

# 高分解能の周波数解析法を用いた スペクトルサブトラクションの改善

正員 広林 茂樹\*      学生員 柴野 洋平\*  
正員 山淵 龍夫\*

## Improvement of the Spectral Subtraction Method by High-resolution Frequency Analysis

Shigeki Hirobayashi\*, Member, Yohei Shibano\*, Student Member, Tatsuo Yamabuchi\*, Member

The Nonharmonic Analysis, recently proposed by the authors, is able to obtain voice and noise spectra independent of each other as a result of high frequency resolution. Therefore, NHA may even enable voices in a noisy environment to be extracted from spectra without distortion. For preprocessing by the spectral subtraction method of effective noise suppression, we used NHA instead of the discrete Fourier transform and the effects on noise suppression were quantitatively verified. In an environment of SNR from -10 dB to +10 dB, the proposed technique showed an improvement of about 4 dB on average.

キーワード：雑音抑圧，スペクトルサブトラクション，信号処理

Keywords : Noise suppression, Spectral subtraction, Signal processing

### 1. はじめに

音声認識など様々な分野において、測定信号から雑音の影響を軽減するための研究が盛んに行われている。それらのアプローチにはハードウェア的に大きく2つに分けて、信号の入力に単一のマイクロホンを使用する手法と、複数個のマイクロホンを使用する手法が提案されている。著者等は装置の小型化や設置の簡略化、また既に収録された信号から雑音を取り除くため、1入力系の雑音抑圧処理について検討を行った。1入力系の代表的な雑音抑圧法にはスペクトルサブトラクション<sup>(1)</sup>がよく知られている。これは雑音を事前に計測しその統計的なスペクトル包絡特性モデル化することで、雑音混合下において目的信号を抽出する手法である。さらに最近ではこのスペクトルサブトラクションの前処理として、Discrete Fourier Transform (DFT)の代わりにウェーブレット変換を用いて周波数を変換する手法も提案されているが、大幅な改善効果が見込めない<sup>(2)</sup>。

一方、著者等は高分解能を有する周波数解析法として最近 Nonharmonic Analysis (NHA)の研究を行っている<sup>(3)</sup>。NHAは非線形方程式の解法を用いて周波数とその振幅、初期位相などのパラメータを求めているため、DFTに比べて分析窓の影響が小さく高い分解能を有する。従って、スペ

クトルサブトラクションにおける周波数変換処理をDFTからNHAに置き換えることで雑音抑圧効果の向上が見込める。本報告では、白色雑音環境を想定して定量的にこの改善効果について調べた。

### 2. NHAを用いたスペクトルサブトラクション

図1左側に示されているように、一般にDFTに代表される調波形周波数解析では分解能が低いため雑音と音声のスペクトルが重複する。そのため、各周波数域で音声の振幅や初期位相が雑音に乱され、音声の抽出が困難になる。一方NHAは非線形方程式の解法を用いて対象信号に対し最適な正弦波のパラメータを推定するものである。すなわち最急降下法を用いて次式のように、対象信号と正弦波モデルの信号の差の二乗和が最小値になるように周波数 $f$ と振幅 $A$ と初期位相 $\phi$ を求めている<sup>(3)</sup>。

$$F(A, f, \phi) = \sum_{n=0}^{N-1} \left\{ x(n) - A \cos \left( 2\pi f \frac{n}{f_s} + \phi \right) \right\}^2 \dots\dots\dots (1)$$

$f_s$ はサンプリング周波数(Hz)である。複合正弦波に対しては逐次残差信号に対し同様に行い、複数の正弦波を抽出する手法である。高い分解能を有し、特性の異なる信号においてスペクトルの独立性が高い。図1右側に示されているように、それぞれ独立にスペクトル情報が保存されていれば、音声のスペクトル情報を抽出することが可能である。

\* 富山大学  
〒930-8555 富山県富山市五福 3190  
Toyama University  
3190 Gofuku, Toyama-shi, Toyama 930-8555

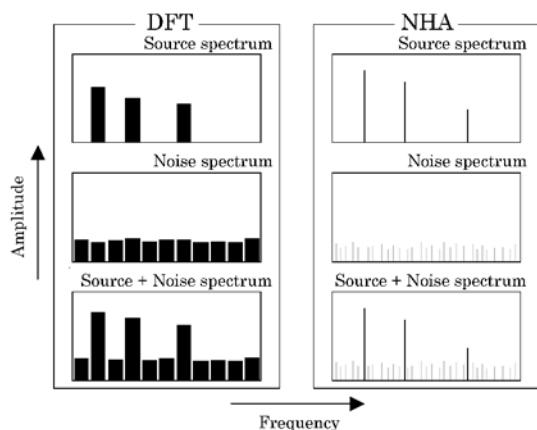


図1 DFTとNHAのスペクトル解析の概略  
Fig. 1. The outline of the spectrum analysis of DFT and NHA.

