

Compte rendu approximatif RDV 2

Super ressource : <https://source-separation.github.io/tutorial/landing.html>

Notes globales :

Risque d'usine à gaz du modèle. Pas sûr pour CLAP mais on peut tenter : "separate what you describe".

Grosse différence entre mélanges acoustiques et artificiels : le rayonnement des instruments est un facteur important qu'on ne retrouve pas avec des haut-parleurs. Les modèles actuels font énormément d'hypothèses simplificatrices.

Enregistrement :

Énormément de logistique, ne pas dépasser 8 micros. Préparer plusieurs jours en avance, **CHECKLIST**. Cartes sons, micros, logiciels, instruments, matériel additionnel, s'assurer que tout marche en avance.

Méthodes spatiales :

- Beamforming : Rendu musical moyen, OK en entrée d'un modèle, mais les données d'acoustique des salles sont aussi pertinentes. Faire gaffe au nombre de micros.
- Méthodes HR de localisation : MNMF (cours d'aujourd'hui).

IA :

- SpeakerBeam : https://www.isca-archive.org/interspeech_2019/ochiai19_interspeech.html : séparation de voix supervisée
- Mixit : <https://arxiv.org/pdf/2006.12701> : Séparation non supervisée

Datasets :

- MusDB : <https://sigsep.github.io/datasets/musdb.html>
- Open-end Mix : pas trouvé

D'autres datasets : <https://source-separation.github.io/tutorial/data/datasets.html>

Antonin :

- ★ ★ ★ A multichannel learning-based approach for sound source separation in reverberant environments : <https://link.springer.com/article/10.1186/s13636-021-00227-2>
- https://www.isca-archive.org/interspeech_2024/ahn24c_interspeech.pdf
- Music AI : <https://music.ai/research/> + <https://www.linkedin.com/pulse/machine-learning-models-behind-musica-is-stem-service-scott-josephson-d5ttf>