

数理・情報のフロンティア  
2020 年度採択研究者

2021 年度 年次報告書
------------------

坂東 宜昭

産業技術総合研究所 人工知能研究センター  
研究員

統計的時空間モデルに基づく雑踏音環境マッピング

## § 1. 研究成果の概要

2021年度は、昨年度確立した雑踏音環境マッピングの枠組みを洗練化させるため、雑踏音環境の生成モデルをより簡素で見通しの良い枠組みに再構築するとともに、非線形生成モデルに基づく音源モデルの開発を行った。変分自己符号化器を用いた非線形音源モデルは、音源分離や音声強調で広く活用されているが、本モデルは音源信号を教師データとして用いた教師あり学習が必要であった。そこで本研究では、多チャンネル混合音のみから音源モデルを教師なし学習する深層フルランク空間相関分析 (Neural FCA) を開発した。Neural FCA は、教師なし学習データに対する非線形ブラインド音源分離とみなすことができる枠組みで、学習データに対する周辺尤度を最大化することで音源モデルを学習する。音響信号に対するブラインド音源分離では、周波数間の音源インデックスが揃わない周波数パーミュテーション問題が存在するが、潜在空間の独立性を最大化する枠組みを採用することで、深層生成モデルの学習とパーミュテーション問題の解決を一挙に実現した。本手法は、多チャンネル音声混合音を用いたベンチマークにおいて、従来の教師なし深層分離法やブラインド音源分離法を凌駕し、最新の教師あり多チャンネル音源分離法の一つである多チャンネル変分オートエンコーダ (MVAE) 法と同程度の性能を達成した。

### 【代表的な原著論文情報】

- 1) “Neural Full-Rank Spatial Covariance Analysis for Blind Source Separation,” IEEE Signal Processing Letters, 28, 1670-1674, 2021, Yoshiaki Bando, Kouhei Sekiguchi, Yoshiki Masuyama, Aditya Arie Nugraha, Mathieu Fontaine, Kazuyoshi Yoshii