Skysong Innovations
Advancing the Arizona State University Knowledge Enterprise

1475 N. Scottsdale Road, Suite 200
Scottsdale, AZ 85287-3538
Phone: 480 884 1996 Fax: 480 884 1984

Case ID:M21-063P^
Published: 10/18/2021

Inventors

Vivek Sivaraman Narayanaswamy
Jayaraman Thiagarajan
Rushil Anirudh
Andreas Spanias

Contact

Shen Yan
shen.yan@skysonginnovations.com

Unsupervised Audio Source Separation Using Generative Priors

-Background Audio source separation, the process of recovering constituent source signals from a given audio mixture, is a key component in downstream applications such as audio enhancement and music information retrieval. Modern under-determined audio source separation systems rely on supervised training of carefully tailored neural network architectures operating either in the time or the spectral domain. However, these methods are severely limited by their requiring of expensive source level labeled data and being specific to a given set of sources and the mixing process, which demands complete re-training when assumptions change. This strongly emphasizes the need for unsupervised methods that can leverage the recent advances in data-driven modeling, and compensate for the lack of labeled data through meaningful priors. Invention Description Researchers at Arizona State University have developed a novel approach for audio source separation based on generative priors trained on individual sources. Through the use of projected gradient descent optimization, the approach simultaneously searches in the source-specific latent spaces to effectively recover the constituent sources. Though the generative priors can be defined in the time domain, using spectral domain loss functions for optimization yields quality source estimates. Empirical studies on standard datasets (spoken digit, drums, and piano) clearly demonstrate the effectiveness of this approach over classical as well as state-of-the-art unsupervised baselines. Potential Applications • Audio source separation • Unsupervised learning • Audio enhancement Benefits and Advantages • Complete inference-time technique is well suited for under-determined source separation • Efficiently handles varying number of known sources in a given mixture • Unlike supervised approaches, does not require re-training or extensive fine-tuning Related Publication: Unsupervised Audio Source Separation using Generative Priors (PDF)Homepage of Professor Andreas Spanias