Music Separation Enhancement with Generative Modeling

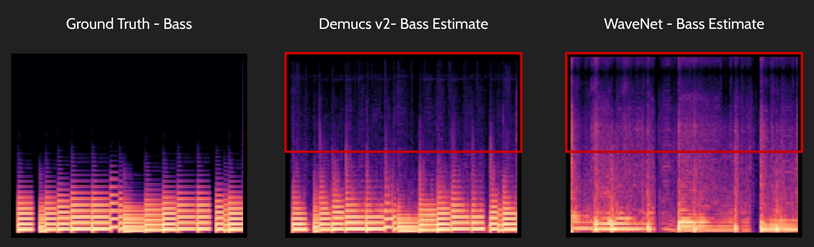
Noah Schaffer [#*], Boaz Cogan [#*], Ethan Manilow [#], Max Morrison [#], Prem Seetharaman [♭], Bryan Pardo [#]

* Equal Contribution

Northwestern University

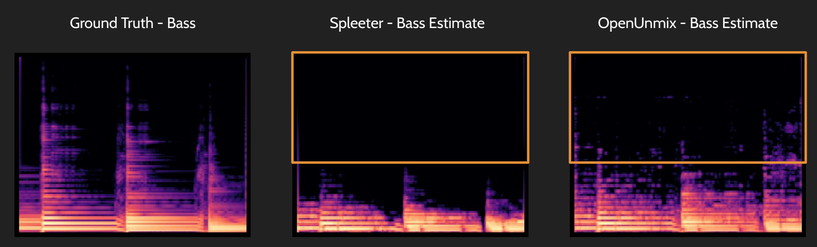
♭ Descript, Inc.

SotA Separators Still Have Issues!!!



Added noise!

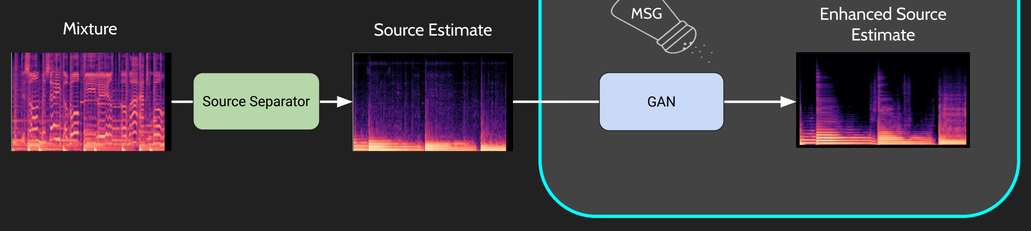
Waveform separators tend to add high frequency noise. Spectrogram separators tend to remove upper harmonics



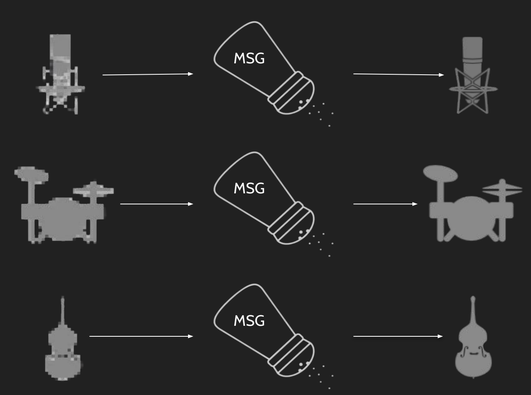
Missing upper harmonics!

How do we fix these issues???

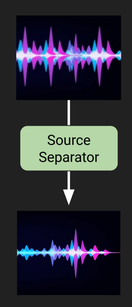
Leverage advances in generative modeling to create a post-processor for source separation (MSG)



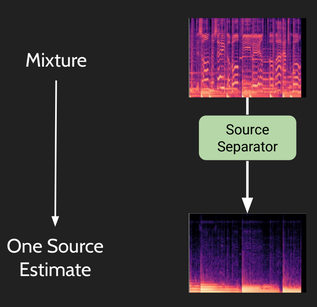
Train three source-specific MSG models for the "vocals", "drums", and "bass" sources of MUSDB18



