

小型マイクロフォンアレイを使った音源分離・定位

Sound source separation and localization with small microphone array

研究分野
Department知識科学
Knowledge Science研究者
Researcher武田 龍
R. Takeda
駒谷和範
K. Komataniキーワード
Keyword対話ロボット、音源分離、音源定位、深層学習
dialogue robot, sound source separation, sound source localization, deep neural network応用分野
Application医療介護、防犯、家電、エンタテインメント
medical care, crime prevention, electrical appliances, entertainment

研究開発段階

基礎

実用化準備

応用化

背景

「百聞は一見に如かず」というように、音には一覽性や可視性がなく、記録や検索が難しいという問題があります。また音は可視光などに比べて波長が長い等の性質があることから、回折や反射による残響などの特性があります。このため、とりわけ周辺雑音や複数の音が存在する場合、特定の音を聞き分けたり、音の到来方向を検出することは単純ではありません。

概要・特徴

ロボットに装着された複数のマイクロフォンなど、小型のマイクロフォンアレイを使って、特定の音を聞き分ける音源分離や音の到来方向を検出する音源定位を開発しています。これらは人の音声を感じ、聞き分け、音声認識するような対話ロボットには必須の技術です。

技術内容

深層学習技術 (Deep Neural Network : DNN) を応用し、マイクロフォンアレイ音響信号処理の高精度化およびリアルタイム処理化などに取り組んでいます。

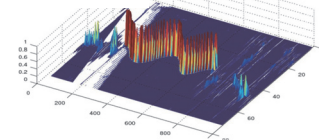
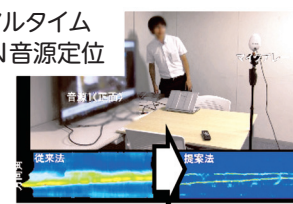
- 音源定位: 音源の到来方向を入力音響信号から検出する DNN モデル。
- 音源分離: 特定方向から到来する音声だけを入力音響信号から抽出する DNN モデル。
- データ拡張: 音の到来過程をシミュレートし、深層学習に必要な大量データを自動生成。無響室等で測定したインパルス応答を利用し、実データに近い音響特性を再現。

社会への影響・期待される効果

- ロボットに装着された複数のマイクロフォンで音源分離・音源定位を実現
- 音の到来方向の記録や可視化、話者毎の音声記録が可能

【論文 Paper】

- [1] Hokuto Munakata, Yoshiaki Bando, Ryu Takeda, Kazunori Komatani and Masaki Onishi: "Joint Separation and Localization of Moving Sound Sources Based on Neural Full-Rank Spatial Covariance Analysis," IEEE Signal Processing Letters, Vol.30, pp.384-388, April, 2023.
- [2] R. Takeda, et al.: Spatial Normalization to Reduce Positional Complexity in Direction-aided Supervised Binaural Sound Source Separation, Proc. APSIPA ASC, pp.248-253 (2021).
- [3] R. Takeda and K. Komatani: Sound Source Localization based on Deep Neural Networks with Directional Activate Function Exploiting Phase Information, Proc. IEEE-ICASSP, pp.405-409 (2016).

DNN 音源定位
スコアリアルタイム
DNN 音源定位リアルタイム
DNN 音源分離