

Title	複素ケプストラム分析を用いた残響音声の基本周波数推定に関する研究
Author(s)	細呂木谷, 敏弘
Citation	
Issue Date	2007-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/3603
Rights	
Description	Supervisor: 鶴木 祐史, 情報科学研究科, 修士

複素ケプストラム分析を用いた残響音声の基本周波数推定に関する研究

細呂木谷 敏弘 (510091)

北陸先端科学技術大学院大学 情報科学研究科

2007年2月8日

キーワード: 基本周波数, 残響音声, 複素ケプストラム分析, 音源フィルターモデル.

音声の基本周波数は, 音声の分析合成, 音声認識, などの様々な音声信号処理にとって非常に重要な特徴である. これまでに, 多くの基本周波数推定法が提案され, 原音声から高精度に推定できる手法が確立されはじめている. しかし, 実環境における音声信号処理を考えた場合, これらの推定法が雑音や残響の影響がある音声に対しても正しく機能することが求められる. 近年, 耐雑音性に関しては議論され, いくつかの手法が提案されているが, 耐残響特性に関してはほとんど議論されていない. 現在, 残響に頑健な基本周波数推定法は報告されておらず, また非常に困難であると考えられている. そこで本論文では, 残響音声から精度良く基本周波数を推定する手法を提案することを目的とする. 残響音声から基本周波数を精度良く推定できれば, 実環境での様々な音声信号処理に応用が可能である.

まず, 従来の基本周波数推定法が残響環境下でどれだけ機能するのかを見極めるため, 従来の基本周波数推定法を概説し, その中で代表的な10個の手法について耐残響性の評価を行った. 評価に用いた手法は, 時間領域における周期性を利用した手法として自己相関法, AMDF法, LPC残差法, ACMWL法, 周波数領域における調波性を利用した手法としてSTFT-combフィルタ法, SHS法, リフター法, ケプストラム法, TEMPO, PHIA法である.

評価に使用したデータは, 音声とEGGが同時収録されたデータベースで, 男女各12名, 4文章を使用した. 音声への残響特性の付加は, 指数減衰型の非最小位相インパルス応答を計算機上で畳み込むことで実現した. 評価尺度は正答率とSNRを利用して評価した. 正答率は正しく推定できた区間の割合であり, 残響に対する頑健性を示す. SNRは基本周波数を推定できた区間内における, 推定した基本周波数の誤差を表すことから, 推定精度を示す.

評価結果から, どの手法も残響時間が長くなるにつれ推定精度が著しく低下し, 残響時間が2sの場合では50%を下回る正答率となっていることが分かった. ケプストラム法は

クリーンな状況でこそ TEMPO に劣ったが、残響時間が長くなっても比較的精度が低下していなかった。

本論文では、この評価結果を踏まえ、複素ケプストラム分析を用いた基本周波数推定法を提案した。提案法は二つのブロックで構成される。

前半部では、複素ケプストラム分析を用いて残響音声から音源情報を取り出す。音源・フィルターモデルを仮定すると、音声のケプストラムは、低ケフレンシー部に声道情報、高ケフレンシー部に音源情報とに分離する。従ってリフターにより高ケフレンシー部の成分だけを取り出せば、基本周波数推定に必要である音源情報のみを抽出することができる。同時にこの操作は、残響の低ケフレンシー部の成分を捨てることができ、残響の大きな成分を取り除くことにもなる。

後半部は残った音源情報から周期性や調波性を利用して基本周波数を推定する。後半部の処理は、抽出した音源情報の周期性を取り扱うか、調波性を取り扱うかで様々な方法が考えられるが、本論文では、周期性を取り扱う方法として時間領域での自己相関法を利用したもの、調波性を取り扱うものとして振幅スペクトルの comb フィルタ法を利用したものの2種類構築した。

提案法の残響に対する有効性を評価するため、従来法と正答率の比較を行った。評価方法は従来法の評価と同じである。この結果、comb フィルタ法を利用した提案法が残響時間が 0.1 s 以上の場合、従来法よりも推定精度が高いことが分かった。また、クリーンな状況での推定精度も高かったことから、提案法は幅広い環境で使用できる手法であることが分かった。従って、複素ケプストラム分析により抽出した音源情報は残響に対して頑健な特徴であることが分かった。

提案法は従来法を上回る推定精度であったことから、分析フレーム内での残響の影響はある程度取り除くことができていると考えられる。しかし、残響のインパルス応答は通常、基本周波数推定で用いられる窓長よりも長いため、過去のフレームの情報が現在のフレームに影響を与えているが、提案法はこの影響を考慮していなかった。そこで、さらに推定精度を上げるため、残響のインパルス応答よりも長いフレーム長を用いた処理を検討した。

まず、音声に大きく影響を与える残響の特徴量を調査した。その結果、残響の非最小位相の位相成分が大きく影響を与えていることが分かった。従って、残響の非最小位相成分を取り除くことができれば、さらに残響に頑健に基本周波数を推定することができると考えられる。そこで、この残響の非最小位相成分をケプストラム平均除去法を用いて取り除くことを検討した。しかし、非最小位相ケプストラムに対しケプストラム平均を求めることには問題があったため、ケプストラム平均除去法では効果がなかった。

今後の課題としては、残響の非最小位相成分を取り除く処理を構築し、推定精度を上げることと、実環境での残響を用いた評価を行うことが挙げられる。