

スマートグラスを用いた視聴覚・認識能力の拡張

実環境を想定した高性能・低遅延システムを開発中

日常生活におけるリアルタイム音声強調・認識・翻訳+ARインタフェース

視線/指差して話者指定・同時発話認識

重複会話の抑制



キーコンセプト:
使えば使うほど賢くなる

- 「正解」が得られない実環境で破綻せずに自律的に学習
- オンラインユーザ&環境適応

ブラインド音源分離と深層教師あり学習の融合

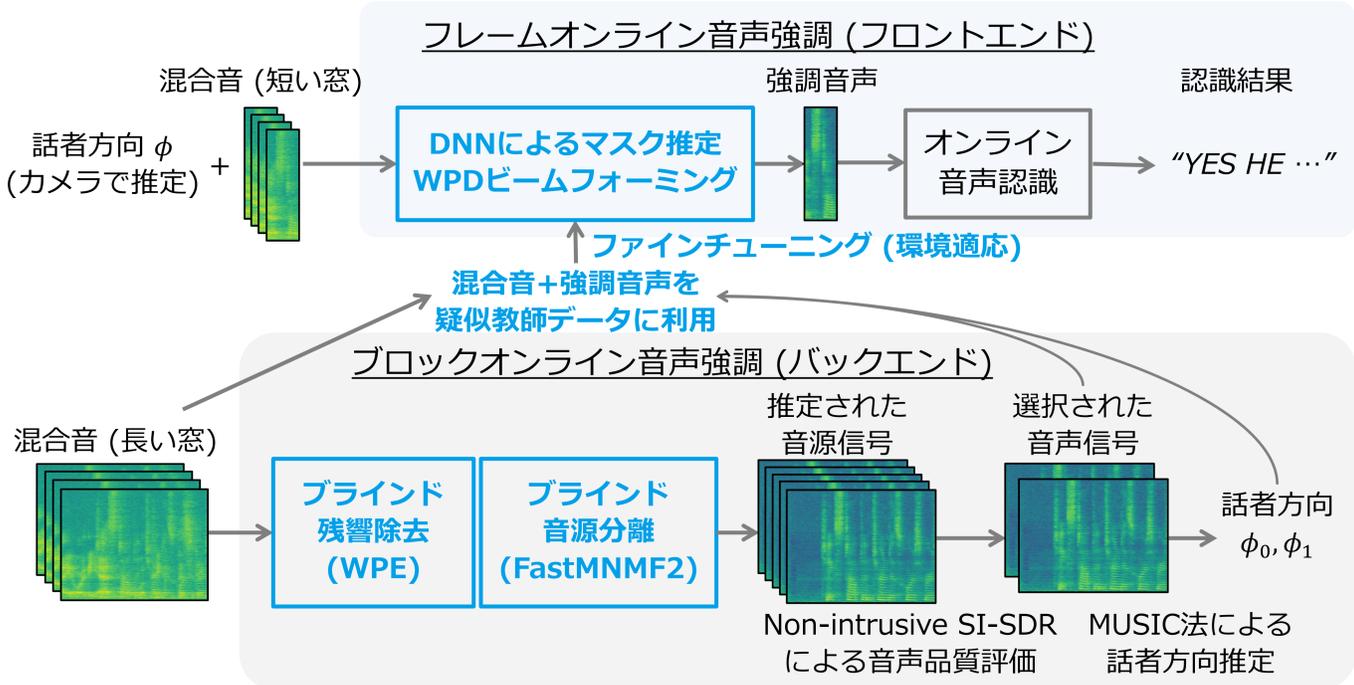
[APSIPA 2024]

未知の環境で適応的に動作する高性能かつ低遅延な音声強調 (雑音・残響除去)

同時的残響・雑音除去
 WPDビームフォーミング
 DNNで空間共分散行列を予測
 バックエンドでDNNを更新

音声認識
 RNNトランスデューサー
 ストリーミング変換 (低遅延)

実装予定の機能
 音声強調 + 音声認識の同時適応
 強化学習によるユーザ適応

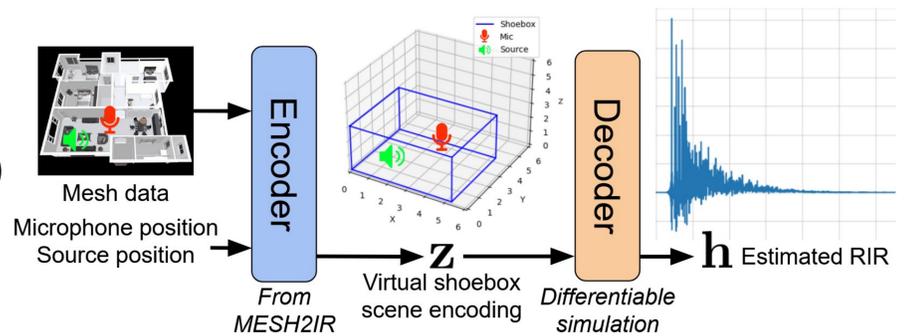


実世界における音声強調のための基盤技術の開発

部屋の3D形状からのインパルス応答推定

[Interspeech 2024]

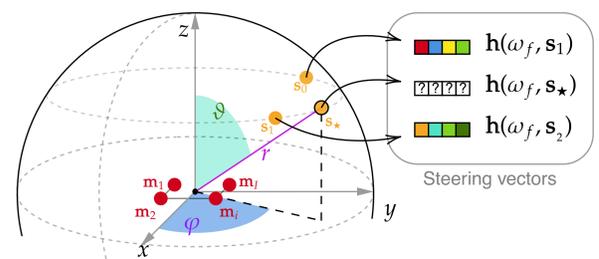
複雑な形状の部屋を直方体型の部屋で近似 (エンコーダ)
 鏡像法でインパルス応答を生成 (デコーダ)
 物理的妥当性を担保した推定が可能



関数表現に基づくステアリングベクトル推定

[HSCMA 2024]

任意の方向・周波数のベクトルを推定
 スマートグラスユーザの姿勢・向きに追従
 適応的ビームフォーミングで効果を確認



分散型マイクアレイを用いた同時的音源分離・定位

[IWAENC 2024]

室内に複数設置したマイクアレイを一挙に利用
 統一的な最尤基準に基づく音源信号・位置の同時最適化
 移動音源対応・オンライン推定も検討中

