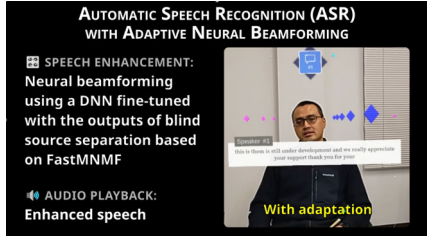


スマートグラスを用いた視聴覚・認識能力の拡張

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



開発中のARシステム (HoloLens2)

重複会話の抑制



着目したい話者に対する音声強調・認識

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

保有するブラインド音源分離技術 (FastMNMF) のアプリケーション展開

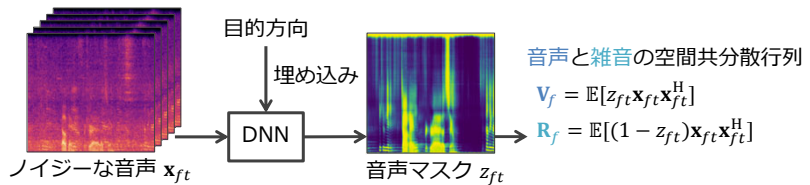
実環境中で頑健に動作する高性能かつ低遅延な音声強調・認識

音声強調：MVDRビームフォーミング

DNNで空間共分散行列を予測

音声認識：RNNトランスデューサー

DNNで音声を文字へストリーミング変換



音声強調部のオンライン適応

[IROS 2022/IWAENC 2022]

デュアルプロセスに基づく適応的音声強調

Teacher-Student学習に基づくオンライン適応

単語誤り率が12ポイント以上改善

音声強調・認識部の同時オンライン適応

[Interspeech 2022]

音声強調・認識部の同時最適化

信頼できる音声認識結果を学習データに利用

単語誤り率が10ポイント改善

高精度音声強調のための基盤技術の開発

ガウス過程に基づく時間領域音声分離

[WASPAA 2023]

音声が非定常ガウス過程に従うことを仮定

深層カーネル学習+時間領域ウィナーフィルタ

位相の整合性が大幅に改善 → 聴感が向上

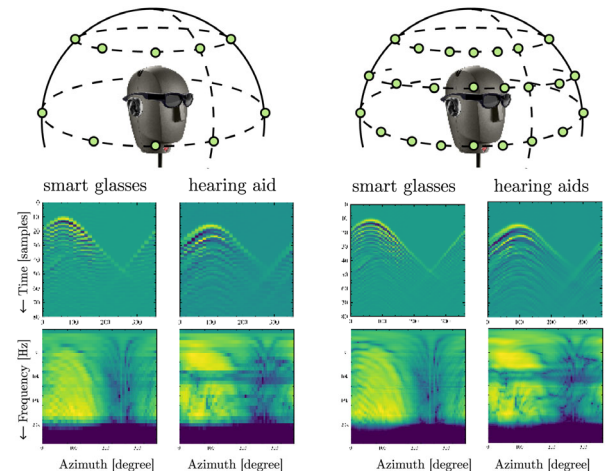
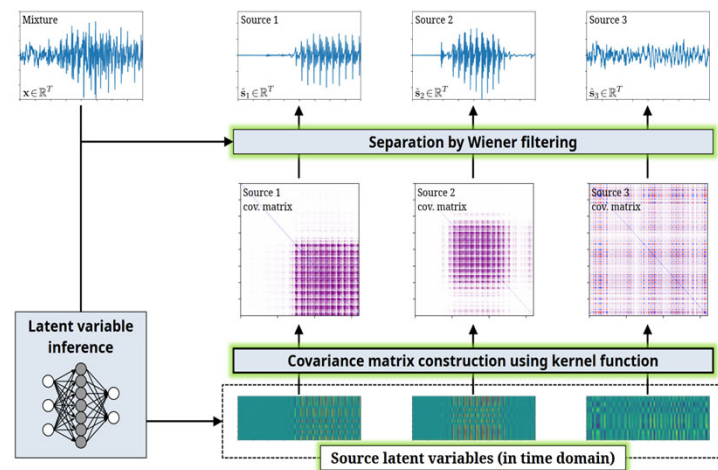
関数表現に基づくステアリングベクトル推定

[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



interpolation via Neural Steerer