

2H-2

十分大きなInterpolation間隔を有する
Short Time DFTの出力信号の挙動

岸 政七 古賀裕規
愛知工業大学 情報通信工学科

1. はじめに

Short Time DFT (ST-DFT) は瞬時スペクトラムを求める事を可能とするが、反面処理量が膨大になる欠点を有する。しかし、Interpolation (補間) を適用すればこの欠点を克服することができる^[1]。補間間隔Rはフレーム内サンプル数Nを越える事は一般には許されない。

しかし、音声等入力信号のスペクトラム構造に関して、RがNを十分越えても出力信号に大きな歪みが生じない現象を新たに見出したので報告する^[2]。

2. Short Time DFT

ST-DFTを用いれば時刻nにおける瞬時スペクトラム $\Phi(n)$ は次の様に与えられる。

$$\Phi(n) = \{ \phi_0(n), \phi_1(n), \dots, \phi_{N-1}(n) \}^T \quad (1)$$

ここに $\phi_k(n)$ は周波数インデックスkに対する瞬時スペクトラム $\Phi(n)$ の周波数成分であり、次のように与えられる。

$$\phi_k(n) = \sum_{r=-\infty}^{\infty} x(r) h(n-r) W_N^{-rk} \quad (2)$$

$$W_N^{-rk} = \exp\{-j(2\pi rk/N)\}$$

ただし、 $x(r)$ は時刻rにおける入力信号、 $h(*)$ はウィンド関数、Nはフレーム内サンプル数。

一方、時刻nの出力信号y(n)は式3に示す様にRサンプル毎のスペクトラム成分 $S_k(r)$ から補間したスペクトラム成分にShort Time 逆DFT (ST-IDFT) を用い求める。

$$y(n) = \sum_{k=0}^{N-1} \left\{ \sum_{r=L^-}^{L^+} f(n-rR) S_k(r) \right\} W_N^{nk} \quad (3)$$

$$S_k(r) = \phi_k(rR)$$

ここに $f(*)$ は補間関数、 $L^+, L^- = [n/R] + Q/2, [n/R] - Q/2 + 1$

[M]は実数Mを越えない最大の整数、Qは $f(*)$ のフレーム数。

3. 処理概要

式2における補間間隔Rは $R \leq N$ である場合、入力情報の欠落は生じない。すなわちRサンプル毎にdecimation (間引き) されて求められる瞬時スペクトラムを補間し毎サンプルの瞬時スペクトラムを正しく求めて行く事ができる。

$R > N$ の場合、間引かれる間隔RがNを越えているため $(R - N)$ サンプル時点の情報の欠落が生じる。したがって $R > N$ では欠落した情報を推定補間することになる。補間された瞬時スペクトラム $\Phi(n)$ は推定した情報も含んでいるので、毎サンプルの瞬時スペクトラムが近似的に得られることになる。

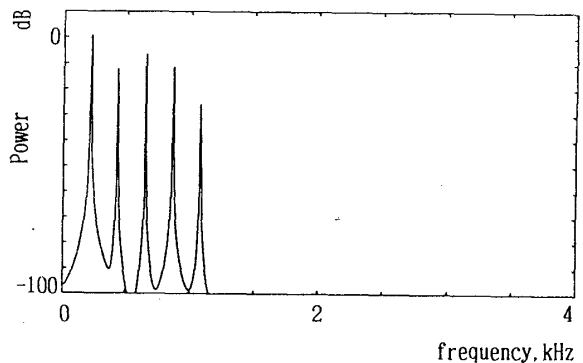
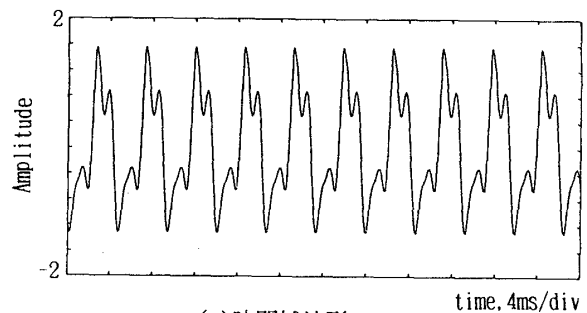


図1 男声“a”の入力信号

Behavior on Output Signals Yielded through Short Time DFT Interpolations

Masahichi KISHI, Hiroki KOGA

Department of Information Network Engineering, A.I.T.

