

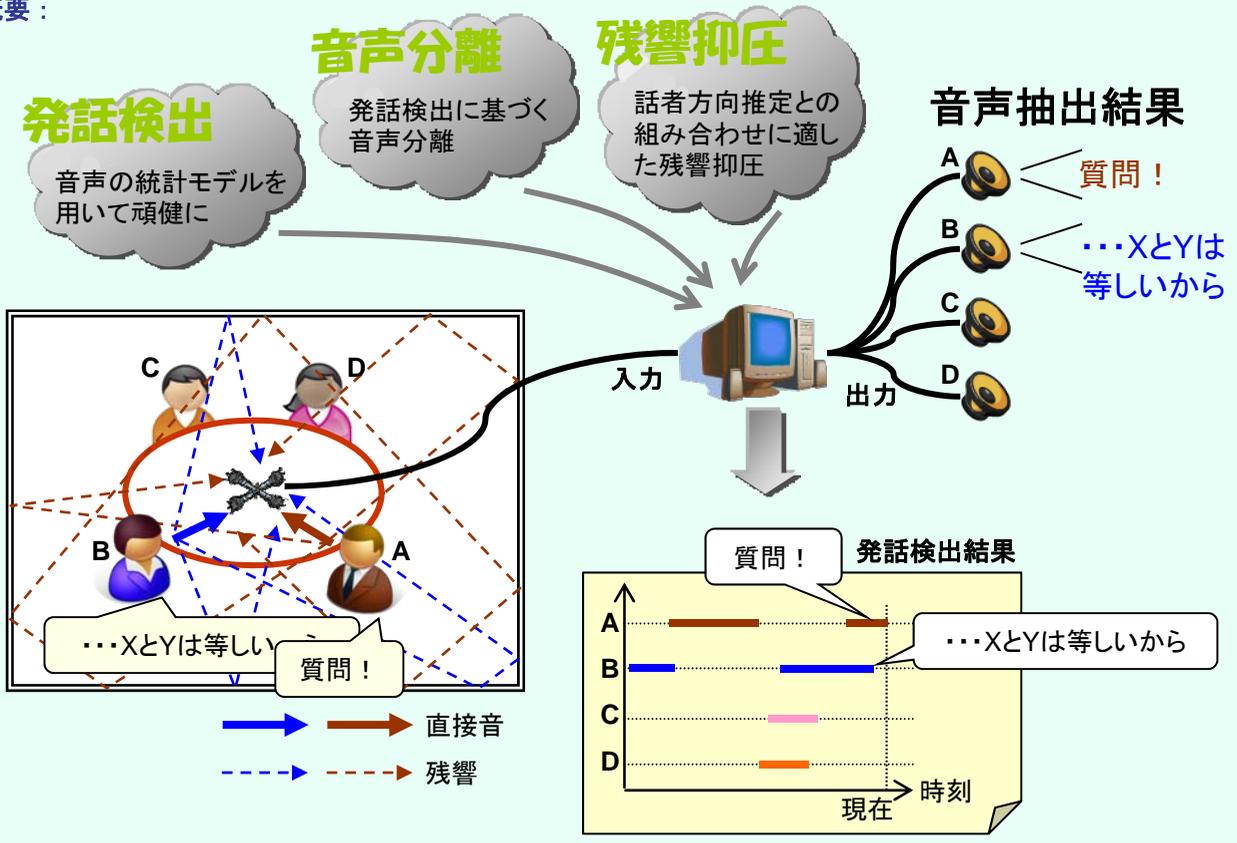
会話の中で一人一人の声を取り出します —オンライン発話抽出技術—

背景・課題：マイクロホンアレイ技術などの進歩により、一人の音声を対象とした音声認識システムが実用化されてきています。しかし、こうしたシステムは、会話のように複数の人が発話する場面では性能が劣化してしまいます。そこで、会話の中で一人一人の音声を抽出する技術が必要とされています。

アプローチ：我々は、発話検出や雑音・残響抑圧などの要素技術のブラッシュアップと会話音声処理のための技術の構築の両方を体系的にすすめています。要素技術の研究では、会話音声処理を含めた様々な用途に応用可能な汎用性の高い技術の構築を目指しています。

到達点：発話検出や残響抑圧、会話音声分離などを組み合わせた処理をオンラインで実行することができるようになりました。しかし、現在の技術では一定の時間遅延が生じるため、さらに遅延を小さくすることが今後の課題です。

概要：



Yoshioka



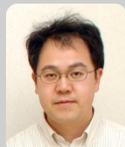
Fujimoto

関連文献

- T. Yoshioka, et al., "Blind separation and dereverberation of speech mixtures by joint optimization," IEEE Trans. Audio, Speech, Lang. Process., available on IEEE Xplore, 2010.
- M. Fujimoto, et al., "Noise robust voice activity detectoin based on switching Kalman filter," IEICE Trans. Inf., Syst. E91-D(3), 467-477, 2008.
- K. Ishizuka, et al., "A speaker diarization method based on the probabilistic fusion of audio-visual location information," Proc. ICMI-MLMI, 55-62, 2009.



Araki



Nakatani

連絡先

吉岡 拓也 (Takuya Yoshioka)

メディア情報研究部 信号処理研究グループ