

学内研究組織（2024 年度）

研究課題：「自動車室内音場へ適用可能なシームレス音像定位モデル」

反射音が支配的な音場を考慮した音像方向知覚の深層学習と XAI を用いた分析

Deep learning for perceived sound image direction considering sound fields dominated
by reflected sounds and analysis using XAI

阪本 浩二 (Sakamoto Koji)

はじめに

音像定位に関する両耳聴モデルはロボット聴覚の重要な一機能であり、対話や聴覚支援、災害救助、音環境デザイン等の様々な場面において幅広く有用とされている。従来の音像定位モデルの多くは主に自由音場で構築されており、自動車内の音源や倒壊した建物内からの音源のように反射音が支配的な音場においては適用が難しかった。そこで、著者は反射音や騒音の外乱が支配的な環境においても音像定位可能なモデルへ性能改善してきたが、推定精度および適用範囲に改善の余地があった。

そのため、本研究ではまず音像定位の一要素である音源方向知覚について、深層学習を用いてモデルの性能改善を行うとともに、XAI (Explainable Artificial Intelligence) を用いて説明可能性と推定精度向上の両立に向け取り組んだ。また、複数音源による合成音像や、その音像方向が狭く明瞭に知覚するか拡がって知覚するかも重要な要素となるため、合成音像および音像の拡がり感についても基礎検討を行った。

モデル学習方法

自由音場や室内音場、車内音場を想定し、聴取者周囲からの音源聴取のシミュレーション環境を構築した。そして、音源、音場の条件が異なる網羅的な学習データを作成し、畳込みニューラルネットワーク (Convolutional Neural Network, CNN) により分類学習した。

モデル学習結果

CNN 学習の結果、十分に狭いとされる分類区間に対し 84%程度の正解率が得られた。また、XAI を用いて学習したモデルの分析を行った結果、先行研究により得られた特徴量と類似の傾向が見られることを確認した [1, 2]。また、合成音像や拡がり感を評価するために、立体音響システムを構築し [3]、聴取者周囲への立体的な合成音像定位や立体的な拡がり感を制御できることを確認した。

今後は、合成音像を含むモデルや音像距離、拡がり感を推定可能なモデルへ拡張し、（走行）騒音環境下での検証を行う予定。

参考文献

- [1] 阪本, 車内音場を考慮した音源方向推定の CNN モデルに関する基礎検討, 日本機械学会関西支部第 99 期定時総会講演会, 2024.
- [2] 阪本, 畳込みニューラルネットワークを用いた車内音源方向知覚のモデル学習における音源特性の影響, 日本機械学会 第 34 回設計工学・システム部門講演会講演論文集, 2024.
- [3] K. Sakamoto, Basic study on the influence of 3D surround on the sense of alarm and comfortability of sound signs, Proceedings of the 19th International Conference on Innovative Computing, Information and Control, 2025.