














The need for causal, low-latency sound demixing and remixing to improve accessibility

Gerardo Roa Dabike ¹¶, Michael A. Akeroyd ², Scott Bannister ³, Jon Barker ⁴, Trevor J. Cox ¹, Bruno Fazenda ¹, Jennifer Firth ², Simone Graetzer ¹, Alinka Greasley ³, Rebecca Vos ¹, and William Whitmer ²

¹ Acoustics Research Centre, University of Salford, UK ² Hearing Sciences, School of Medicine, University of Nottingham, UK ³ School of Music, University of Leeds, UK ⁴ Department of Computer Science, University of Sheffield, UK ¶ Corresponding author

Published: 04 November 2023

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/)).

Abstract

Soundtracks and music sound different to those with hearing loss. Dialogue and lyrics can be difficult to pick-out and the loss of high frequencies make sounds less impactful. The most common treatment for hearing loss is hearing aids. But about 40% of people who could benefit from hearing aids do not use them often enough citing poor perceived performance. Being able to demix sounds on a hearing aid would allow a personalised rebalancing of sounds in a remix to improve audio quality. For soundtracks, you might amplify parts that are important to the narrative, such as the dialogue and key sound effects [Shirley et al. (2019)](Ward et al., 2019). For music, the requirements for remixing are still being researched but it might include amplifying the lyrics so these are easier to hear. A challenge is that hearing loss is heterogenous. Furthermore, the tolerance of hearing aid users to distortion artefacts from processing varies. Consequently, it is important any remix is personalised. The Cadenza project is running a series of machine learning challenges to improve the processing of music for those with a hearing loss. Our ICASSP 2024 grand demixing/remix challenge encourages causal and non-causal approaches. Non-causal musical source separation with deep neural networks has achieved impressive results (Mitsufuji et al., 2022) for separating pop/rock into vocals, bass, drums, and other (Rafii et al., 2019). For hearing aids and live music, a low-latency, causal approaches are needed. Furthermore, hearing aid processing power is limited, which is challenging for very large models. The paper will outline the opportunities and challenges of using source separation to improve accessibility.

Mitsufuji, Y., Fabbro, G., Uhlich, S., Stöter, F.-R., Défossez, A., Kim, M., Choi, W., Yu, C.-Y., & Cheuk, K.-W. (2022). Music demixing challenge 2021. *Frontiers in Signal Processing*, 1, 808395.

Rafii, Z., Liutkus, A., Stöter, F.-R., Mimitakis, S. I., & Bittner, R. (2019). *MUSDB18-HQ - an uncompressed version of MUSDB18*. <https://doi.org/10.5281/zenodo.3338373>

Shirley, B., Ward, L., & Chourdakis, E. T. (2019). *Personalization of object-based audio for accessibility using narrative importance*.

Ward, L., Paradis, M., Shirley, B., Russon, L., Moore, R., & Davies, R. (2019). Casualty accessible and enhanced (a&e) audio: Trialling object-based accessible TV audio. *Audio Engineering Society Convention 147*.