JAIST Repository
https://dspace.jaist.ac.jp/

Title	時間情報と周波数情報を用いた雑音環境における基本周波数推定に関する研究
Author(s)	石本，祐一
Citation	
Issue Date	2004-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/951
Rights	
Description	Supervisor:赤木　正人，情報科学研究科，博士

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

時間情報と周波数情報を用いた
雑音環境における音声の基本周波数推定に関する研究

石本 祐一
北陸先端科学技術大学院大学

2004 年 1 月 22 日

論文の内容の要旨

本研究では、雑音環境においても頑健で高精度な基本周波数推定を目指し、ヒトの聴覚機構と同様に時間情報と周波数情報の両方を利用した雑音に対しても頑健な手法と、帯域幅可変櫛形フィルタによる雑音抑圧、及び瞬時周波数の不動点を利用した高精度な手法の 3 つを組み合わせた基本周波数推定法を構築する。

音声の基本周波数は多くの音声情報処理において用いられる重要な特徴であるため、これまでにも様々な基本周波数推定法が提案されている。しかし、雑音の存在に対応した決定的な手法はまだ確立されていない。これらの多くは、基本周波数を表す情報として、時間領域に現れる音声の周期性の特徴 (時間情報) と周波数領域に現れる調波性の特徴 (周波数情報) のどちらかを利用している。低周波数帯域に強い雑音のエネルギーが存在すると音声波形は雑音の影響により歪み、音声波形の自己相関法のような時間情報のみを利用する推定法では推定誤差が大きくなる。一方、中・高周波数帯域の雑音エネルギーが強い場合には音声スペクトルの調波構造が歪み、ケプストラム法のような周波数情報のみを利用する推定法は雑音の影響を受けやすい。実環境雑音は特定の時間-周波数帯域に偏ったエネルギー分布を持つため、雑音の特性によって基本周波数抽出に有用な音声の特徴や周波数帯域は異なり、時間情報と周波数情報のどちらかのみを利用する従来手法では実環境雑音に対応できない。ヒトは雑音が存在していても基本周波数に対応する音の高さを知覚できるが、ヒトの音高知覚には時間情報と周波数情報の両方が用いられると考えられている。本研究では、上記の実環境雑音の特性から音声のエネルギーに対して雑音エネルギーが相対的に小さい帯域が存在すると考え、その帯域の時間情報と周波数情報の両方を利用し統合することで、雑音に対して頑健な推定を行う。ただし、この推定はクリーン音声に対する高精度な手法と比べると推定精度の点で劣るため、次に雑音を抑圧した後に高精度な基本周波数推定を行うことを考える。実環境雑音の特性は時々刻々と変化するために、雑音の特性を予め仮定する雑音抑圧手法は適していない。そのため、時間情報と周波数情報を用いて得られた推定基本周波数を利用して構築した櫛形フィルタを用いて雑音を抑圧する。その後、雑音抑圧音声に対して高精度な基本周波数推定を行う。以上の方策より、提案手法は次の 3 つの段階から構成される。

(1) 初期推定部 - 雑音を含む音声に対して瞬時振幅の時間-周波数表示に現れる時間情報と周波数情報を利用し雑音に対して頑健な基本周波数推定を行う。雑音エネルギーの小さい帯域に現れる音声の基本周波数の時間情報と周波数情報を利用することで、雑音が存在する環境でも頑健な推定が可能となる。この初期推定部における推定結果は信号対雑音比が 3 dB 程度であっても推定可能区間の減少を抑えることができる。

(2) 雑音抑圧部 - 初期推定部で得られた基本周波数を用い、音声の調波成分を残す帯域幅可変櫛形フィルタによって雑音付加音声の雑音抑圧を行う。このとき、初期推定部で推定された基本周

波数にはある程度の誤差が含まれているため、音声の調波成分が誤って除去されないように櫛形フィルタを設計する必要がある。よって、初期推定基本周波数の誤差の度合を推測し、櫛形フィルタの帯域幅を調節する。

(3) 最終推定部 - 雑音抑圧された音声に対して、瞬時周波数の不動点を利用した高精度な基本周波数推定を行う。瞬時周波数は雑音の影響を受けやすい特徴であるが、雑音抑圧部によって調波成分付近以外の雑音を抑圧できるため、雑音の存在にかかわらず高い精度を保ったままの基本周波数推定が可能となる。

計算機シミュレーションにより提案手法の耐雑音性能と推定精度を調査した結果、白色雑音やピンク雑音のような従来から評価に用いられていた人工的な雑音だけではなく、走行自動車内雑音のようにエネルギー分布が大きく偏った実環境雑音に対しても、クリーン音声に対する基本周波数推定とほぼ同程度の高い推定精度を得ることができ、本研究による手法が「雑音に対して頑健」と「高精度」の両方を満たす推定が可能であることがわかった。

雑音環境においても音声の基本周波数を推定できるという本研究の成果は、音声情報処理の幅広い分野に応用可能である。例えば、一般に雑音環境での自動音声認識では認識精度が大幅に低下することが問題とされているが、韻律辞書を作成し基本周波数の概形を特徴量として認識器に用いることにより認識精度の向上を図ることができる。また、基本周波数推定の精度が合成音声の自然性を左右する音声分析合成符号化においては、雑音の存在する実環境でのシステムの運用に貢献できる。さらに、様々な音が混じりあった状態から目的音声を抽出するカクテルパーティ効果の概念を基にした音源分離では、基本周波数を音源の違いを示す特徴として用いているため、本研究は聴覚情景解析の研究に対しても貢献できる。

キーワード: 瞬時振幅, 周期性, 調波性, 櫛形フィルタ, 瞬時周波数