

《小特集》

音響信号処理とウェーブレット変換

阪 本 正 治*・小 林 メ イ*

ABSTRACT Wavelet analysis has emerged as a powerful new tool for the study of non-stationary signals. In contrast to Fourier and Gabor transform methods which use a single analysis window, wavelet methods use short windows at high frequencies and long windows at short frequencies, i.e., constant-Q analysis. In this article, we give an elementary introduction to acoustical signal processing using wavelets and provide examples of some industrial applications.

1. はじめに

音響あるいは音声信号の解析は、FFT（高速フーリエ変換）アルゴリズムの発明を期に急速に発展し、以来、短時間フーリエ変換が、音響信号の時間周波数分析の主役となった。しかし、本来、フーリエ変換は定常信号の分析手法であり、短時間フーリエ変換は便宜的に非定常信号の分析ができるように拡張したに過ぎない。また、その時間一周波数分解能は時間周波数平面上のどこでも一定であり、非定常信号の分析に用いるには不便である。しかし、非定常信号の分析においては、高い周波数領域では信号の始まりの時間を、低い周波数領域ではその周波数成分を詳しく知りたいことが多い。そのような場合は、複数の時間周波数分解能で分析する必要が生じる。

一方、ウェーブレット変換は、本来、非定常信号を解析するための道具として発明された手法である。その基底関数は、時間的にも周波数的にも局在し、しかも、相似性を有しているため、低い周波数では高い周波数分解能が得られ、高い周波数では高い時間分解能が得られるような定Q型分析が可能となる。また、基底関数の相似性は、不連続性の検出にも有利にはたらく。

しかし、定Q型分析それ自体は、ウェーブレット変換が発明される以前から、サブバンドコーディングとして実用化されていたし、人間の聴覚では定Q型分析が行われていることが知られており、そのような

バンドパス特性を持つフィルターバンクも音声認識のフロントエンドとして用いられていた。このように音響処理の分野では、“ウェーブレット的な処理”は決して新しい処理方法ではないが、ウェーブレット変換は時間分解能と周波数分解能の最適配分、基底関数の選択の自由度、逆変換の存在、理論的な明確さなどから、利用が広まってきている。

ウェーブレット変換については、すでに優れた成書^{1)~3)}や解説^{4)~8)}があり、これらを読めばウェーブレット変換について十分理解できると思われるので、本稿では、ウェーブレット変換の数学的定義を述べた後、ウェーブレット変換による音声信号の時間一周波数分析の例、聴覚フィルターとウェーブレット変換の関連、雑音除去、ピッチ抽出、音声認識のフロントエンドとしての適用、音響信号の Classification など適用例を紹介する。

2. ウェーブレット変換

ウェーブレット変換は、時間的にも周波数的にも局在した関数を伸縮あるいは並行移動したものの内積として定義される。ウェーブレット変換の核関数（ウェーブレット基底）であるための必要十分条件はアドミッシブル条件と呼ばれる。それは単に直流成分がゼロというものであり、緩い条件である。そのため、ウェーブレット基底の選択には自由度がある。また、時間一周波数空間でのサンプリングにも豊富なバリエーションがある。

フーリエ解析ではフーリエ変換とフーリエ級数は非常によく似ており、考える区間が有限区間か無限区間かという違いしかない。これに対してウェーブレット解析では、連続ウェーブレット変換とその離散形で、

Wavelet Analysis of Acoustic Signals. By Masaharu Sakamoto and Mei Kobayashi (Tokyo Research Laboratory, IBM Japan, Ltd.).

*日本アイ・ビー・エム株式会社東京基礎研究所

数学的な枠組みや性質がかなり異なっている⁵⁾。連続ウェーブレット変換では、時間的にも周波数的にも局在した基本ウェーブレット関数を用意し、これにスケール変換 a とシフト操作 b を施し、ウェーブレットファミリー $\Psi_{a,b}$ を構成する。

$$\Psi_{a,b} = \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right) \quad (1)$$

ある関数 $x(t)$ のウェーブレット変換は次式のように $x(t)$ との内積で与えられる。

$$CWT_x(a, b) = \frac{1}{\sqrt{a}} \int x(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad (2)$$

