

雑音・残響に頑健な基本周波数推定法に関する研究

Study on Fundamental Frequency Estimation Method
with Robust to Noise and/or Reverberation

情報コミュニケーション研究室

山川 拓真¹⁾

指導教員 三輪 賢一郎

1) サレジオ工業高等専門学校 情報コミュニケーション研究室

キーワード：基本周波数推定，非同期検波，雑音，残響

1. 緒言

音声の基本周波数（以下， F_0 ）は，音の高さを決定づける重要な要素である．例えば，楽曲には必ずメロディが存在しているが，これは F_0 の変化によるものである．他にも F_0 は，音声分析合成，音声強調処理，音声認識のほか，非言語情報の認識，変換といった様々な音声信号処理で重要な特徴量として利用されている．したがって F_0 推定の手法は，雑音や残響が存在する実環境に対応できるものが望ましい．

実環境に対応した手法として AM 復調技術を用いた FreeDAM が提案されている[1]．本研究では，FreeDAM の課題であった計算量を大幅に低減した改良版である FreeDAM Lite を検討してきており，雑音と残響のそれぞれに対し十分な耐性を持つことまでは確認している．本稿では，同手法 (FreeDAM Lite) が雑音と残響の両方が混在する環境において有効であるかを検討する．

2. 理論

2.1 観測領域の拡大

提案法においては，3 本の隣り合う調波を AM 信号（変調波）に見立てて復調を実行することで F_0 推定を実現している．当初の提案法のやり方では，1 組の変調波だけを用いて復調を実行していた．しかしながら，さらに性能を向上させるためには用いている情報の拡大を図る必要がある．そこで今回は，

分析対象とする変調波を 1 組から 8 組に変更することとした．具体的には，図 1 に示すように，入力信号（10 次調波複合音）の基本音から第三倍音の変調波 1 から第八から第十倍音の変調波 8 までの計 8 つの変調波を抽出し，非同期検波（ヒルベルト変換）により順次復調を実行する．次に，復調した 8 つの F_0 推定値から，次の 2 つの指標を用いて尤もらしい F_0 を決定する．

- (1) 復調波形の振幅値：振幅値が大きいほど尤もらしいと判定．正解の復調波形は，入力信号の本来の調波をフルに用いることから振幅が大きくなることを利用．対して，不正解の復調波形は，雑音のスペクトラムを復調に用いるため振幅が小さくなる．
- (2) 復調波形の歪度：歪度が小さいほど尤もらしいと判定．正解の復調波形は綺麗な正弦波となることを利用．具体的には，復調波形と正弦波との相互相関係数により判定する．

これら 2 つの指標をそれぞれ点数化し，総合得点が最も高いものを尤もらしい F_0 として出力する．

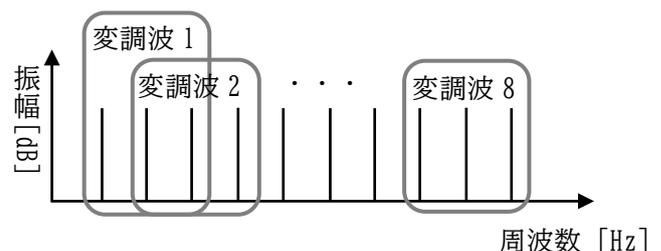


図 1 AM 変調波の作成方法

2.2 雑音除去機構

耐雑音性を向上させるにあたり、スペクトルサブトラクション法(spectral subtraction method)と呼ばれる雑音除去機構を採用する。

本実装では、雑音を含む入力信号のパワースペクトルの中央値を求め、中央値以下を除去することで雑音除去を実現している。対象信号の入力直後にこの処理を実行することとしている。

3. シミュレーション方法

今回は、理想的な調波構造を持つ音信号を対象に、MATLAB 上にてシミュレーションを実施した。

入力信号は、音声合成ソフトウェア「VOICEBOX」を用いて男声・女声・子供の声により/aoiumihaemoi/(青い海はエモイ)を発話させ、それぞれの発話に対して 0.1 秒ごとに TEMPO 法[2]を用いて F0 を計測する。それをもとに、図 2 に示すような階段状の 10 次調波複合音(時変)を作成した。