4. 実験結果

$R > N$ の範囲において入力信号 $x(n)$ に対し、合成信号 $y(n)$ がどのような特性を示すのかシミュレーション実験した。入力信号として図 1.(a) に示す男声 "a" の信号を考える。図 1.(a) の入力信号を周波数解析して得られる時刻 n におけるスペクトラムを図 1.(b) に示す。同図に示すように、様々な多くの成分を含む音声信号の瞬時スペクトラムを $\Phi(n)$ とする。瞬時スペクトラムの周波数成分 $\phi_k(n)$ は、式 2 に示す様に ST-DFT を用いて求めることができる。ST-DFT のウィンド関数 $h(*)$ はフレーム数 $m = 8$ とした Nyquist 関数を、補間関数はフレーム数 $Q = 8$ なる Lagrange 関数を用いた。

ここで 2 つのパラメータ δ と κ を定義する。入力信号に含まれる各サブチャネルにおける周波数成分がそのサブチャネルの中心周波数からどれだけ離れているかを表すパラメータを δ とする。サブチャネルの中心周波数とは $2\pi k/N$ である。例えば、 $\delta = 0$ は、入力信号の周波数成分がサブチャネルの中心に、 $\delta = 1$ はサブチャネルの上限フリッジに、 $\delta = -1$ はサブチャネルの下限フリッジ

に存在する事を意味する。また、補間間隔 R とフレーム内サンプル数の比、すなわち R/N を圧縮率 $1/\kappa$ とする。

図 2, 3 は、入力信号の周波数成分に対して δ を 0, ± 1 とし、 κ を 10 とした場合の合成結果である。図 2, 3 はともに (a) は合成信号 (b) はパワースペクトラムについて示している。 $\delta = 1$ においては合成信号の歪みが顕著にみられ、時間域波形ではビート現象が生じる。 $\delta = 0$ のとき、得られる合成信号の最大歪みは -55dB 以下と大きな歪みは認められない。時間域波形においては、入力信号とはほぼ一致する。以上のことから補間間隔 R が分岐点 N を大きく越えても、入力信号の周波数に対して δ を 0 とする事で歪みを低く抑えられることが明らかとなった。

【文献】

- [1] M. Kishi, "A Proposal of Short Time DFT Hilbert Transformers and Its Configuration", Trans. IEICE, vol. E71, No. 5, May 1988, pp. 466-468.
- [2] 古賀, 岸, 後藤, "R > N (R = 1.1N) の場合の Short Time DFT Interpolation 特性", 昭 63 年度電気学会東海支部連合大会, No. 430.

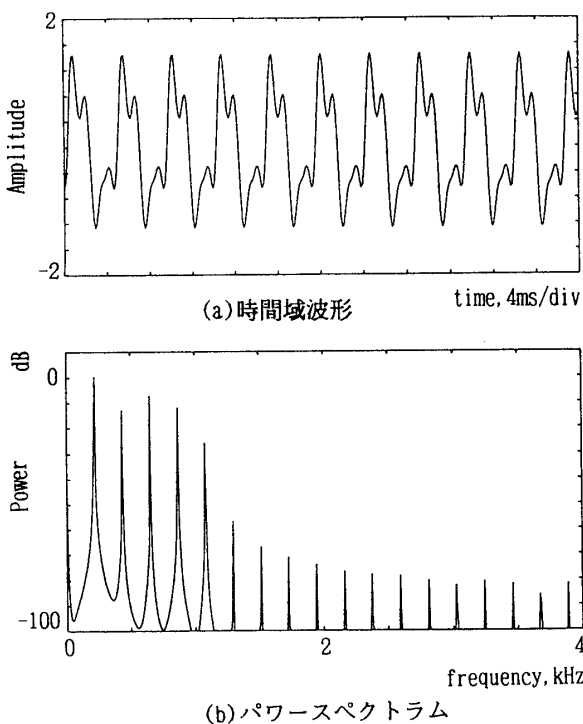


図 2 $\delta=0$ の合成信号

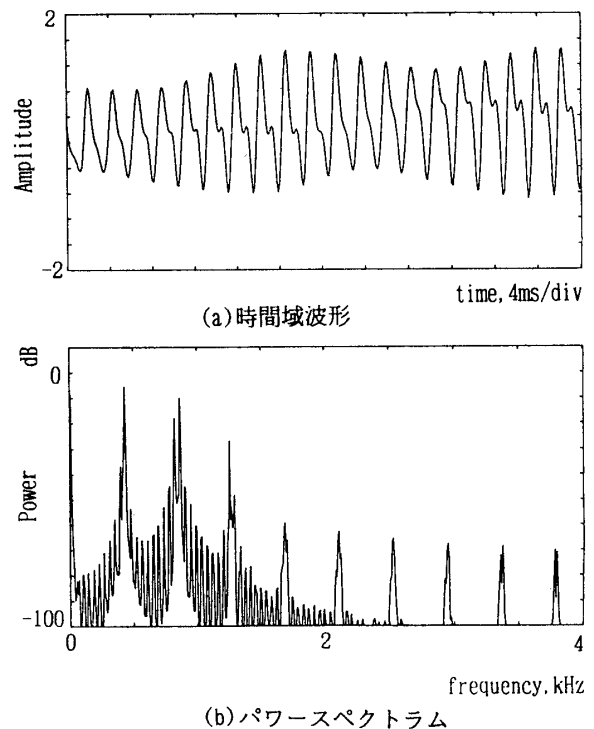


図 3 $\delta=1$ の合成信号