ただし、逆変換の存在を保証するためには次式で表される条件(アドミッシブル条件)を満たす必要がある。

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < \infty \quad (3)$$

$\hat{\psi}(\omega)$ は $\psi(t)$ のフーリエ変換である。この条件を満たす時 $\psi(t)$ は、アナライジングウェーブレットあるいはマザーウェーブレットと呼ばれる。この条件は直流分が0であるということにすぎず、アナライジングウェーブレット関数の選択の自由度は高い。 $\psi(t)$ が式(3)を満たせば、ウェーブレット逆変換が次式で与えられる。

$$x(t) = \frac{1}{C_\psi} \int \frac{1}{\sqrt{a}} CWT_x(a, b) \Psi \left(\frac{t-b}{a} \right) \frac{da db}{a^2} \quad (4)$$

連続ウェーブレットでは、基底関数系は直交しないので、 $CWT(a, b)$ を時刻 b におけるスケール a の成分と解釈することはできない^{5),7)}。このことを回避するために、基底関数を直交化したウェーブレットがある。直交化されたウェーブレットでは、 a, b のパラメータはもはや連続的ではあり得ない。離散ウェーブレットでは基底関数を直交系に選ぶと次の展開関係が成り立つ。

$$f(x) = \sum_j \sum_k \alpha_{j,k} \Psi_{j,k}(x) \quad (5)$$

$$\alpha_{j,k} = \int_{-\infty}^{\infty} \Psi_{j,k}^*(x) f(x) dx \quad (6)$$

スケール a とシフト b のは次のように選ばれ、ほとんどの場合、 a_0 の値として2が用いられる。

$$a = a_0^m, \quad b = nb_0 a_0^m \quad (7)$$

離散ウェーブレット変換は、QMF フィルターとして計算でき計算量が少なくてすむ。また、分析—合成系が容易に構成できるのでサブバンドコーディングとして利用される。しかし、スケールとシフトが2のべき乗になり、音響信号の分析に用いるにはサンプリン

グが荒らすぎるため、分析の目的で用いられることは少ない。

3. アナライジングウェーブレットの例

表1, 2にそれぞれ代表的なウェーブレットの一覧を示す^{31),32)}。

表1のウェーブレットは、多重解像度解析から導かれる直交ウェーブレットである。直交ウェーブレットで最も簡単なものは、Haar ウェーブレットであるが、時間軸上で不連続な関数であり使いにくい。Daubechies のウェーブレットは時間軸上でコンパクトで、モーメント N を大きくすると時間軸上でのサポートは広がるが滑らかなウェーブレットが得られる。Meyer のウェーブレットは、フーリエ空間上でコンパクトである。ただし、時間軸上では無限に広がっている。表2は非直交ウェーブレットの一覧である。非直交ウェーブレットは、その制約条件が緩いので無数に考えられるが、ここには、実用に供されることが多いものをあげた。

4. ウェーブレット変換による音声信号の分析

音声信号は、声帯の振動によって断続された気流を音源とする有声母音や、閉じていた口を急激に開くことによって生じるインパルスのような破裂音などがある。有声母音は、周期的に繰り返される比較的定常な音である。一方、破裂音は、10 ms から20 ms という短い時間に観測される音である。(音声信号の生成の仕組みや性質に関しては、例えば、文献8)を参照されたい。) 音声信号のように幅広い特徴を持つ信号を、短時間フーリエ変換を用いて調べる場合は、有声母音の周期性を見たい場合には、窓幅を長くして周波数分解能を高くし、一方、破裂音がどこで起こったかを調べるには窓幅を短くして時間分解能を高くして調べる必要がある。両方の信号を同時に都合よく調べることができない。これに対して、ウェーブレット変換では、その時間—周波数分解能の最適配分によって、両方の信号を同時に都合よく観察できる。

ここで、音声信号をウェーブレット変換した例を図1示す^{10),11)}。図1は、音響信号のウェーブレット変換の絶対値の時間的変化を濃淡によって目で見える形にしたものであり、スカログラム (scalogram) とも呼ばれる。ここでは、アナライジングウェーブレットとして、Cubic Spline ウェーブレットと Gabor ウェーブレットを用いた例をあげている。Cubic Spline ウェ

