

Upsampling Artifacts in Music Source Separation

Jordi Pons¹, Santiago Pascual¹, Giulio Cengarle¹, and Joan Serra¹

Arxiv DOI:

<https://arxiv.org/abs/2010.14356> ¹ Dolby Laboratories

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

A number of recent advances in neural audio synthesis rely on upsampling layers, which can introduce undesired artifacts. In computer vision, upsampling artifacts have been studied and are known as checkerboard artifacts (due to their characteristic visual pattern). However, their effect has been overlooked so far in audio processing. Here, we address this gap by studying this problem from the audio signal processing perspective. We first show that the main sources of upsampling artifacts are: (i) the tonal and filtering artifacts introduced by problematic upsampling operators, and (ii) the spectral replicas that emerge while upsampling. We then compare different upsampling layers, showing that nearest neighbor upsamplers can be an alternative to the problematic (but state-of-the-art) transposed and subpixel convolutions which are prone to introduce tonal artifacts.

References

Link to the original ICASSP 2021 publication: <https://ieeexplore.ieee.org/document/9414913>

Link to a preprint on arXiv: <https://arxiv.org/abs/2010.14356>