BREAK-INFORMED AUDIO DECOMPOSITION FOR INTERACTIVE REDRUMMING

Patricio López-Serrano[1] **Matthew E. P. Davies**[2] **Jason Hockman**[3]
Christian Dittmar[1] **Meinard Müller**[1]
[1] International Audio Laboratories Erlangen, Germany
[2] INESC TEC, Sound and Music Computing Group, Portugal
[3] DMT Lab, Birmingham City University, UK
patricio.lopez.serrano@audiolabs-erlangen.de

ABSTRACT

Redrumming or *drum replacement* is used to substitute or enhance the drum hits in a song with one-shot drum sounds obtained from an external collection or database. In an ideal setting, this is done on multitrack audio, where one or more tracks are dedicated exclusively to drums and percussion. However, most non-professional producers and DJs only have access to mono or stereo downmixes of the music they work with. Motivated by this scenario, as well as previous work on decomposition techniques for audio signals, we propose a step towards enabling full-fledged redrumming with mono downmixes.

1. PROPOSED METHOD

Figure 1 gives an overview of our method for break-informed redrumming, inspired by [5]. The figure consists of two sides: the left side contains the track that will be redrummed, and the right side contains the track that will provide new timbral information. We describe each side in the following.

As input for the track that will be redrummed, we need the monaural (mono) downmix of a song that contains a drum break, as well as a segmentation indicating the location of the drum break. A method for automatically finding percussion-only regions in digital music recordings is described in [4]. Since we are working with mono input data, we need a way to extract multiple drum kit tracks from the mixture. At this stage we use the drum source separation (DSS) method by Dittmar and Müller [1] (which is based on non-negative matrix factor deconvolution, NMFD) to learn timbral templates for the drum kit elements present in the break.

Following the assumption that the drum timbre remains largely unchanged throughout the track, we fix the drum templates and learn a further non-negative matrix factorization (NMF) for the entire track, following [2]. In principle, this is equivalent to NMF-based harmonic-percussive source separation (HPSS), such as recently proposed by [3]. We now have spectral information for all the non-percussive instruments (or the *harmonic* part, shown as horizontal orange bars), which includes lead melody as well as accompaniment instrumentation. Our NMF model has also learned activations for the fixed drum templates—these activations are shown as light purple curves for kick drum (KD), snare drum (SD), and hi-hat (HH), and will be used as insertion positions for the redrumming.

The right side of Figure 1 shows a second song containing a drum break, along with an indication of where the break is. This drum break will provide the timbral information for the redrum, with sounds being "copied" onto appropriate timepoints in the left-hand track. Again, we use NMFD-based DSS [1] to learn the timbral properties of individual drum hits in the break, in order to combine these spectral templates with the activations found in the left-hand track. We encourage listening to audio examples of our results at the accompanying website. [1]

2. REAL-WORLD SCENARIO

Fully automatic redrumming is a difficult task. Especially when working with audio mixtures, the main challenge lies in avoiding crosstalk—not only between the harmonic and percussive parts, but also among the individual drum kit parts themselves. In a studio setting, redrumming is usually done by selecting (one-shot) drum hit sounds from a collection or database—in this contribution, instead of having a ready-made drum sound database, we construct this collection from a drum break of the user's choosing. Thus, for instance, we can imagine that a user selects James Brown's "Funky Drummer" as a track to be redrummed, using the sounds (or timbral properties) from "Amen, Brother" by The Winstons.

Redrumming is also usually done manually or semi-manually, using a DAW such as `Logic Pro` or `Cubase`, or a specialized plugin, like `Drumagog` or `Superior Drummer`. Assuming that a user wishes to substitute KD