表1 スケーリング関数とウェーブレット

Type	Scaling Function Formula	Wavelet Formula
Haar	$\phi(x) = \chi_{[0,1]} = \begin{cases} 1 & , x \in [0,1[\\ 0 & , \text{otherwise.} \end{cases}$	$\chi_{[0,1/2]} - \chi_{[1/2,1]}$
Cardinal Splines	$N_m(x) = \frac{x}{m-1}N_{m-1}(x) + \frac{m-x}{m-1}N_{m-1}(x-1)$; $N_1(x) = \chi_{[0,1]}$	$S_m(x) = \sum_n q_n n_m(2x-n)$; $q_n = \frac{(-1)^n}{2^{m-1}} \sum_{l=0}^m \binom{m}{l} N_{2m}(n+1-l)$; $n = 0, 1, \dots, 3m-2$.
Daubechies	generated iteratively $\phi = \lim_{i \rightarrow \infty} \phi_i$ $\phi_i(x) = \sqrt{2} \sum_n h(n) \phi_{i-1}(2x-n)$; $\phi_0 = \chi_{[-1/2, 1/2]}$	for N vanishing moments $\psi(x) = \sum_{n=0}^{2N-1} (-1)^n h_N(-n+1) \phi_N(2x-n)$; $h_N(n)$ are filter coefficients.
Meyer	inverse Fourier transform of $\hat{\phi}_N(\omega)$ $= \begin{cases} 0 & , \omega \geq \frac{4\pi}{3} \\ \cos(\frac{\pi}{2} s_N(\frac{3}{2} \omega -1)) & , \frac{2\pi}{3} \leq \omega \leq \frac{4\pi}{3} \\ 1 & , \omega \leq \frac{2\pi}{3} \end{cases}$ s_N is any corner smoothing function s.t. $= \begin{cases} s_N(x) = 0 & , x \leq 0, \\ s_N(x) + s_N(1-x) = 1 & , 0 \leq x \leq 1, \\ s_N(x) = 1 & , x \geq 1, \end{cases}$ N is the order of s .	$\psi_N = \sum_{k=-\infty}^{\infty} p_{1-k} \phi_N(2t-k)$
Shannon	$\phi(t) = (\sin \pi t)/(\pi t)$	$\psi(t) = (\sin 2\pi t - \cos \pi t)/(\pi(t - \frac{1}{2}))$
Battle-Lemarié	The inverse Fourier transform of $\hat{\phi}_m = \hat{N}_m(\omega)/\sqrt{E_m(e^{-j\omega})}$; $N_m(t)$ is a cardinal B-spline , $E_m(z) = \sum_{k=-m+1}^{m-1} N_{2m}(m+k)z^k$ the Euler-Frobenius-Laurent polynomial.	$\psi_m(t) = \sum_k (-1)^k p_{1-k} \phi_m(2t-k)$; p_k are two scale sequences.

表2 非直交ウェーブレット

Type	Wavelet Formula
Gabor	$\exp(-\alpha^2 t^2) \cdot \exp(j2\pi f_k t)$
Mexican Hat	$M(t) = \frac{2}{\sqrt{3}} \pi^{-1/4} (1-t^2) \exp\left(\frac{-t^2}{2}\right)$
Gammatone	$At^{N-1} \exp(-2\pi b_f t) \cos(2\pi f_c t + \phi)$

ーブレットは実数型であり, Gabor ウェーブレットは複素数型である. 音声信号は英国人が³/stop/と発声した音であり, これを16ビット, 20 KHzでサンプリングしたものである. 分析の範囲は, 6 オクターブであ

り, 1 オクターブを12分割している.

図1から一見してわかるのは, 実数型のウェーブレットである Cubic Spline ウェーブレットでは特徴的な縞模様が現れることである. しかし, この縞模様は, 同じ音声信号を分析しているにもかかわらず, ウェーブレットの形の影響を受けて異なることもわかっている^{10)~12)}. 実数型のウェーブレットは, スカログラムを描くのには向いていない. 一方, 複素数型のウェーブレットでは, 連続的な模様となり, 有声音の周期性による横縞模様と, 破裂音/t/の位置がはっきり現れている.