雑音残響環境は入力信号に白色雑音と人工的な残響(統計的室内インパルス応答)を付加してシミュレーションを実施した。SNR(信号対雑音比)は 20, 10, 0, -5, -10 [dB]の 5 通り、残響時間は 0.1, 0.3, 0.5, 1.0, 2.0 [s]の 5 通りで、これらを組み合わせた計 25 通りで実施した。比較対象は雑音に強いとされる SWIPE 法[3], 残響に強いとされる複素ケプストラム法(CmpCep)[4], 雑音と残響の両方に強い FreeDAM とした。評価指標は真値の F0 との許容誤差を 5 [%]以内とした正答率を用いた。

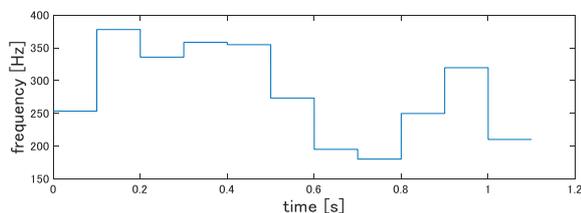


図 2 入力信号の F0 の例 (女声)

4. シミュレーション結果

図 3 に、女声領域での F0 推定結果を一例として示す。雑音残響環境において FreeDAM Lite の F0 推定精度は SWIPE, 複素ケプストラム法に対し十分高い結果となった。また、より緻密な同期検波方式

を用いる FreeDAM に対しても遜色のない結果を得られた。

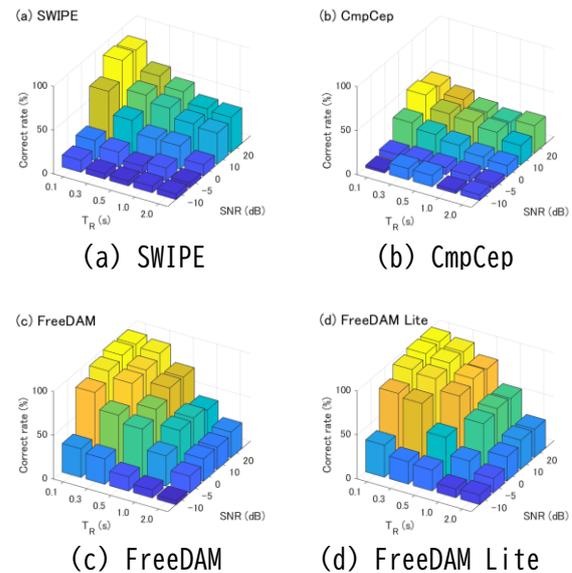


図 3 雑音残響環境における F0 推定結果 (女声)

5. 結論

本研究では、FreeDAM Lite が雑音と残響の両方が混在する環境において、一般的な従来法よりも高精度な F0 推定が可能であることをシミュレーションにより確認した。今後は、提案法の用途を音声信号にも拡大するべく検討を行う予定である。

謝 辞

本研究には、北陸先端科学技術大学院大学鷺木研究室の研究成果物[4]を利用しています。

参考文献

- [1] 三輪賢一郎, 鷺木祐史, “振幅変調音のピッチ知覚に基づいた調波複合音の基本周波数推定法,” 電子情報通信学会論文誌 A, Vol. J98-A, No.12, pp. 668--679, 2015 年 12 月.
- [2] Kawahara, H., Masuda-Katsuse, I. and Cheveigné, A., “Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds,” Speech Communication, vol. 27, pp.187-207, 1999.
- [3] A. Camacho, “SWIPE: A Sawtooth Waveform Inspired Pitch Estimator for Speech and Music,” Ph.D. Thesis, University of Florida, 2007.
- [4] 細呂木谷敏弘, 鷺木祐史, “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”, 北陸先端科学技術大学院大学修士論文, 2007 年.