Waveform Separators



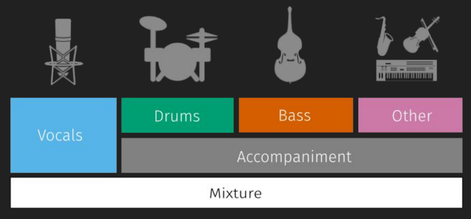
Spectrogram Separators



MSG enhances the output of **both** waveform separators **and** spectrogram separators

Experimental Setup

Dataset



Use the MUSDB18 test to evaluate the performance of each source-specific MSG model

Separation Models

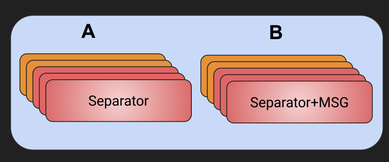
Spectrogram: Spleeter, OpenUnmix

Waveform: Demucs v2, WaveNet

Demucs v3 — Hybrid separator unseen during training

Generate source estimates from five separators and pass them through MSG

Subjective Evaluation



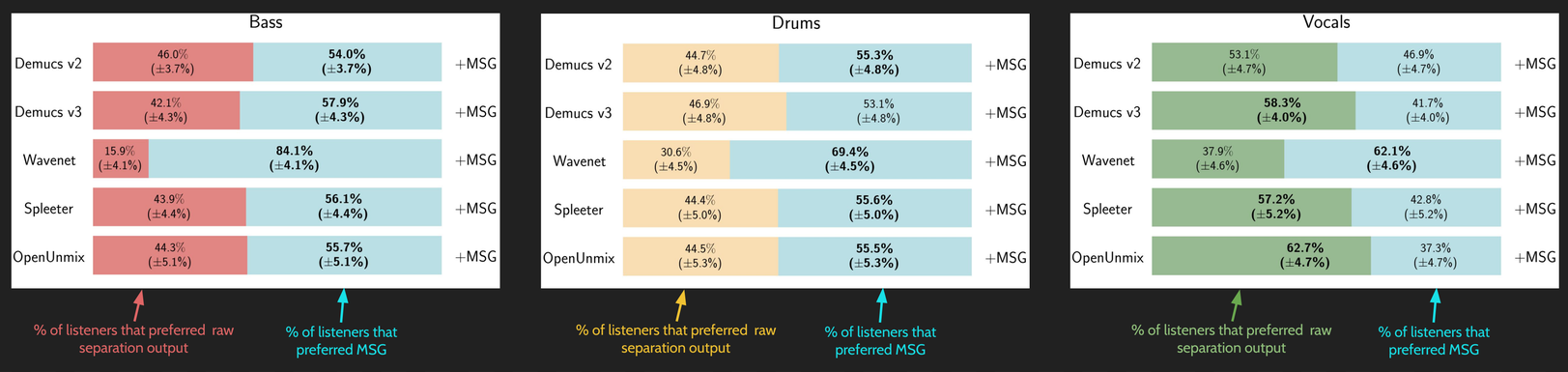
Pairwise listener study where subjects choose between raw separation output and MSG output

Artifact Analysis

A heuristic evaluation to measure artifacts heard in source separation output

Subjective Evaluation

The subjective evaluation results for the "vocals", "drums", and "bass" sources of MUSDB18. A **bolded** result indicates statistical significance

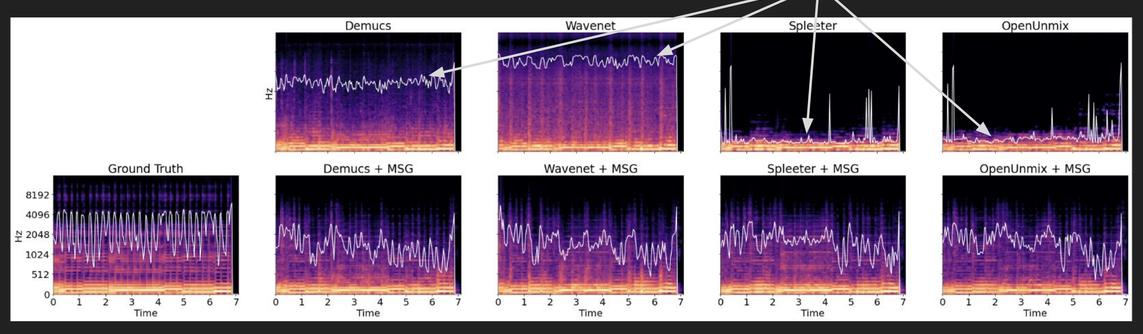


Listeners prefer MSG for all 5 separators for bass, all 5 separators for drums and 1 for vocals

