

会話中のあなたの話、何でも聞いて理解します

～話し言葉音声認識のための高度な学習、適応、正規化技術～

どんな研究？

現在、携帯電話やスマートフォンを用いた音声検索や音声翻訳等が普及しつつありますが、現行では単語や定型句等の音声発話を認識するに留まっています。本展示では、会議や雑談等における、くだけた話し言葉音声の高精度な認識方法について紹介いたします。

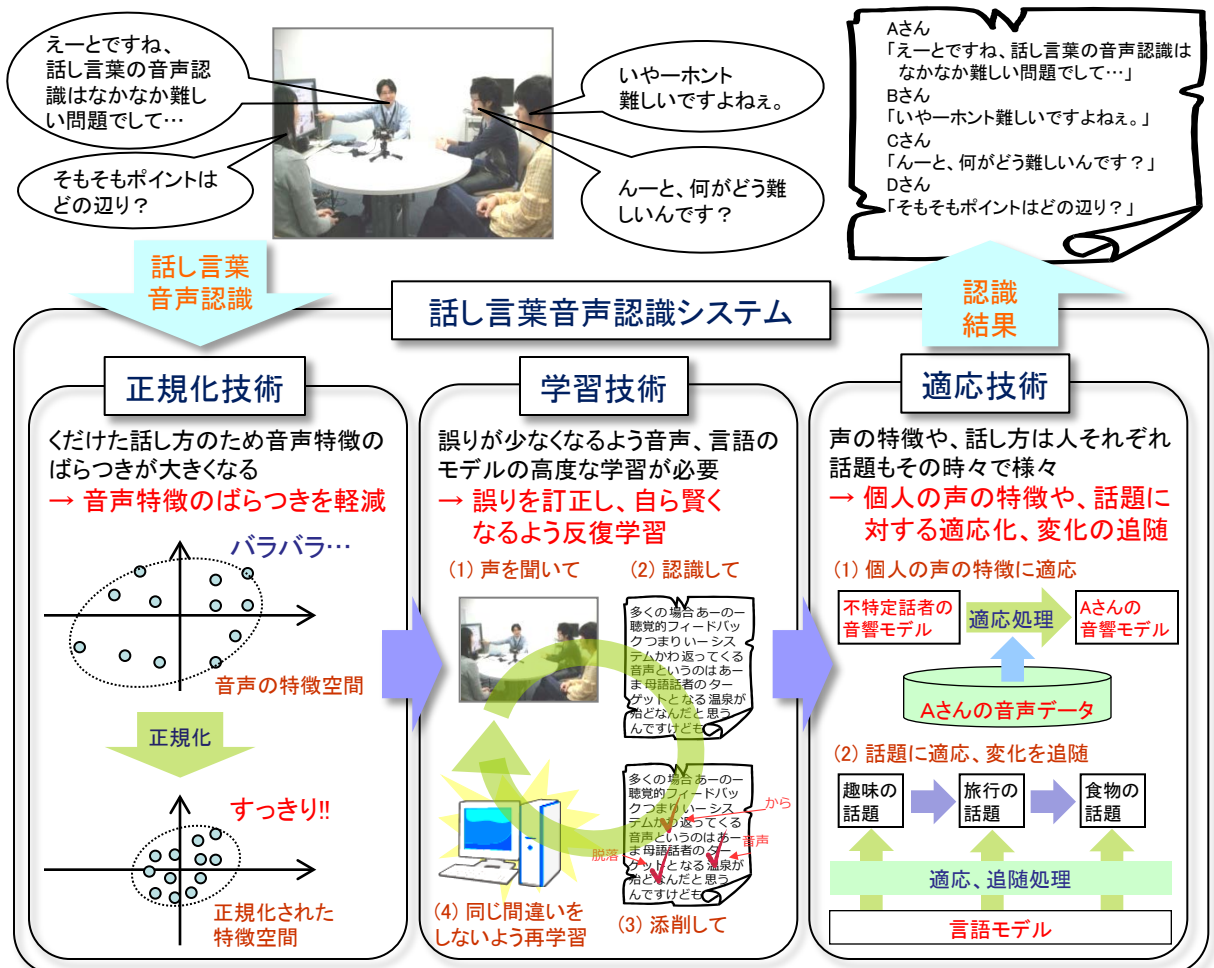
どこが凄い？

話し言葉の高精度な認識には、音声特徴の正規化（揺らぎの軽減）、音声と言語のモデル学習、個人の音声や話題への適応が必要不可欠です。本技術では、高度な学習理論に基づきこれらを構成し、世界でも稀にみる高精度な話し言葉音声認識を実現しました。

どんな風に役立つ？

本技術を用いることにより、会議における議事録の自動生成、映像コンテンツに対する実時間自動字幕作成等が実現されます。また、スマートフォンの音声アプリケーションにおいても定型句ではなく、話し言葉による、より自然で複雑な情報入力が可能となります。

会話シーン分析技術に導入した最先端の音声認識技術をピックアップしてご紹介
会議や雑談における、くだけた話し言葉音声を高精度に認識！！



関連文献

- [1] T. Hori, S. Araki, T. Yoshioka, M. Fujimoto, S. Watanabe, T. Oba, A. Ogawa, K. Otsuka, D. Mikami, K. Kinoshita, T. Nakatani, A. Nakamura, and J. Yamato, "Low-latency real-time meeting recognition and understanding using distant microphones and omnidirectional camera," IEEE Trans. ASLP, vol. 20, no. 2, pp. 499-513, Feb. 2012.
- [2] 大庭隆伸, 渡部晋治, 藤本雅清, 小川厚徳, 堀貴明, 中村篤, "いつ誰が何を話したかを即座に認識するオンライン会話分析システム - (3) 音声認識・イベント検出・トピック抽出 -," 日本音響学会講演論文集, 2-9-8, Sept. 2010.

連絡先

藤本雅清 (Masakiyo Fujimoto) メディア情報研究部 信号処理研究グループ
E-mail : fujimoto.masakiyo[at]lab.ntt.co.jp ({at}の部分をも@に置き換えてください)