

それっぽくしゃべります

～非負値時空間分解法に基づく発話リズムの抽出～

どんな研究

外国語でのコミュニケーションの困難さは、しゃべること、聞くことの両方において存在します。我々は、言語による**発話リズム**の違いを考慮することが、外国語の正しい聞き取りや発話において重要であると考え、音声信号から発話リズムを抽出し、制御する技術を考案しました。

どこが凄い

発話時の音声信号から発話リズムを抽出するための**非負値時空間分解法**というアルゴリズムを考案しました。従来法（非負値行列因子分解）にはない音声に特化した制約を考慮しながら、高い精度で発話リズム（時間情報）を抽出でき、また発話リズムを制御できることが分かりました。

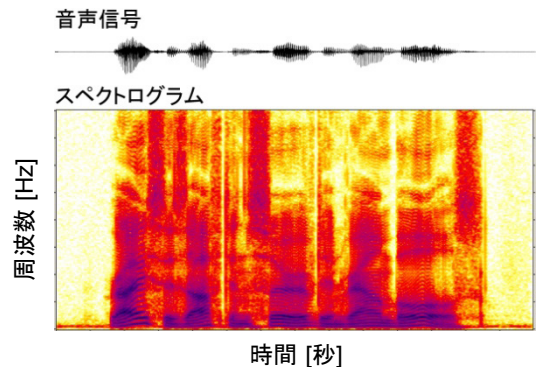
めざす未来

本技術は、外国語の聞き取りにおける発話リズムの重要性の検証のために大きく貢献できると考えられます。将来、日本語母語話者の英語の発話リズムを変換することで、英語が英語母語話者に伝わりやすくなり、英語でのコミュニケーションの苦勞が軽減されるかもしれません。

スペクトログラムを周波数情報と時間情報に分解

$$Y \cong A \Phi$$

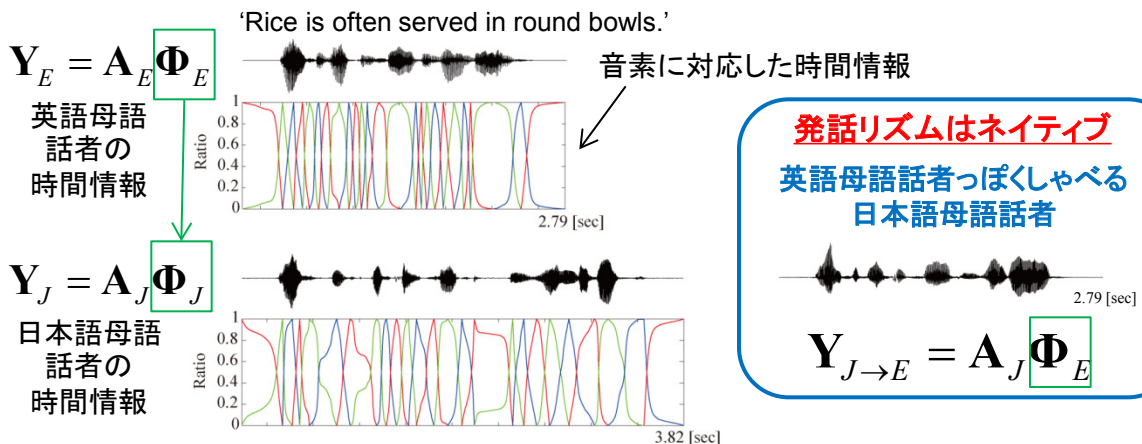
スペクトログラム = 周波数情報 × 時間情報 = 発話リズム



音声であることを考慮した時間情報への制約

- ①発話リズム（時間情報）を音素単位で表現するモデルを導入
- ②隣り合う音素のみが時間情報に影響を及ぼすこと（調音結合）を考慮

時間情報を入れ替えることによりそれっぽくしゃべらせる



関連文献

- [1] S. Hiroya, "Non-negative temporal decomposition of speech parameters," in *Proc. The 35th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2010.
- [2] S. Hiroya, T. Kitamura, "Generation of a vocal-tract movie based on sparse sampling," in *Proc. The 9th International Seminar on Speech Production (ISSP)*, 2011.

連絡先

廣谷 定男 (Sadao Hiroya) 人間情報研究部 感覚運動研究グループ
E-mail : hiroya.sadao[at]lab.ntt.co.jp ({at}の部分をもに置き換えてください)