

Sound Scene Understanding Team

Kazuyoshi Yoshii

音響情景理解チーム 吉井 和佳



Center for
Advanced Intelligence Project

目標：ウェアラブルデバイスを用いた視聴覚統合環境理解・支援

健常者向け：視聴覚機能の拡張

カクテルパーティにおける会話支援

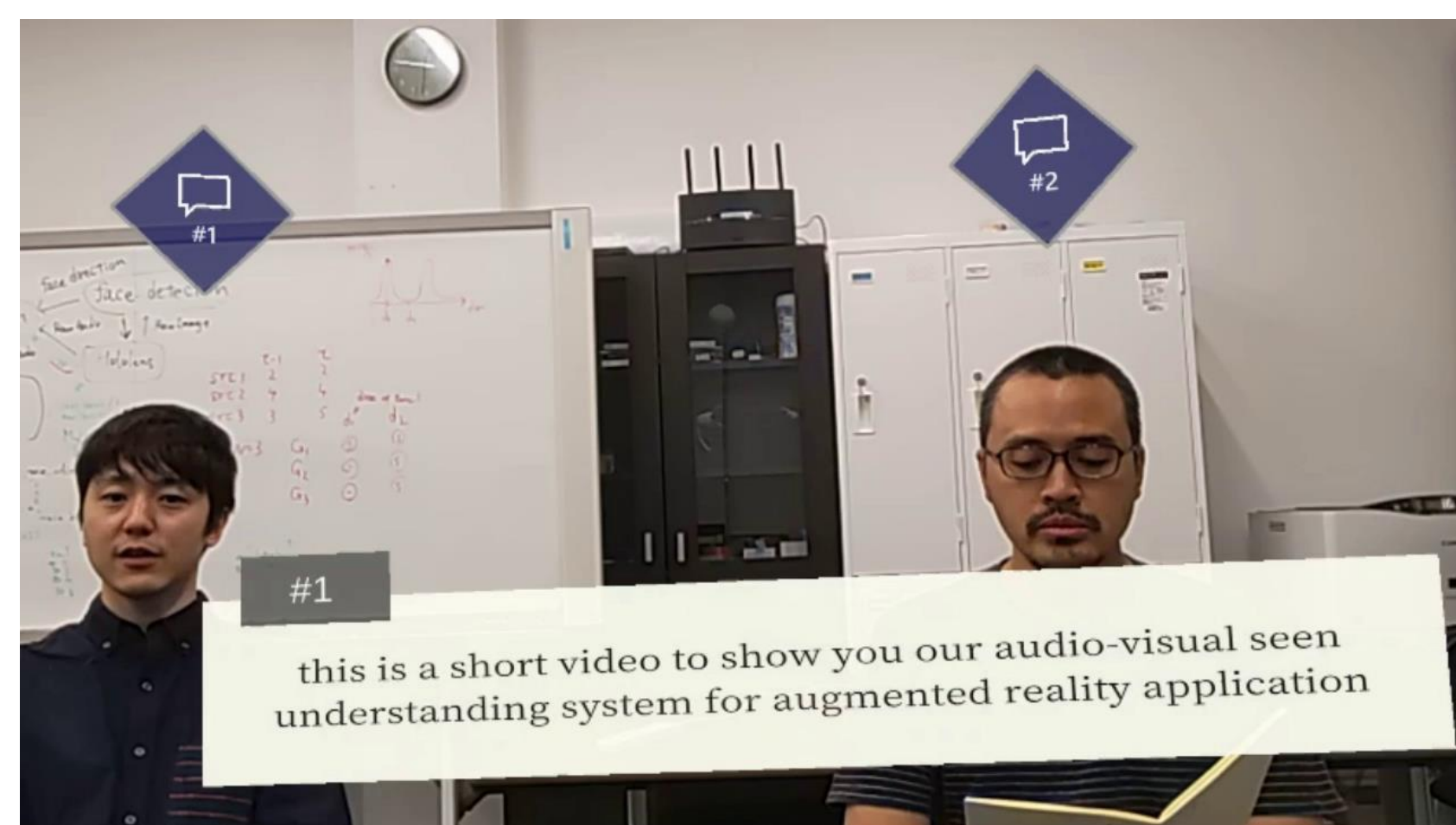
音声強調・認識・翻訳・AR表示

聴力障害者向け：“知的”健康の回復

屋内外での安心・安全な日常生活の支援

視野内外の音イベント解析・可視化

視線/指差しで話者指定・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

街案内との同時表示



音環境の可視化 (視野内)・注意喚起 (視野外)

最新の成果：実環境で動作するプロトタイプシステムを開発

基礎研究

既存の最先端手法を凌駕する汎用

ブラインド音源分離法**FastMNMF2**を提案

K. Sekiguchi et al., "Fast Multichannel Nonnegative Matrix Factorization with Directivity-Aware Jointly-Diagonalizable Spatial Covariance Matrices for Blind Source Separation", IEEE/ACM TASLP, Vol. 28, pp. 2610-2625, 2020 (特願2021-025864).

応用研究

音声強調の低遅延化・オンライン適応と

ARスマートグラス (HoloLens 2) への実装

K. Sekiguchi et al., "Direction-Aware Adaptive Online Neural Speech Enhancement with an Augmented Reality Headset in Real Noisy Conversational Environments", IEEE/RSJ IROS, submitted, 2022.

重要な成果：音声強調・雑音抑圧・残響除去を一挙に行う統計的アプローチを確立

高速・高精度なFastMNMF2をコア技術に、指向性/拡散性雑音・残響に対する頑健性を向上
最尤推定の枠組みを堅持したまま、音源・空間モデルの深層化・表現力向上を実現 (世界初)

音源モデル

空間モデル

尤度関数

