

音声雑音除去に関する研究

- ランニングスペクトルフィルタ(RSF)の効果 -

大槻 典行* ・ 宮永 喜一**

Speech Noise Reduction Method on Frequency Domain

- Noise Reduction using Running Spectrum Filtering (RSF) -

Noriyuki OHTSUKI Yoshikazu MIYANAGA

Abstract - In this report, We propose a noise reduction method on frequency domain for speech signals. An additive noise in a speech signal is reduced this method using Running Spectrum Filtering (RSF). RSF is a filter on frequency domain. The noise reduction method using RSF dose not need the estimated noise spectrum.

Key words : noise reduction, frequency domain, digital filter, additive noise

1. はじめに

近年、情報機器の発達には目覚ましいものがあり、中でも携帯電話や音声認識装置などの音声入力を使用した機器の利用が増えている。しかし、音声を利用した機器を使用する場合、環境によっては使用に耐えない状況がある。これには、周囲でおこる雑音がこれらの機器のマイク等の入力装置から混入する場合、マイクの特性的変化などの機器自体あるいは周囲の反響等による伝達特性的変化が起きる場合等がある。何れも、音声非常に聞き取りにくくなったり、音声認識率が極端に低下する等、音声の持つ情報を正確に取得する事が困難となる。このような問題に対して従来より雑音を除去する手法が提案されてきた。この雑音除去法には、複数のマイクロホンを用いて雑音を取り除く方法[1]、適応フィルタリングにより雑音を取り除く方法[2]、周波数領域で雑音のスペクトルを取り除く方法[3]等数多くの手法が提案されている。この中の、周波数領域における雑音除去法として、スペクトルサブトラクション法[4]がある。この雑音除去手法は、複数のマイクロホンを必要とせず非常にシンプルな処理で効果が非常に高いものである。しかし、雑音のスペクトル

を推定する必要があり、この推定に誤りがあると十分な雑音除去効果が得られないと言う問題があった。そこで我々は、雑音のスペクトルを推定すること無く雑音除去可能なRunning Spectrum Filter(RSF)を用いた雑音除去法を提案する。この手法は、周波数領域で時間的に変化する短時間スペクトル(Running Spectrum)に対して、各周波数毎に時間軸方向にデジタルフィルタを通過させる事で雑音を除去する方法である。本報告では、RSFのアルゴリズムを示し、雑音除去効果が非常に高いことを実験で示す。

2. 周波数領域における雑音の除去

2.1. スペクトルサブトラクション

周波数領域における雑音除去手法としてスペクトルサブトラクション法[4]（以下SS法）が提案されている。この手法は、シンプルなアルゴリズムであるにも関わらず非常に高い雑音除去効果が得られ、この手法を改良したアルゴリズムも多数提案されている。以下に、この手法を簡単に説明する。

SS法は、雑音が付加された音声信号のパワースペクトルから、別途推定した雑音のパワースペクトルを差し引き、そのパワースペクトルをフーリエ逆変換することで雑音を除去した音声信号を復元するものである（図1）。この時、位相情報は、原信号の位相を用

* 釧路高専情報工学科

** 北海道大学工学部

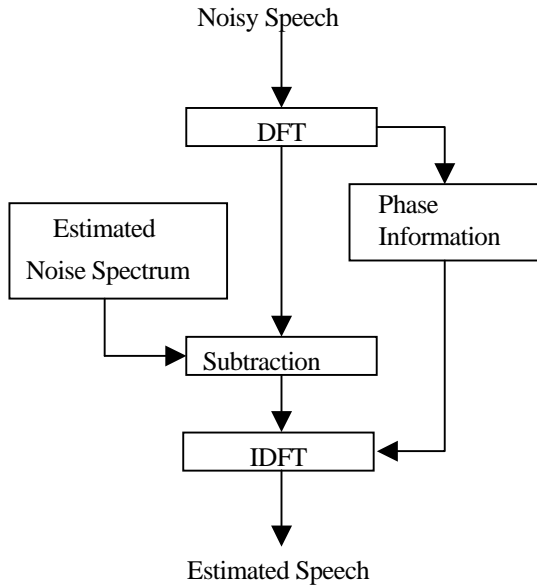


図1 . スペクトルサブトラクション法

いる .

