



個人研究

スマート社会

表現豊かな音声の理解と合成



機械 (AI) が人と共生する時代には、音声による情報交換手段が重要かつ自然な役割を果たします。音声には文字で表せる言語情報だけでなく、話者に関する情報やニュアンスの違いの情報が含まれ、豊かなコミュニケーションを実現できます。心のもった音声を扱う研究をしています。

KEYWORDS 韻律解析ツールの提供と講習、感情音声の認識と合成、外国語発音の評価と矯正の自動化

RESEARCHER

コンピュータサイエンス学部 教授 大野澄雄

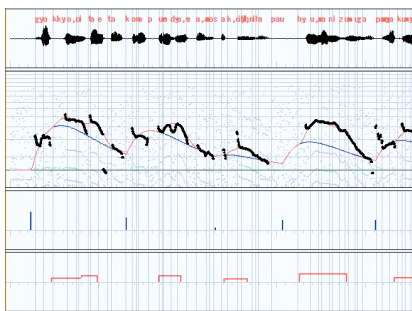


学会発表・論文・著書・社会活動

- [1] J. Shinzawa, S. Chen, J. She, H. Kameda, and S. Ohno, "The First Step towards Automatic Quality Evaluation of Chinese Vowel Pronunciations for Foreign Learners for Self-training," International Journal of Information and Education Technology vol. 9, no. 1, pp. 70-73, 2019.
- [2] Y. Arimoto, Y. Horiuchi, and S. Ohno, "Consistency of base frequency labelling for the F0 contour generation model using expressive emotional speech corpora," in Proceedings of Speech Prosody 2018, pp. 398-402, 2018.
- [3] T. Kawamura and S. Ohno, "A Study on the Difference of Emotional Perception Between Japanese and Chinese," in Procedia - Social and Behavioral Sciences, vol. 27, pp.95-104, 2011.

01 | 音声の基本周波数パターンの解析と生成

音声の高さの時間変化パターンを基本周波数パターン (FOパターン) と言います。自然な発声のFOパターンは複雑な形状をしています。その生成過程に基づいたモデルを利用することで、発声時に筋肉に与える僅かな数の指令で表現できます。FOパターンは強調表現や感情表現など言語表現だけでは伝わらない情報を担っているため、モデルへの指令を解析することにより、自動で細かなニュアンスを理解したり、適切な指令を与えることでそれらを表現する研究を行っています。



02 | 外国語学習者の発音の評価と矯正

外国語発音の学習には、お手本となり直してくれる先生が不可欠ですが、先生に代わってコンピュータが発音の良さを評価し、どのように直せば正しい発音になるかを自動的に矯正・指導するシステムを作成しています。アクセントやイントネーションの把握には上記のFO生成過程モデルを利用し、発音の把握には口の形状や舌の位置の情報を波形から抽出します。