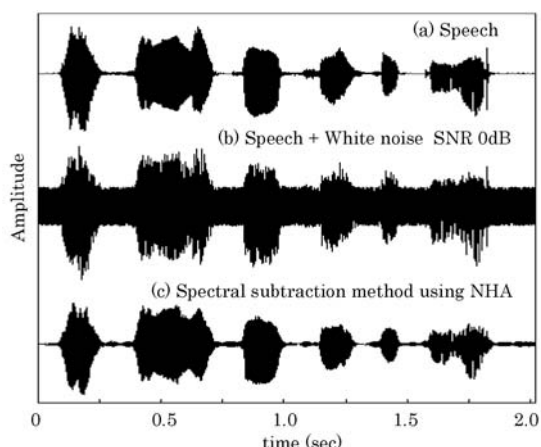


図2 SNR 0dBの白色雑音環境下におけるNHAを用いた回復信号例

Fig. 2. The example of a recovery signal using NHA under the white noise environment of SNR0dB.

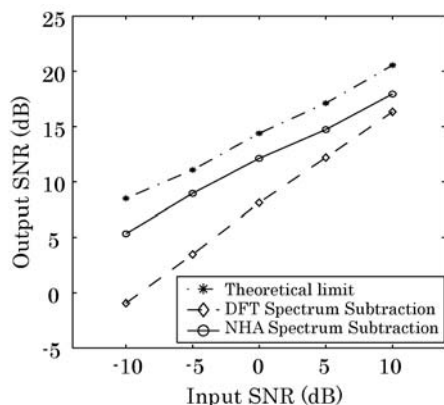


図3 スペクトルサブトラクションによる各S/Nの変化  
Fig. 3. Transition of S/N by Spectral subtraction.

### 3. 実験結果

入力信号はサンプリング周波数を 8kHz とし、約 2 秒間の音声信号に白色雑音を加えて生成した。一般的な評価に用いられている SNR -10dB から 10dB までの雑音環境下で 5dB 刻みで変化させた。単純な白色雑音環境を想定しているため、NHA を用いたスペクトルサブトラクションの閾値を、簡易的におおよそ雑音スペクトルの最大値とし、それ以下のスペクトルは雑音とみなすことにする。計算ではフレーム長 32ms , 1/16 オーバラップとした。図 2 に SNR 0dB 白色雑音環境下における NHA を用いたスペクトルサブトラクションの例を示す。0dB 環境の信号は白色雑音が付加された分だけ絶対的な大きさが増加しているが、回復信号では原信号の大きさに近づいていることがわかる。

図 3 で各雑音環境下で DFT を用いた一般的なスペクトルサブトラクションを行ったときの改善度を比較する。Theoretical limit は各雑音環境下の信号から得られる位相情報と原信号の振幅を用いて信号を回復したものである。また破線の DFT を用いたスペクトルサブトラクションの計算では、NHA の手法と同様とし、さらにスペクトル包絡はケプストラム域で 20ms のハニング窓で求めた<sup>(4)</sup>。全域で NHA を用いた方法が、従来の DFT を用いた方法を上まわる結果になった。Input SNR 0dB においては従来法と比べて約 4dB の改善度の向上が認められる。提案手法はほぼ Theoretical limit と平行に推移し、高雑音下では従来法との差が大きい。低雑音下ではやや改善度の向上が鈍るが、NHA の計算コスト削減のため原信号に対し 35dB 程度の品質で計算した影響があらわれているためである。

### 4. まとめ

本報告ではスペクトルサブトラクションの前処理として DFT の代わりに NHA を用いて改善効果の実験を行った。その結果、白色雑音環境下では SNR -10dB から 10dB の雑音環境下で平均約 4dB の改善が確認された。今後はさまざまな雑音環境下による回復効果を検討する予定である。  
(平成 16 年 6 月 11 日受付, 平成 16 年 8 月 11 日再受付)

### 文 献

- (1) S. F. Boll : "Suppression of Acoustic Noise in Speech Using Spectrum Subtraction", *IEEE Trans., ASSP*, Vol.27, No.2, pp.113-120 (1979)
- (2) R. Nishimura, F. Asano, Y. Suzuki, and T. Sone : "Speech Enhancement Using Spectral Subtraction with Wavelet Transform", *IEICE Trans*, Vol.J79-A, No.12, pp.1986-1993 (1996-12)
- (3) S. Hirobayashi, F. Ito, T. Yoshizawa, and T. Yamabuchi : "Estimation of the frequency of non-stationary signals by the steepest descent method", Proceedings of the Fourth Asia-Pacific Conference on Industrial Engineering and Management Systems, pp.788-791 (2002)
- (4) H. Nakashima, T. Usagawa, and M. Ebata : "Speech enhancement using harmonic-type filter", Technical Report of IEICE, EA99-82, No.12, pp.69-76 (1999)