2.2. 雑音推定スペクトル

スペクトルサブトラクション法は定常的な雑音が音声に付加した場合の雑音除去に関しては非常に有効であるが、非定常な雑音や雑音のパワースペクトルを推定できない場合には、その効果が十分に得られない。特に、雑音を推定する区間が無音区間である事、推定した区間の雑音と音声に付加されている雑音のパワースペクトルが等しいことが条件となる。つまり、推定した区間から雑音が増加した場合には、雑音の引き残しあるいは引き過ぎが生じ、復元した音声は歪んだものとなる。スペクトルサブトラクション法におけるこの問題を改善する手法が多数提案されている。しかし、いずれも雑音のパワースペクトルを推定する手法であり、推定を誤る可能性が残されている。

3. Running Spectrum Filtering

周波数領域における雑音除去手法の一つとして、Running Spectrum Filtering による雑音除去手法を提案する。以下にその手法を解説する。

3.1. 雑音と音声のパワースペクトル

白色雑音のように定常的な雑音は、十分に長い時間ではパワースペクトルの分布が一様で、しかも、どの時刻においても同じ強さを持つ。つまり、この雑音は周波数領域においてパワースペクトルの時間的変化は

どの周波数でも無いものとして良い。このような雑音が音声信号に加わった場合、この雑音を取り除くには、スペクトルサブトラクション法が有効である。では、白色雑音と同等のスペクトル構造を持ち、時間的に緩やかな変化をする雑音の場合はどのようになるであろうか。このような雑音にもスペクトルサブトラクション法が利用できるが、スペクトル全体の強さが時間的に緩やかに変化するため、雑音のパワースペクトルに相当する値を正確に求めなければ、雑音のパワースペクトルの引き残しや引き過ぎが生じる(図2)。しかも、雑音のパワースペクトルは無音区間で推定する必要があり、時間的に変化している場合は、正確に雑音のパワースペクトルに相当する値を求めることは困難である。

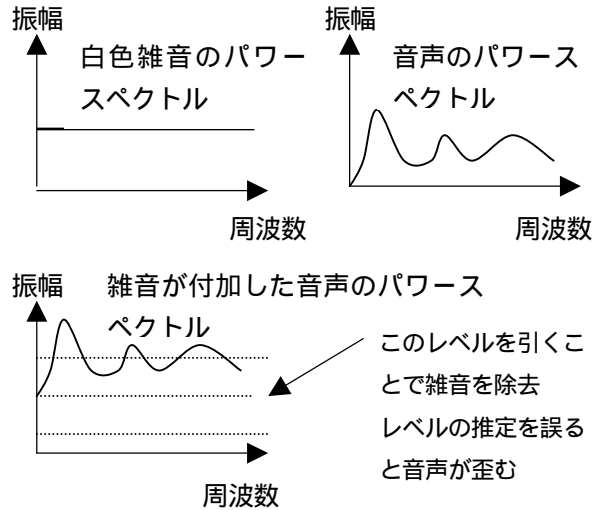


図2 . 周波数領域における雑音除去の様子

3.2. ランニングスペクトル

周波数領域における雑音除去を行う場合、音声信号を時間領域から周波数領域に変換する必要がある。この変換には、音声を短時間で区切り各々DFTにより短時間スペクトルを求め、時間軸上に並べる。つまり、ランニングスペクトルを求める。提案する手法は、このランニングスペクトルの各周波数毎の時間的変化に注目し、雑音除去を行うものである。