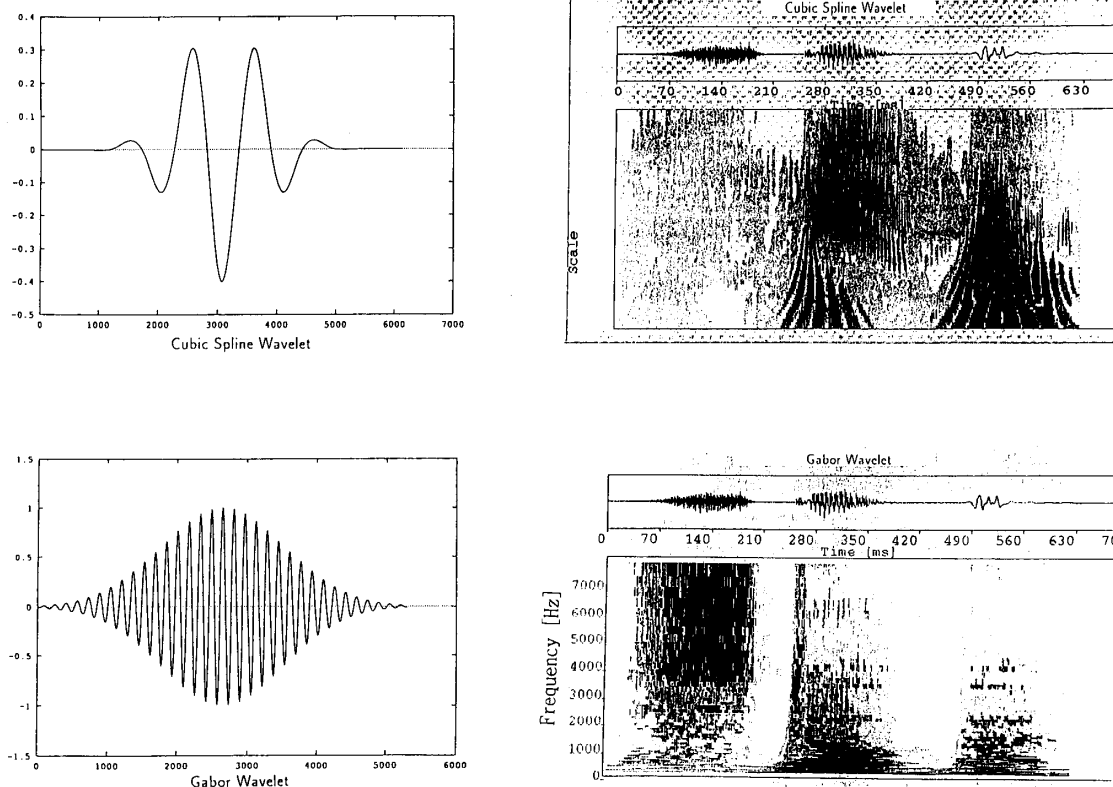


図1 英国人男性による発声“stop”の音声波形とウェーブレット変換（上段左：cubic spline ウェーブレット，上段右：音声波形と cubic spline ウェーブレットによる変換，下段左：Gabor ウェーブレット，下段右：音声波形と Gabor ウェーブレットによる変換）

5. 聴覚フィルターとウェーブレット変換

音響信号は、外耳道、鼓膜、耳小骨を経て、蝸牛内のリンパ液につたわり、基底膜とよばれる膜の振動になる。この膜の振動の振幅が最大となる位置とそこでの最適周波数の対数はほぼ直線的な関係にあり、また心理学的な手段で測定した聴覚フィルターの比帯域幅はほぼ一定になることが知られている¹³⁾。このことは、音響信号が聴覚の初期の段階において定Q型バンドパスフィルター群で分析され、ウェーブレット分析に近い分析が施されていることを意味する。この聴覚フィルターの周波数特性を模擬したフィルターの代表的なものとして、ガンマトーンフィルター (Gammatone filter) があげられる。このフィルターのインパルス応答は次のような式で与えられる。

$$g(t) = at^{N-1} e^{-2\pi b t} \cos(2\pi f_c t), \quad t \geq 0 \quad (8)$$

フィルターのインパルス応答と周波数特性を図2に示す。ここで、パラメータの値は、文献16)を参照した。ガンマトーンフィルターは、アドミッシブル条件を満たしており、基本ウェーブレットとして利用できる。

しかし、単に音響信号を聴覚フィルターで分析するだけであれば、ウェーブレット変換を使う積極的な理由はない。聴覚の研究でウェーブレット変換が使われる理由は、従来の聴覚モデルでは実現できなかった合成系の定義された聴覚モデルを簡単に構築できることにある^{8),15)}。

河原、入野らは聴覚モデルのインパルス応答からアナライジングウェーブレットを構成し、高品質の音声分析合成系を実現している¹⁴⁾。ガンマトーンフィルターは、聴覚による情景理解をモデル化し、例えば、複数の人が同時に話している音響信号から、特定の人の声だけを分離するカクテルパーティー効果など人間の聴覚機能を工学的に実現するための研究や、ノイズが付加された音から音声を抽出するための研究にも用いられている^{16),17)}。