In our paper, we hypothesize why MSG may struggle with vocals more than other sources

Analyzing Artifacts

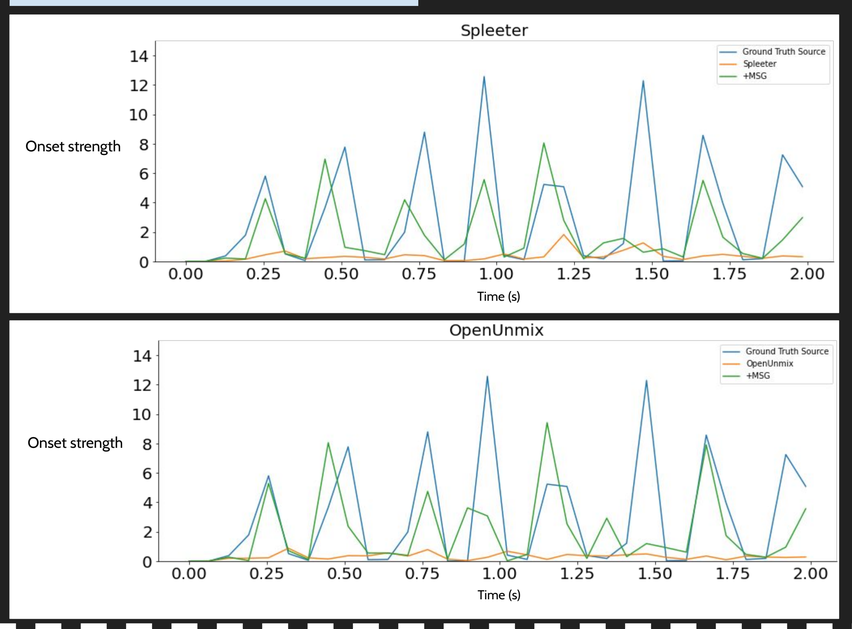
Measuring Missing Harmonics and Added Noise



Spectral rolloff at 98% is the frequency bin below which 98% of the energy of a signal lies

Above, we see that MSG removes noise from waveform separators and recovers missing harmonics from spectrogram separators. **Both** improvements are captured by spectral rolloff!

Measuring Blurred Transients



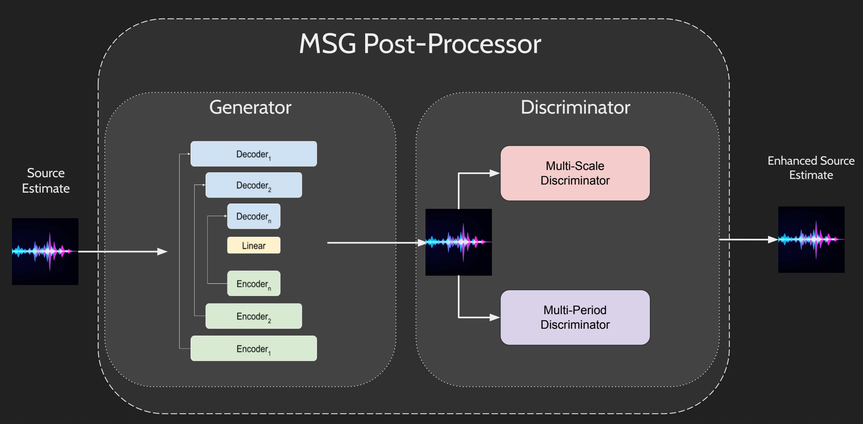
Spectrogram separators struggle with transient recovery. We see that Spleeter and OpenUnmix have a considerably weaker onset strength values than ground truth

MSG improves onset strength in spectrogram separators and is able to recover blurred transients!

Conclusion

State-of-the-art source separation models contain audible errors in its output

MSG improves these errors in both waveform and spectrogram separators !



See and listen to more examples at out project page: https://interactiveaudiolab.github.io/project/msg