まず、定常な白色雑音のランニングスペクトルについて考える。定常な白色雑音は、スペクトル分布が一様で時間的に変化しない。このランニングスペクトルの任意の一周波数の時間変化に注目すると、時間的に変化しない一定の値を持つものとなる。つまり直流とみなす事ができるのである。これはランニングスペクトル全ての周波数に対して言える。次に、白色雑音と同

様のスペクトル構造を持ち時間的に緩やかに変化する雑音について考える。ランニングスペクトルは時間的に緩やかに変化するため、任意の一周波数の時間変化は、非常にゆっくりとしている。この時間変化を一つの信号とみなしてスペクトルを求めると、非常に低い周波数成分を持つ信号と考えられる。次に、音声信号のランニングスペクトルを考える。音声信号は、短時間的に見ると定常であるが、十分長い時間では非定常と考えられ、ランニングスペクトルも時間的に変化したものとなる。従って、このランニングスペクトルの任意の一周波数の時間変化は、一定にはならず雑音に比べ変化が大きい(図3)。つまり、任意の一周波数の時間変化を信号とみなすと高い周波数成分の信号になると考えられる。

提案する手法は、ランニングスペクトルの時間変化を信号とみなし、その信号をフィルタリングすることで時間変化の少ない雑音を除去するものである。

3.3. ランニングスペクトルフィルタ

本報告では、加法性の雑音の除去を対象とする。雑音が付加した音声のランニングスペクトルは、式(1)で示される。

$$Y(\omega, t) = S(\omega, t) + N(\omega, t) \quad (1)$$

ここで、 $S(\omega, t)$ は、時刻 t における音声信号の短時間パワースペクトル、 $N(\omega, t)$ は、時刻 t における雑音の短時間パワースペクトル、 $Y(\omega, t)$ は、時刻 t における雑音が付加した音声信号(入力信号)の短時間パワースペクトルである。白色雑音の場合 $N(\omega, t)$ は、時刻 t によらないスペクトルとなる。

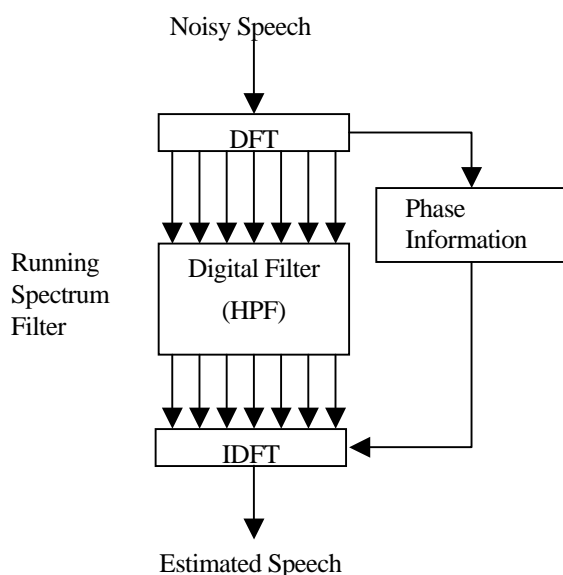


図4. Running Spectrum Filtering のブロック図

この雑音が付加した音声信号のランニングスペクトル $Y(\omega, t)$ に対して、 t に関するデジタルフィルタを施す。これを各周波数 ω について行う。このデジタルフィルタは、非常に低いカットオフ周波数を持つ HPF である。雑音のスペクトル $N(\omega, t)$ が時刻 t に対して変化しないか非常に緩やかに変化する場合、その変化は、直流分と非常に低い周波数成分となる。従って、HPF を通過する事によって $N(\omega, t)$ が除去される。この HPF を Running Spectrum Filter (RSF) と呼ぶ。

この RSF は、直線位相特性を持つ FIR 型デジタルフィルタまたは DFT によって構成されるデジタルフィルタである(図4)。

RSF を通過した短時間パワースペクトルは、各々逆フーリエ変換によって音声信号へ変換される。この時、用いる位相情報は、RSF を通過する前のランニングスペクトルを求める時に得られた位相情報を保存して用いる。

