

研 究 主 論 文 抄 録

論文題目 逆フィルタ制御ホルマント推定法の精度評価と音声の分析的研究への応用

熊本大学大学院自然科学研究科 情報電気電子工学専攻 人間環境情報講座  
( 主任指導 上田 裕市 教授 )

論文提出者 坂田 聡  
(Tadashi Sakata)

主論文要旨

近年の音声認識などの主要な音声処理技術では、ケプストラム係数など、音声のスペクトルに関連したパラメータを確率論的なモデルで扱うことが主流となっている。一方、古くから、音声分析の分野では、音声の生成や聴取モデルと密接に関連する少数のパラメータを音声信号から探索する研究が多数行われてきた。声帯の振動を表現するピッチと声道の共鳴の共振周波数としてのホルマントはその代表であるが、これまで、多様な実音声から、ホルマントを高精度で推定する手法は確立されていないため、実用システムにホルマントを用いる試みは少ない。

このような背景の中、これまで聴覚障害、構音障害の補償を目的とした音声処理のために、音声の画像化システムやその発話訓練システムへの展開などの研究に取り組んできた。その中には、実用システムに使用できる安定で高精度のホルマント推定法として考案、開発された逆フィルタ制御(IFC)法が利用されている。本研究は、そのホルマント推定法(IFC法)の厳密な精度評価と声道長比推定法など音声の分析研究への応用に関して行われたものである。

論文の構成は以下のとおりである。

第1章は、まえがきとして、研究の背景と目的について述べる。

第2章では、本論文における評価対象である IFC 法の概要と特徴について解説している。

第3章では、零交差情報を用いてホルマントを推定する IFC 法と、スペクトル情報を用いる方式の代表として線形予測分析(LPC)法の推定精度を、合成音声を用いた評価実験によって比較した。合成音声の分析過程における詳細な視察によると、零交差情報はスペクトルの変形の影響を受けにくく、かつ、IFC 法の処理において零交差情報の荷重平均により高精度ホルマント推定が可能になることが示された。次に、モンテカルロ法による評価によって、零点を付加した合成音声では推定精度が極端に落ちる LPC 法に対し、IFC 法のホルマント推定精度には大きな低下は見られず、高精度の推定を可能にすることが確認され

た。また、IFC のシステムを LPC のパラメータで制御する Hybrid-IFC 法の結果から、分析パラメータとして零交差情報の有効性が確認された。

第 4 章では、従来、確立されているとは言えない実音声におけるホルマント推定精度の評価法として“データ縮減法”を提案し、IFC 法と LPC 法の推定精度を比較した。本研究で提案するデータ縮減法は、あるホルマント推定法を用いて得た実音声の F1-F2 ダイアグラムから、重み付き主成分分析を用いてエラーとみなされるデータを削除しながら、真の正規分布と見なされる領域としての  $3\sigma$  確率楕円(収束楕円)を推定する方法である。その結果、得られた収束楕円の外側にあるデータをホルマント推定エラーと定義することにより精度比較が可能になる。本章では、IFC 法を LPC 法と比較した結果、実音声においても、IFC 法が高精度ホルマント推定を可能にすることを示し、第 3 章の合成音声による結果を裏付ける。

第 5 章では、IFC 法で得たホルマント周波数比からの声道長比の推定を行った。これは、一様断面積音響管の長さとその共振周波数の関係を基礎として、多数の単語音声のホルマント軌跡から 2 人の話者の平均的な声道長比を推定するものである。その方法には、異なる話速の同一単語音声のホルマント軌跡から声道形状が相似となる瞬間を求め、その時のホルマントだけを用いる厳密法と、ホルマント軌跡全体を用いる直接法がある。結果として、厳密法では 100 単語以上の資料を用いることで 95%信頼区間が  $\pm 0.14\%$  程度の安定した推定が可能となり、直接法でも通常の数値比(2.0 以下)においては、10 秒以上の音声を用いることで信頼区間  $\pm 2\%$  以下で推定できることが示された。

第 6 章では、IFC 法のホルマント推定と第 5 章までに提案し評価を行った IFC 法とデータ縮減法、声道長比推定法の有効性を、4 種の音声データベースを用いた男女間平均声道長比の推定によって確認した。すなわち、ひとつの方法として、IFC 法と第 4 章で述べたデータ縮減法を用いて推定した母音ホルマント分布の中心点から男女間の平均声道長比を求めた。また、もうひとつの方法として、同じ音声データを用いて、第 5 章で述べた 2 話者間の声道長比推定から男女間平均声道長比を求めた。異なる 2 つの方法によって推定された男女間平均声道長比がどのデータベースでも約 1%以内の誤差で一致することから、両推定法の過程で用いる分析手法が有効かつ妥当なものと結論した。

第 7 章では、IFC 法による高精度ホルマントを用いた実用研究例として、リアルタイム音声画像化システムによる発話訓練と、構音障害音声の定量評価方法の検討について述べている。

最後に、第 8 章において、本研究を総括し、今後の課題について述べている。