UNMIXDB: A DATASET FOR DJ-MIX INFORMATION RETRIEVAL

Diemo Schwarz **Dominique Fourer**
Ircam Lab, CNRS, Sorbonne Université, Ministère de la Culture
F-75004 Paris, France

ABSTRACT

DJ techniques are an important part of popular music culture but are also not sufficiently investigated by researchers due to the lack of annotated datasets of DJ mixes. This paper aims at filling this gap by introducing a publicly available DJ-mix dataset. It contains automatically generated beat-synchronous mixes based on freely available music tracks, and the ground truth about the placement of tracks in a mix. Each mix is generated in several variants with different effects and time-stretching methods applied. Possible applications are to test novel methods for track identification in mixes, or for automatic annotation and deconstruction of recorded mixes for which the constituent tracks are known.

1. INTRODUCTION

This freely available dataset [1] offers a missing brick in a larger research agenda to understand DJ practices—an important part of popular music culture. The outcomes from such an understanding are many, for instance musicological research in popular music, cultural studies on DJ practice and reception, music technology for computer support of DJing, automation of DJ mixing for entertainment or commercial purposes. So far, DJ techniques are not very well researched for the lack of annotated databases of DJ mixes.

Possible uses of the dataset are the evaluation of methods for the identification of tracks for monitoring DJ mixes [3], or the precise annotation or even reverse engineering of DJ-mixes when the constituent tracks are available [1]. In the latter project, we perform alignment to determine the exact offset of each track in the mix, and then estimate the cue points [2] and volume fade curves, in order to learn about the decisions a DJ makes when creating a mix.

The *UnmixDB* dataset is based on the curatorial work of Sonnleitner et. al. [3], for identification of the tracks within human-made DJ mixes by fingerprinting. They collected Creative-Commons licensed source tracks of 10 free dance music mixes from the *Mixotic* netlabel. [2] Their dataset [3] provides the mixes, the full tracks, and the ground-truth playlists with hand-annotated time points from which only the next track in the playlist is present in the mix. Unfortunately, this does not give information about the start point of the track in the mix, and is not accurate enough for our aims of DJ mix analysis, let alone reverse engineering.

We used their collected full tracks to produce our track excerpts, but regenerated beat-synchronous and thus "ecologically valid" artificial mixes with perfectly accurate ground truth (see figure 1).

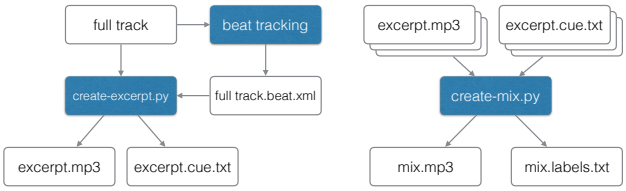

Figure 1. Data flow and file types of the *UnmixDB* dataset.

2. THE *UNMIXDB* DATASET

In order to evaluate the DJ mix tagging, analysis, and reverse engineering methods, we created a dataset of excerpts of open licensed dance tracks and automatically generated mixes based on these. We use track excerpts because of the runtime and memory requirements, especially for methods such as DTW, which is of quadratic memory complexity. We could also not have scaled the dataset up to the many playlists and variants when using full tracks.

Each track excerpt contains about 20s of the beginning and 20s of the end of the full source track (not included in the dataset, but available from [3]). However, the exact choice is made taking into account the metric structure of the track. The cue-in region, where the fade-in will happen, is placed on the second beat marker starting a new measure (as analysed by the beat tracker IRCAMBEAT), and lasts for 4 measures. The cue-out region ends with the 2nd to last measure marker. We assure at least 20s for the beginning and end parts by extending them accordingly. The cut points where they are spliced together is again placed on the start of a measure, such that no artefacts due to beat discontinuity are introduced.

Each mix is based on a playlist that mixes 3 track excerpts beat-synchronously, such that the middle track is

1 . http://zenodo.org/record/1422385 DOI 10.5281/zenodo.1422385
2 . http://www.mixotic.net
3 . http://www.cp.jku.at/datasets/fingerprinting

embedded in a realistic context of beat-aligned linear cross fading to the other tracks. The first track's BPM is used as the seed tempo onto which the other tracks are adapted.

Each playlist of 3 tracks is mixed 12 times with combinations of 4 variants of effects and 3 variants of time scaling using the treatments of the *sox* open source command-line program. [4] The 4 effects are:

none: no effect

bass: +6 dB bass boost using a low-shelving biquad filter below 100 Hz

compressor: heavy dynamics compression (ratio of 3:1 above -60 dB, -5 dB makeup gain)

distortion: heavy saturation with +20 dB gain

These effects were chosen to cover treatments likely to be applied to a DJ set (EQ, compression), and also to introduce non-linear treatments (distortion) to test the limits of MIR methods.

The 3 timescale variants are:

none: no time scaling, ie. the tracks are only aligned on the first beat in the cue region and then drift apart

resample: linked time and pitch scaling by resampling (sox *speed* effect)

stretch: time stretching while keeping the pitch (sox *tempo* effect using WSOLA)

These 3 variants allow to test simple alignment methods not taking into account time scaling, and allow to evaluate the influence of different algorithms and implementations of time scaling.

The dataset is organised in 6 individually downloadable sets of tracks and mixes, between 500 MB and 1 GB in size, for a total of 4 GB. Table 1 shows more details about the content one of the sets. In the near future, the dataset could be extended by more songs, more mixes, and mixes of the full source tracks. We also publish the Python source code [5] to generate the excerpts and mixes, such that other researchers can create test data from other track collections or in other variants.

2.1 File Formats

The *UnmixDB* dataset contains the ground truth for the source tracks and mixes in `.labels.txt` files with tab-separated columns *starttime, endtime, label*. For each mix, the start, end, and cue points of the constituent tracks are given, along with their BPM and speed factors. We use the convention that the label starts with a number indicating which of the 3 source tracks the label refers to.

The song excerpts are accompanied by their cue region and tempo information in `.txt` files in table format.

Additionally, we provide the `.beat.xml` files containing the beat tracking results for the full tracks available from [3].

Number of tracks	37
Number of playlists	37
Number of tracks per playlist	3
Number of variants per playlist	12
Number of mixes	444
Average duration of tracks [s]	46
Average duration of mixes [s]	107
Total duration of tracks [min]	1016
Total duration of mixes [min]	2743
Median tempo of tracks [bpm]	128
Minimum tempo of tracks [bpm]	67
Maximum tempo of tracks [bpm]	140

Table 1. Statistics of set 123 of the *UnmixDB* dataset v1.

3. CONCLUSION

The presented work is a first step towards providing the missing link in a chain of methods that allow the retrieval of rich data from existing DJ mixes and their source tracks.

4. ACKNOWLEDGMENTS

This work is part of the ABC_DJ project [6], which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 688122.

5. REFERENCES

[1] Diemo Schwarz and Dominique Fourer. Towards Extraction of Ground Truth Data from DJ Mixes. In *Proc. ISMIR*, Suzhou, China, 2017.

[2] Diemo Schwarz, Daniel Artur Schindler, and Severino Spadavecchia. A heuristic algorithm for DJ cue point estimation. In *Sound and Music Computing (SMC)*, Limassol, Cyprus, July 2018.

[3] Reinhard Sonnleitner, Andreas Arzt, and Gerhard Widmer. Landmark-based Audio Fingerprinting for DJ Mix Monitoring. In *Proc. ISMIR*, New York, NY, 2016.

4 . http://sox.sourceforge.net

5 . http://github.com/Ircam-RnD/unmixdb-creation

6 . http://abcdj.eu