以上のように RSF を用いた雑音除去手法では雑音のパワースペクトルを推定する過程を必要とする事無く雑音除去が可能である。しかも、HPF のカットオフ周波数を変更することにより緩やかに変化する雑音に関しても効果が期待できる手法となっている。

4. 評価実験

提案する RSF を用いた雑音除去手法の有効性を確認するために、雑音を付加した実音声に対する評価実験を行った。

4.1. 実験条件

実験は、白色雑音を付加した実音声に対して行った。付加する雑音は SNR を幾つか変えて行ったが、本報告では SNR=10dB および SNR=0dB について結果を示した。音声は 8kHz, 16bit のサンプリングである。ランニングスペクトルは、短時間スペクトルを 256 ポイント FFT で求めパワースペクトルを計算している。フレームシフトは 20 ポイント(2.5msec), 50 ポイント(6.25msec), 128 ポイント(16msec)とした。このフレームシフト量は、ランニングスペクトルの時間軸方向のサンプリング周期に相当する。RSF の次数は 64 次, 128 次および 256 次の HPF について実験した。HPF のカットオフ周波数は 2.5Hz である。

また、時間的に変化する雑音を付加した実音声について、RSF 法と SS 法の比較を行った。

4.2. 実験結果

実験は、雑音除去処理前と処理後の信号波形およ



(a) 白色雑音を付加した音声信号

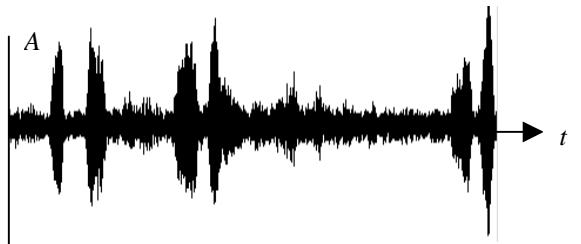


(b) RSF による雑音除去後の音声信号

図5 . SNR=10dB の音声に対する実験結果



(a) 白色雑音を付加した音声信号



(b) RSF による雑音除去後の音声信号

図6 . SNR=0dB の音声に対する実験結果

び試聴の比較を行った . 信号波形の比較について図5 , 図6 に示す . この波形は HPF の次数が 256 次のものを示した .

図5の(a)の波形は , SNR=10dB の雑音除去前の信号波形である . (b)の波形は , RSF を用いて雑音除去した後の音声波形である . 波形上では若干の雑音が残っている様子があるが , 試聴比較では , 白色雑音がほとんど聴こえない状態まで除去された . しかし , 僅かなミュージカルノイズが知覚された .

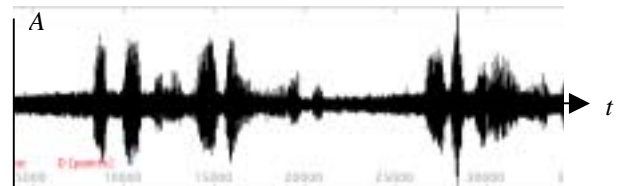
図6の(a)の波形は , SNR=0dB の雑音除去前の信号波形である . (b)の波形は , RSF を用いて雑音除去した後の音声波形である . 波形上では , 完全に雑音を取り除かれてはいないが , 試聴比較では , 白色雑音はほとんど知覚できない . しかし , SNR=10dB に比較してミュージカルノイズが大きく聴こえる .



(a) 時間的に変化する白色雑音を付加した音声信号



(b) RSF 法による雑音除去後の音声信号



(c) SS 法による雑音除去後の音声信号

図7 . 時間的に変化する雑音の除去

時間的に変化する雑音除去効果を SS 法と RSF 法で比較した実験については , 実験結果を図7 に示した . SS 法に比べ RSF 法の方が雑音除去後の雑音の時間変化が少ないことが判る . 試聴実験においても , 雑音除去後の波打つような雑音が SS 法に比べ RSF 法の方が遥かに小さくなることが確認できる .

