

Title	パーティクルフィルタを用いた雑音に頑健な音声スペクトル上の複数ローカルピーク推定に関する研究
Author(s)	友池, 誠二
Citation	
Issue Date	2008-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/4312">http://hdl.handle.net/10119/4312</a>
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 修士

# パーティクルフィルタを用いた雑音に頑健な音声スペクトル上の複数ローカルピーク推定に関する研究

友池 誠二 (0610061)

北陸先端科学技術大学院大学 情報科学研究科

2008年2月7日

キーワード: 複数ローカルピーク推定, 非定常雑音, パーティクルフィルタ, 多次元尤度.

人間の音声の主な特徴量として高調波がある．高調波は音声認識，基本周波数推定，音声強調といった音声情報処理における重要な役割を担っている．高調波は周波数領域における音声スペクトルのローカルピークと密接な関係がある．音声スペクトル上のローカルピーク推定が可能になれば，高精度な音声の高調波が得られる．例えば，高調波成分のスペクトル強調に基づいた音声認識や，高調波の周波数を規範とした楕型フィルタによる基本周波数の推定などに応用できる．そのため，音声スペクトル上のローカルピーク推定は，様々な音声信号処理における特徴量として重要な役割を持つ．

音声スペクトル上のローカルピーク推定に関して，従来法では過去のフレームで推定されるローカルピークに関する知識を利用しておらず，現在のフレームで得られた情報のみを使ってローカルピーク推定を行っている．そのため，従来法の推定精度は現在のフレームに存在する雑音に大きく影響されるという問題点がある．過去のフレームで推定されるローカルピークに関する知識を学習することができれば，現在のフレームにおけるローカルピークの存在の確からしさをを用いて雑音による影響を最小化したローカルピーク推定が可能となる．

本論文では，雑音に頑健な音声スペクトル上の複数ローカルピーク推定を考える．つまり，従来の音声スペクトル上のローカルピーク推定では扱うことのできなかつた雑音環境下での音声スペクトル上の複数ローカルピーク推定を目指す．その方法として，過去のフレームのパラメータから現在のフレームのパラメータを推定する手法であるパーティクルフィルタを用いる．パーティクルフィルタはシステムの状態推定に際して，事後分布を正確に近似することで，システムの状態を推定する方法である．入力される音声によって，ローカルピークの個数は異なり，各ローカルピークの動きは独立している．そのため，音声スペクトル上のローカルピーク推定を行うためには，ローカルピークの個数を仮定すべきではない．そこで，提案法では，ローカルピーク推定に際して，音声スペクトル上の

複数のローカルピークを同時に推定する尤度を導入する．この尤度を用いることで複数のローカルピークの存在確率を同時に推定可能である．ローカルピークの存在確率が高い部分は近接フレームにおいてもピークが存在する確率が高い．そのため，提案法は個数が未知である複数のローカルピーク推定に適用可能である．さらに，状態推定に尤度と事後分布のみを用いているため，ローカルピークの動きのモデル化を必要としない．

提案法は，大きく分けて2つの手順で構成される．第1の手順は，ケプストラムから得られるスペクトル包絡を尤度とするピークの存在確率推定である．この方法で得られるピークの存在確率は尤度を用いて動的に更新される．そのため，提案法は音声スペクトル上のローカルピークの変化に関するモデルを用いることなく，ピークの存在確率を推定可能である．高調波は，高域成分ほど基本周波数の整数倍にはなっていない．提案法では，個数が未知である各倍音がわずかなゆらぎを持つような独立した動きを持つ場合であっても，パーティクルフィルタによってピークの存在確率を細かく表現することによって，複数のローカルピークを推定するようなピークの存在確率を計算可能とする．第2の手順は，ピークの存在確率から得られる，ローカルピークとなりうる候補からローカルピークを抽出する手法である．

提案法が雑音環境下で精度良くローカルピークの推定が行われているかを検証するために，2つの実験を行った．第1の実験では，提案法と従来法に対して，音声区間全体に白色雑音，ピンク雑音を付加した合成音声を用いて定常雑音に対する評価を行った．第2の実験では，提案法と従来法に対して，初期フレームでない音声区間に，時間長可変の狭帯域雑音を付加した合成音声を用いて非定常雑音に対する時間的な頑健性の評価を行った．第1の実験によって，定常雑音に対する提案法と従来法の精度を比較する．第2の実験によって，クリーン音声で学習した提案手法が非定常雑音に有効であるかどうかを検討する．評価は，推定されるローカルピークの個数がどれだけ正確であるかを測るために，正解のローカルピークの個数よりどれだけ過剰推定したかを評価尺度の一つとする．また，推定されるローカルピークの周波数位置が，正解のローカルピークとかけ離れていないかを測るために，推定ピークと正解ピークとの距離をもう一つの評価尺度とする．

2つの実験の結果，推定ピーク数と正解ピーク数との差，推定ピークの正解との距離の両方の尺度から，提案法は非定常雑音環境下においても従来法を上回る推定精度であった．また，正解ピークの基本周波数の変化の差異，話者の差異，母音内容の差異によらず，同様の傾向が得られた．よって，提案法は従来法の問題点であった雑音に対する頑健性の高めることがわかった．