



通話音声から声の状態を診断

～遠隔医療における構音障害診断システムの開発に向けて～

<概要>

豊橋技術科学大学 IT 活用教育センターの細田侑也助教は、人間の発声メカニズムにおける声帯振動のピッチを通話音声から推定する手法を開発しました。本手法では、音声の振幅及び位相スペクトルから抽出された特徴量を複素平面上で統合してピッチを推定します。実験では、通信規格で周波数帯域が制限されている通話音声に対しても有効であるだけでなく、背景雑音が付加された環境でも提案法が頑健に機能することを示しました。

<詳細>

パーキンソン病などの神経変性疾患の重症化を防止するためには、初期症状である構音障害を早期に診断することが望ましいです。構音障害の特徴として、声の震えや呼気の乱れが挙げられます。臨床検査では患者の声から症状を診断しますが、時間や労力がかかるだけでなく、山間部などの遠隔地での対面実施は困難です。そこで本研究では、通信機器を介して回診する遠隔医療により、自動的に構音障害を診断するシステムの開発を目指します。

構音障害の患者には、肺から放出された空気で生じる声帯振動が喉や口腔で共鳴して声が生成されるという発声メカニズムに異常が生じています。本研究では、声帯振動に関する状態を診断するために振動周期（ピッチ）を推定することを目的とします。これまでに、音声の周波数解析で得られる振幅スペクトルに関する特徴量に基づいて、背景雑音に対しても頑健なピッチ計測手法が考案されています。しかし、遠隔医療における通話音声では、通信規格により振幅スペクトルが一部欠如しています。そのため、情報が削減されている振幅スペクトルから特徴量を抽出すると、ピッチを誤って推定する恐れがあります。

本研究では、振幅スペクトルに加えて、周波数解析の副産物である位相スペクトルから追加で特徴量を抽出する手法を提案しました。まず、時間方向及び周波数方向における位相のズレとピッチの関係式を導出して、観測された位相のズレを関係式に当てはめることで、ピッチを推定できることを検証しました。この知見に基づいて、関係式への当てはまり具合を定量的に評価する特徴量を新たに位相スペクトルから抽出しました。最終的に、振幅スペクトルから抽出される特徴量と複素平面上で統合することで、背景雑音に対する頑健性を維持しつつ、通話音声のピッチ推定で生じる特徴量不足を補いました。

図は、通話音声に対するピッチ推定結果を表します。振幅スペクトルのみを使用する先行研究では、帯域制限で情報が削減されているため、本来よりも高くピッチを推定しました。一方で提案法では、振幅及び位相スペクトルに関する特徴量を用いて通話音声から正確にピッチを推定しました。また、誤差が生じた区間の割合を表す評価指標 Gross Pitch Error (GPE) では、先行研究の 42.2% に対して、提案法は 9.5% まで改善しました。加えて、背景雑音が付加された通話音声に対しても 15.2% を達成して頑健性を示しました。

<今後の展望>

本研究では、声帯振動の異常を検出するためにピッチ推定に着目しましたが、呼吸器や口腔の異常も構音障害を引き起こす要因です。これらの症状を検出するために、振幅スペクトルから特徴量を抽出する手法が考案されていますが、位相スペクトルに関しては十分に検証されていません。将来的には、他の症例に関しても位相スペクトルから関連する特徴量を抽出することに取り組みます。また、それらの特徴量を総合的に解析することで、遠隔医療において有効に機能する構音障害診断システムを開発します。

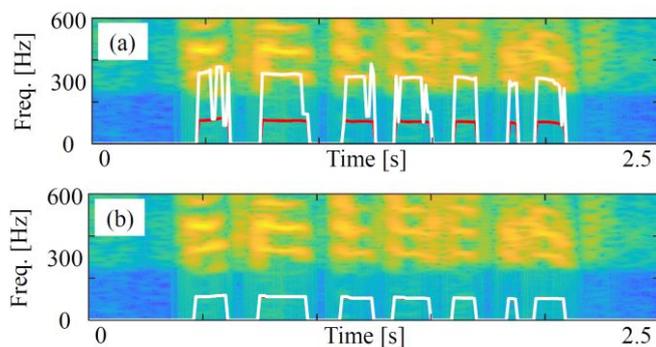
<謝辞>

本研究は、公益信託小野音響学研究助成基金の助成を受けたものです。

<論文情報>

Yuya Hosoda, Arata Kawamura and Youji Iiguni, "Complex-Domain Pitch Estimation Algorithm for Narrowband Speech Signals," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 31, pp. 2067-2078, 2023,
doi: 10.1109/TASLP.2023.3278488.

図：通話音声に対するピッチ推定結果。
赤色が実測値で白色が推定値を表す。
(a) 先行研究 (b) 提案法



本件に関する連絡先

広報担当：総務課広報係 岡崎・高橋

TEL：0532-44-6506 FAX：0532-44-6509