以上より , RSF を用いた雑音除去手法で , 雑音のパワースペクトルを推定すること無く白色雑音を取り除くことが可能であることが明らかになった . しかし , SNR が低い音声信号に対してはミュージカルノイズが目立つ . 特に , RSF に用いるデジタルフィルタの次数が低い場合にミュージカルノイズの発生が顕著であった . また , 時間的に変化する雑音に対する除去効果も SS 法に比べ大きいことが明らかになった .

5. 考察

スペクトル領域における雑音除去法として RSF を用いた手法を提案した . 試聴比較実験においては , SNR=10dB , SNR=0dB の音声信号に対して高い雑音除去効果が得られている . しかし , SNR=0dB の音声の試聴においては , ミュージカルノイズが目立っている . このミュージカルノイズの発生原因として , 本手法が白色雑音の性質を利用した雑音除去であるためと考えられる . 白色雑音は , 十分に長い時間でスペクトルが一様になり , 時間的変化も少ない . しかし , ランニン

グスペクトルを用いている本手法では，ランニングスペクトルを求める時点で音声信号は，短時間スペクトルを求めるために時間の短い区間に分割されることになる．時間の短い区間に分割されることで白色雑音の性質であるスペクトルが一様，時間的に変化しないという性質が成立しない部分が現れ，ミュージカルノイズ発生の原因の一つになっていると考えられる．このことは，RSFに用いるデジタルフィルタの次数を高くすることでミュージカルノイズが軽減されることから推測される．このデジタルフィルタの次数を高くすることは，十分長い時間でスペクトルが変化しないという白色雑音の性質を捉えることになる．

時間的に変化する雑音に対しては，RSFにHPFを用いることで，比較的緩やかに変化する雑音が除去できる．このことは，比較的緩やかに変化する雑音のランニングスペクトルが時間軸方向では非常に低い周波数成分を持っていたため，HPFによって除去されていることを示している．

6. まとめ

本報告では，周波数領域で行う雑音除去手法として，Running Spectrum Filtering (RSF) を用いた手法を提案した．時間領域における加法性の雑音除去手法として従来より用いられている Spectral Subtract (SS) と比べて，雑音のパワースペクトルを推定することなく雑音除去が可能な手法となっている．従って，無音で雑音のみの区間が無く，雑音のパワースペクトルが推定できない状態でも有効に働く手法である．また，時間的に緩やかに変化する雑音に対しても SS 法に比べ雑音除去効果が大きい．

提案する手法の有効性を示すために，実音声に加法性の雑音を付加し，その雑音を除去する評価実験を行った．SNR が低い音声に関してミュージカルノイズが発生するが，雑音のパワースペクトルを推定すること無く雑音を除去することが可能なことを示した．

今後，本手法において SNR が低い時に顕著になるミュージカルノイズの低減および乗法性の雑音の付加に対する対応を行う予定である．

文 献

- [1] 水町光徳, 赤木正人, マイクロホン対を用いたスペクトルサブトラクションによる雑音除去法, 電子情報通信学会論文誌, J82-A, 4, 503-512, 1999.4
- [2] B. Widrow, " Adaptive Noise Cancelling: Principles and Applications ," Proceedings of the IEEE, vol. 63, pp. 1692--1716, Dec. 1975.

- [3] 中西功, 鳥越良充, 伊藤良生, 副井裕, "変形離散フーリエ変換対を用いた音声雑音除去システムにおける雑音減算係数に関する検討," 信学技報, DSP2002-66, pp.1-6, June 2002.
- [4] BOLL, S. F. *Suppression of Acoustic Noise in Speech Using Spectral Subtraction*. IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-27, pp. 113-120, Apr. 1979.