

イントネーションを分析、合成、変換

～音声基本周波数パターンの確率モデル～

どんな研究

線形予測分析 (LPC) は近代式の音声分析合成系を誕生させ、携帯電話という新たなコミュニケーション手段と統計的音声情報処理という研究パラダイムをもたらしました。LPCは音声の中でも音韻 (言語情報) の分析合成系を実現するのに対し、本研究は、**非言語情報に関する韻律的特徴の分析合成系**を実現します。

どこが凄い

LPCでは声道の物理モデルによる音声信号生成過程を確率モデル化し、統計的手法により声道パラメータを推定する枠組を与えました。本研究では、**甲状軟骨の物理モデル (藤崎モデル) による基本周波数パターン生成過程を確率モデル化**し、統計的手法により韻律パラメータを推定する枠組を与えることに成功しました。

目指す未来

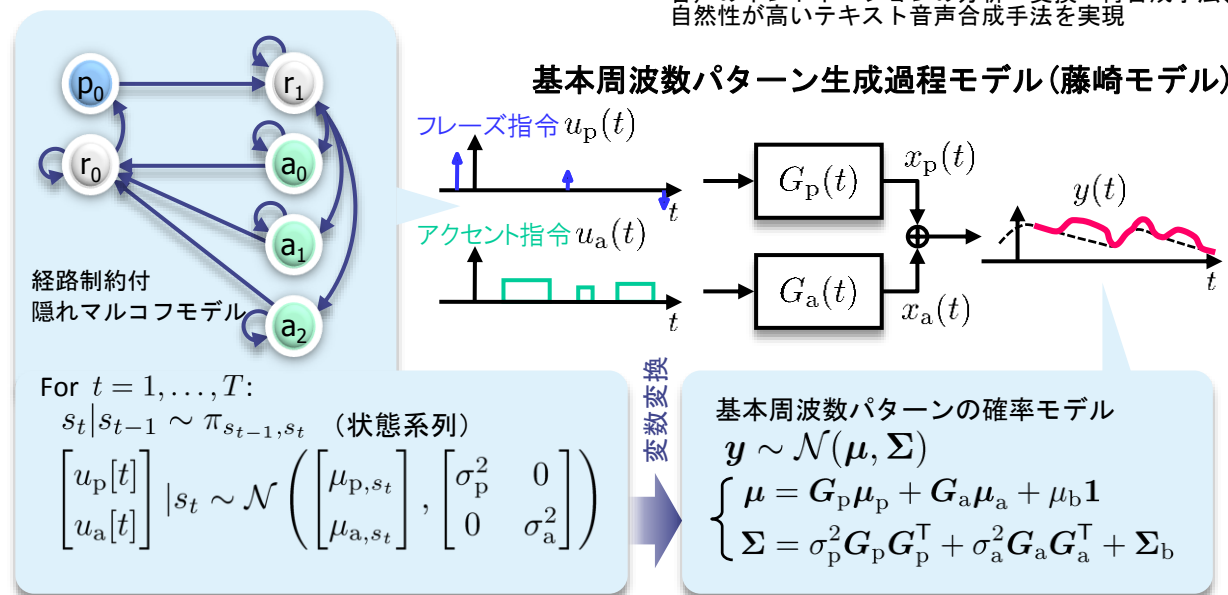
かつてLPCが統計的音声情報処理という研究パラダイムを生んだと同様、本研究は**音声の韻律に関する情報処理の新分野**を拓くきっかけになる可能性があります。現在、イントネーションが自然な音声合成手法や、音声の表情や個性などを認識・分析・変換する手法を開発することを目指しています。

基本周波数パターンとは

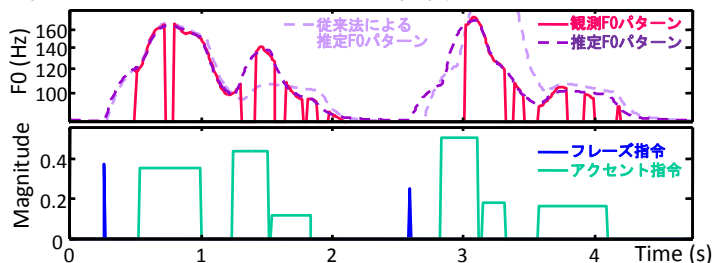
- ・ 音声の個性・調子・感情・意図などを反映した物理量
- ・ コミュニケーションにおいて重要な非言語的役割を担う

本研究のポイント

- ・ 藤崎モデルを確率モデルに翻訳 ⇒ 統計的手法による藤崎モデルのパラメータ推定法の確立
- ・ 音声のイントネーションの分析・変換・再合成手法、自然性が高いテキスト音声合成手法を実現



提案法によるフレーズ・アクセント指令推定例



何ができるようになる?

音声デモあり

- ・ フレーズ・アクセント指令を操作 ⇒ 自然性を保ったままイントネーションを自在に操作できる
- ・ テキストからフレーズ・アクセント指令を予測する帰帰問題に本モデルを適用 ⇒ テキスト入力から自然なイントネーションの音声合成できる

関連文献

- [1] H. Kameoka, K. Yoshizato, T. Ishihara, K. Kadowaki, Y. Ohishi, K. Kashino, "Generative modeling of voice fundamental frequency contours," *IEEE/ACM Transactions on Audio, Speech and Language Processing*, to appear, 2015.
- [2] K. Kadowaki, T. Ishihara, N. Hojo, H. Kameoka, "Speech prosody generation for text-to-speech synthesis based on generative model of F0 contours," in *Proc. The 15th Annual Conference of the International Speech Communication Association (Interspeech 2014)*, pp. 2322-2326, 2014.

連絡先

亀岡弘和 (Hirokazu Kameoka) メディア情報研究部 メディア認識研究グループ
E-mail: kameoka.hirokazu(at)lab.ntt.co.jp

