styles and includes both the stereo mixtures and the original sources, divided between a training subset and a test subset.

Its purpose is to serve as a reference database for the design and the evaluation of source separation algorithms. The objective of such signal processing methods is to estimate one or more sources from a set of mixtures, e.g. for karaoke applications. It has been used as the official dataset in the professionally-produced music recordings task for SiSEC 2018, which is the international campaign for the evaluation of source separation algorithms.

Presentation

The *musdb18* is a dataset of 150 full lengths music tracks (~10h duration) of different styles along with their isolated *drums*, *bass*, *vocals* and *others* stems.

musdb18 contains two folders, a folder with a training set: "train", composed of 100 songs, and a folder with a test set: "test", composed of 50 songs. Supervised approaches should be trained on the training set and tested on both sets.

All files from the *musdb18* dataset are encoded in the Native Instruments stems format (.mp4). It is a multitrack format composed of 5 stereo streams, each one encoded in AAC @256kbps. These signals correspond to:

- `0` - The mixture,
- `1` - The drums,
- `2` - The bass,
- `3` - The rest of the accompaniment,
- `4` - The vocals.

For each file, the mixture correspond to the sum of all the signals.

⚠
Since the mixture is separately encoded as AAC, there there is a small difference between the **sum of all sources** and the **mixture**. This difference has *no impact* on the bsseval evaluation performance.

All signals are stereophonic and encoded at 44.1kHz.

The data from *musdb18* is composed of several different sources:

- 100 tracks are taken from the DSD100 dataset, which is itself derived from The 'Mixing Secrets' Free Multitrack Download Library. Please refer to this original

resource for any question regarding your rights on your use of the DSD100 data.
- 46 tracks are taken from the MedleyDB licensed under Creative Commons (BY-NC-SA 4.0).
- 2 tracks were kindly provided by Native Instruments originally part of their stems pack.
- 2 tracks are from the Canadian rock band The Easton Ellises as part of the heise stems remix competition, licensed under Creative Commons (BY-NC-SA 3.0).

Have a look at the detailed list of all tracks.

Download

⚠
The dataset is hosted on Zenodo and requires that users request access, since the tracks can only be used for academic purposes. We **manually** check this requests. Please *do not* fill the form multiple times, it usually takes as less than a day to give you access.

- Download Full Dataset (4.4 Gb) (MD5: af06762477334799bfc5abf237648207)
- Sample Data (11 MB)

When the download is done, you can use the following tools to use the stems-encoded musdb in your scripts:

Associated Tools

Parsers

- musdb: Python based dataset parser
- mus-io: Docker scripts for decoding/encoding STEMS <=> wav (i.e. MATLAB users go there)
- musdb.jl: Julia based dataset parser

Evaluation

- museval: BSSEval v4 Evaluation tools
- SiSEC 2018: Signal Separation Evaluation Challenge 2018

Further Tools

- cutlist-generator: Scripts to generate 30s and 7s excerpt annotations from the full dataset based on the activity of all sources.
- preview-generator: Scripts to cut and recode the dataset based on provided cutlists.

Oracle Methods

- oracle: Python based oracle method implementation like Ideal Binary Mask, Softmasks, Multichannel Wienerfilter

SiSEC 2018 Evaluation Campaign

- SiSEC 2018: Submissions of raw scores
- SiSEC 2018 - Analysis: Analysis of 2018 Submissions
- Paper: all results, to be published at International Conference on Latent Variable Analysis and Signal Separation.

Acknowledgements

We would like to thank Mike Senior, Rachel Bittner, and also Mickael Le Goff, not only for giving us the permission to use this multitrack material, but also for maintaining such resources for the audio community.

Citation

If you use this dataset, please reference it accordingly:

```
@misc{musdb18,
  author       = {Rafii, Zafar and
                  Liutkus, Antoine and
                  Fabian-Robert St{\"o}ter and
                  Mimilakis, Stylianos Ioannis and
                  Bittner, Rachel},
  title        = {The {MUSDB18} corpus for music separation},
  month        = dec,
  year         = 2017,
  doi          = {10.5281/zenodo.1117372},
  url          = {https://doi.org/10.5281/zenodo.1117372}
}
```