また、入野らは、時間—スケール表現で最小不確定性を有し、聴覚フィルターの周波数特性の非対称特性を模擬できるガンマチャープ関数を提案し、800 Hz以上の帯域ではウェーブレット変換で構成できることを示している^{18),19)}。しかし、同時に800 Hz以下では

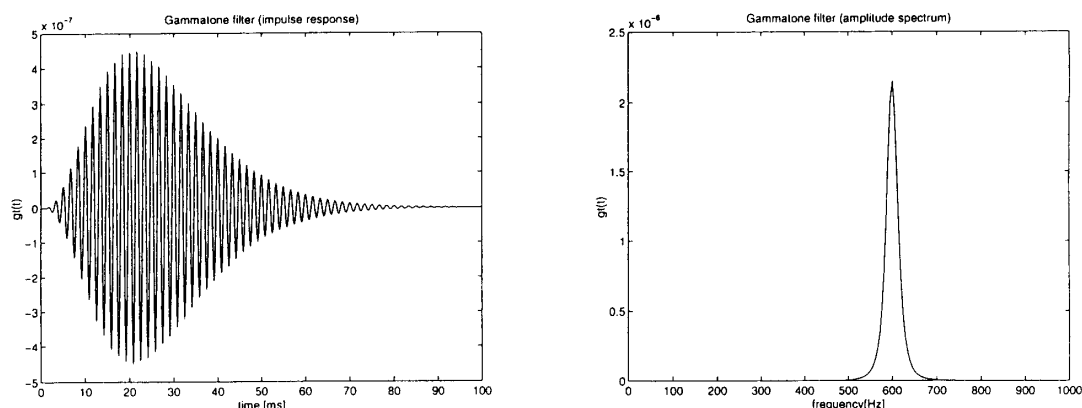


図2 Gammatone filter のインパルス応答と周波数特性（中心周波数：600 Hz, $N=4$, $bf=22.9945$ ）

ウェーブレットとは異なる原理でフィルターバンクを構成する必要があると述べている。

6. 音声強調、雑音除去

背景雑音の存在下で音声による通信や音声認識を行う場合、音声強調/雑音抑制技術が不可欠である。西村らは、従来、短時間フーリエ変換を用いて定義されていたスペクトルサブトラクション法にウェーブレット変換を適用し性能を評価している²⁰⁾。基本ウェーブレットとして、Gabor, Daubechies のほかに、有声音が減衰正弦波で近似できることから、以下のような減衰正弦波を基本ウェーブレットとして用いている。

$$\varphi(t) = e^{-\alpha t} \sin(\omega t), \quad t > 0 \quad (9)$$

この手法による、SN 比の改善の度合いは、Gabor ウェーブレット、減衰正弦波、Daubechies、短時間フーリエ変換の順序となり、また、最大 6 dB の SN 改善があったと報告されている。基底関数の選び方によって、従来のフーリエ変換を用いた手法より上回る結果を得ていることは興味深い。彼らは、ウェーブレット変換を音声強調処理に用いる利点として、基底関数の選択の自由度と定 Q 型分析の 2 点を挙げているが、ウェーブレット関数の時間一周波数平面上での局在性の違いが性能の違いとなって表れているものと考えられる。

7. 音声認識のフロントエンド

短時間フーリエ変換では、時間・周波数分解能が周波数帯域によらず一定であるのに対して、ウェーブレット変換では、低域で高い周波数分解能、高域で高い時間分解能が得られる。筆者らは、このようなウェーブレット変換の特長が音声認識のフロントエンドとして有用なものであるか否かを確かめるために、ウェー

ブレット変換出力を用いた音声の認識実験を行った。従来法との比較のため、特徴量抽出に短時間フーリエ変換出力から求めた臨界帯域フィルター群を用いた場合と認識率の比較を行った²¹⁾。臨界帯域フィルターとは、心理実験で求められた聴覚での周波数分析の帯域幅を模したフィルターであり、ほぼ定 Q 型のフィルターバンクになっている。

基本ウェーブレットとしては、Gabor 関数を選んだ。この関数は、時間一周波数の局在性が最も良い関数として知られている。また、関数が複素正弦波を含んでいるので周波数との対応付けがよい。スケールとシフトは、(4)式に従い、 $a_0 = 2^{1/4}$ とした。これは 1 オクターブをさらに 4 分割することに相当する。

