

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

駒谷和範 武田 龍
K. Komatani R. Takeda

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

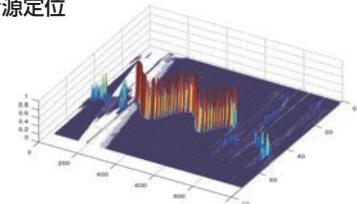
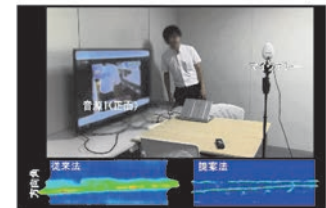
ロボットの頭部に装着された複数のマイクロフォンなど、小型のマイクロフォンアレイを使って、特定の音を聞き分ける音源分離や音の到来方向を検出する音源定位を開発しています。これは人の音声に対して応答する対話ロボットには必須の技術です。

技術内容

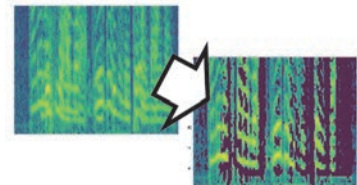
特にマイクロフォン間の位置関係や特性をキャンセルするために、無響室において測定したインパルス応答を用い、入力音響信号と方向ラベルの間の写像をDNN (Deep Neural Network) を使って学習することで、音源定位の高精度化を図っています。また、音響信号と注目する方向ラベルを入力とし、その方向に存在する信号を出力とした写像をDNN で学習させることで、高精度な分離を実現します。インパルス応答を用いて混合音をシミュレートすることで、DNN 学習に必要な大量の混合音データを作り出すことが可能です。

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

DNN音源定位
スコアリアルタイム
音源定位

DNN音源分離



【論文 Paper】

- [1] H. Munakata, R. Takeda and K. Komatani: "Training Data Generation with DOA-based Selecting and Remixing for Unsupervised Training of Deep Separation Models," Proc. of Interspeech, pp.861-865 (2022).
- [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).