

Unsupervised Audio Source Separation Using Differentiable Parametric Source Models

Kilian Schulze-Forster¹, Clément S. J. Doire², Gaël Richard¹, and Roland Badeau¹

¹ LTCI, Télécom Paris, Institut Polytechnique de Paris ² Sonos Inc.

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

In partnership with



Abstract

Supervised deep learning approaches to underdetermined audio source separation achieve state-of-the-art performance but require a dataset of mixtures along with their corresponding isolated source signals. Such datasets can be extremely costly to obtain for musical mixtures. This raises a need for unsupervised methods. We propose a novel unsupervised model-based deep learning approach to musical source separation. Each source is modelled with a differentiable parametric source-filter model. A neural network is trained to reconstruct the observed mixture as a sum of the individual source signals by estimating the source models' parameters given their fundamental frequencies. At test time, soft masks are obtained from the source signals synthesized by the source models. The experimental evaluation on a vocal ensemble separation task shows that the proposed method outperforms learning-free methods based on nonnegative matrix factorization and a supervised deep learning baseline. Integrating domain knowledge in the form of source models into a data-driven method leads to high data efficiency: the proposed approach achieves good separation quality even when trained on less than three minutes of audio. This work makes powerful deep learning based separation usable in domains where training data with isolated source signals is expensive or nonexistent.

This abstract and the presentation given at the workshop are based on the paper "Unsupervised Audio Source Separation Using Differentiable Parametric Source Models" which is currently under review.