

雑多な音の中からあなたの声だけ聞きとります

～時間・空間・周波数情報を統合した高精度音声強調技術～

どんな研究？

収録音に目的話者以外の人の声や雑音が含まれている場合、目的話者の声のみを聞き取ることは、困難なことです。この展示では、このような音から目的話者の声のみを抽出（音声強調）し、人や機械にとって聞き取りやすい音を作る技術を紹介します。

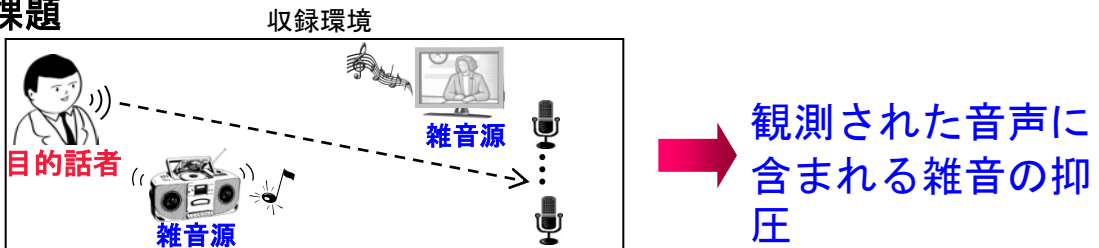
どこが凄い？

話者・雑音に関する全情報（時間・空間・周波数情報）を統合的に活用する音声強調方法を提案しました。この方法を用いることで、雑音下での音声認識に関する学会コンテストで世界トップの成績を収めました。

どんな風に役立つ？

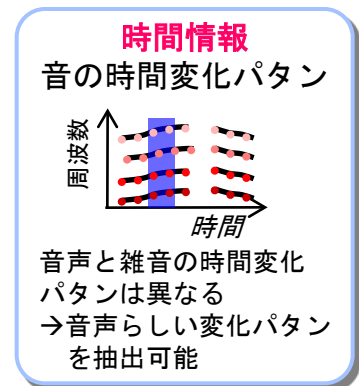
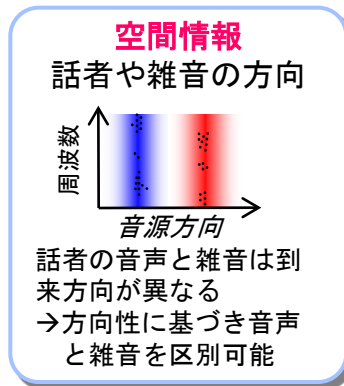
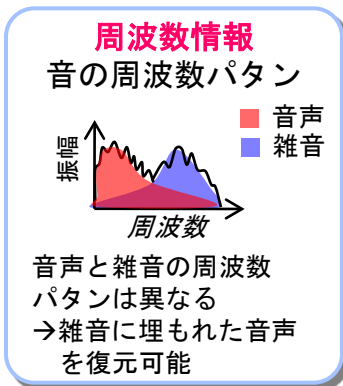
この研究を発展させることで、さまざまな収録環境で、目的の話者の声を高品質に収録できるようになります。例えば、遠隔コミュニケーションシステムにおける音声品質の改善や、ロボットとの音声コミュニケーションの円滑化などに役立つことが期待されます。

■ 課題



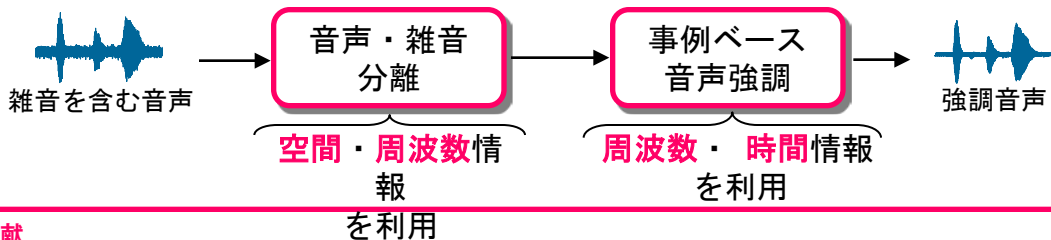
■ 着眼点

観測音声信号は以下の3つの情報を含む



■ アプローチ

これらの情報を活用するため、以下の二つの音声強調を利用



関連文献

- [1] Delcroix, M., Kinoshita, K., Nakatani, T., Araki, S., Ogawa, A., Hori, T., Watanabe, S., Fujimoto, M., Yoshioka, T., Oba, T., Kubo, Y., Souden, M., Hahm, S.-J. and Nakamura, A., "Speech Recognition in the Presence of Highly Non-Stationary Noise Based on Spatial, Spectral and Temporal Speech/Noise Modeling Combined with Dynamic Variance Adaptation," Proc. of CHiME International Workshop on Machine Listening in Multisource Environments, pp. 12-17, 2011.

連絡先

デルクロア・マーク (Marc Delcroix) メディア情報研究部 信号処理研究グループ
E-mail : marc.delcroix[at]lab.ntt.co.jp ([at] の部分を @ に置き換えてください)