[1] `https://www.audiolabs-erlangen.de/resources/MIR/2018-ISMIR-LBD-Redrum`

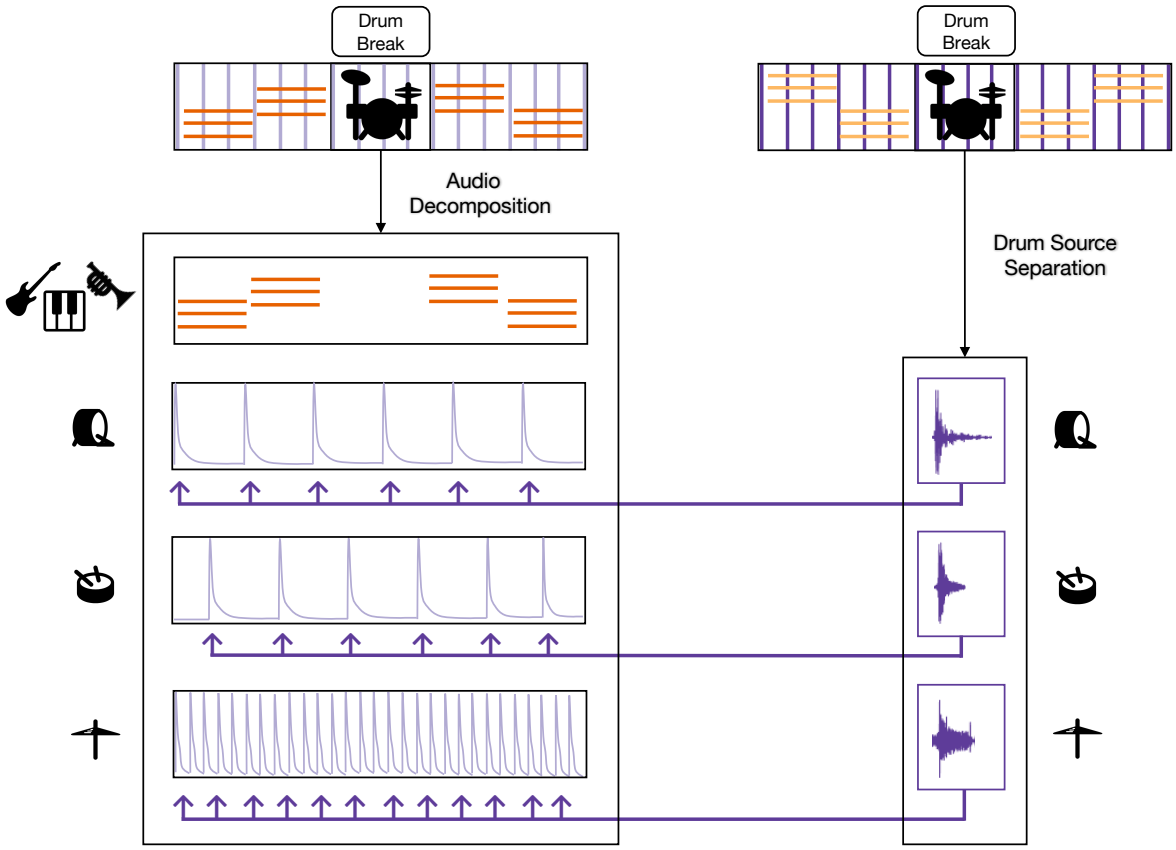

Figure 1. Overview of our break-informed redrumming method. For a detailed explanation, see Section 1.

hits, the procedure would be to step through all the transients in the drum track, placing the cursor at the beginning of each KD hit. Then, the user would select a sound from the library and use it to either replace the current KD hit, or place it in a new track to enhance the timbral properties of the existing one. An advantage of our system is that all hits of a certain type (e.g., all KD hits) are tracked simultaneously by our decomposition model, enabling a once-through automatic substitution.

Furthermore, in a professional setting, musicians often request a certain *sound* or *aesthetic* for the mixing process, wishing to sound like other artists that they admire. Thus, another advantage of our method is that users can directly achieve this effect by selecting a drum break recording with their desired properties.

3. ACKNOWLEDGMENTS

Project TEC4Growth - Pervasive Intelligence, Enhancers and Proofs of Concept with Industrial Impact/NORTE-01-0145-FEDER-000020, financed by the North Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, and through the European Regional Development Fund (ERDF). MD is supported by national funds through the FCT - Foundation for Science and Technology, I.P., under the project IF/01566/2015. PLS is supported in part by a scholarship from CONACYT-DAAD. CD and MM are supported by the German Research Foundation (DFG-MU 2686/10-1). The International Audio Laboratories Erlangen are a joint institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Fraunhofer Institute for Integrated Circuits IIS.

4. REFERENCES

[1] Christian Dittmar and Meinard Müller. Reverse engineering the amen break – score-informed separation and restoration applied to drum recordings. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(9):1531–1543, 2016.

[2] Minje Kim, Jiho Yoo, Kyeongok Kang, and Seungjin Choi. Nonnegative matrix partial co-factorization for spectral and temporal drum source separation. *IEEE Journal of Selected Topics Signal Processing*, 5(6):1192–1204, 2011.

[3] Clément Laroche, Matthieu Kowalski, Hélène Papadopoulos, and Gaël Richard. Hybrid projective nonnegative matrix factorization with drum dictionaries for harmonic/percussive source separation. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 26(9):1499–1511, 2018.

[4] Patricio López-Serrano, Christian Dittmar, and Meinard Müller. Finding drum breaks in digital music recordings. In *Proceedings of the International Symposium on Computer Music Multidisciplinary Research (CMMR)*, pages 68–79, Porto, Portugal, September 2017.

[5] Tomohiko Nakamura, Hirokazu Kameoka, Kazuyoshi Yoshii, and Masataka Goto. Timbre replacement of harmonic and drum components for music audio signals. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7520–7524, Florence, Italy, May 2014.