用いた音声は男性一名が発声した単語発声の中から /b/, /d/, /g/ だけを抜き出したものである。認識は、各チャネル出力を 1 状態の無相関混合正規分布でモデル化したものを用いた。フーリエ変換出力から求めた臨界帯域フィルターは定 Q 型分析であり、ウェーブレット変換と似ているが、時間軸上でのサンプリングが異なる。すなわち、フーリエ変換を用いた分析では、どの周波数成分も同じ時間レートで変換結果が出力されるが、ウェーブレット変換では、スケールが小さいほど出力レートが高くなる。このように、チャネルごとに出力レートが異なるような特徴量を、従来の音声認識のモデルで扱うのは難しく、ここでは、時間構造を含まないモデルを用いて、そのような問題を避けた。

結果としては、ウェーブレット変換を用いたときの認識率は、臨界帯域フィルターを用いたときの認識率に及ばなかった。ウェーブレット変換では高域でシフト幅が短くなるとピッチに起因する変動を拾うためか、臨界帯域フィルターを用いた場合と同等の認識率を得るためには、臨界帯域フィルターよりも正規分布

の混合数を増やさなければならない。また、音韻の違いによる認識率の差もウェーブレット変換の方が大きい。

一方、木村らは、各オクターブ間の帯域通過フィルター出力の時間的かつ周波数的な生起関係を表現した、多重解像度制御 HMM を提案し、無声破裂音/p/, /t/, /k/ の認識実験を行い、良好な結果を得たと報告している²²⁾。

8. 音声信号のピッチ抽出

波形が急激に変化する時点は、過渡的な信号や画像の特徴を分析する上で重要である。Mallat ら²³⁾は、アナライジングウェーブレットをスケーリング関数の一次微分を選ぶことによって、ウェーブレット変換の絶対値が局所的な最大値をとる時点が、信号が急激に変化する時点（不連続点）に一致することを示している。不連続点の検出には、連続ウェーブレット変換は冗長であり、計算量も多く必要とする。そこで、スケールを2のべき乗で離散化した2進ウェーブレットを用いている。信号 $x(t)$ の $DyWT$ は、スケール a を2の累乗 2^j ($j=1, 2, \dots$) で離散化することによって得られる。

$$DyWT_x(b, 2^j) = \frac{1}{2^j} \int x(t) \Psi^* \left(\frac{t-b}{2^j} \right) dt \quad (10)$$

先に有声音が声帯の振動を伴うことを説明したが、声道は声帯が閉じることによって呼吸が遮断された時（これを声門閉鎖点と呼ぶ）に駆動され、音声波形が急激に立ち上がる。したがって、音声波形が急峻に立ち上がった時点を検出することによって声門閉鎖点を見つけることができる。声門が閉じた時点から次に声門が開いた時点までが1ピッチ区間であり、1ピッチ区間の時間長の逆数がピッチ周波数である。ピッチ周波数は男性より女性、大人より子供のほうが高く、平均的には女性のピッチ周波数は男性のピッチ周波数の約2倍（1オクターブ）である。さらには、同じ話者でもアクセントや感情によってピッチ周波数は大きく変化する。また、音声波形は声道の共振特性を含んでいるため、非常に複雑な波形をしており、音声波形から直接に声門閉鎖点を見つけたのは困難である。

Kadamba^{24), 25)}らは、音声波形が声門閉鎖点で急峻に変化することに着目し、音声波形の Dyadic Wavelet 変換のローカル・ピークをサーチすることによって声門閉鎖点を抽出し、ピッチ周期を推定する方法を提案した。この方法は、1) 雑音に強く、抽出精度が高い、2) 低ピッチ音声と高ピッチ音声のどちらにも適用できる、3) 計算量が少ない、などの特徴を有する。

筆者らは、テキスト音声合成の音声辞書の作成に Kadamba らの方法を適用し良好な結果を得た²⁶⁾。筆者らは、ウェーブレット変換の絶対値の局所的な最大点を見つけるための閾値の設定にウェーブレット変換出力の統計量を利用することによって、実時間処理向きに変更した²⁷⁾。声門閉鎖点を抽出するために、ピッチ周波数のレンジ（70 Hz-500 Hz）をカバーするように、3つのスケールで2進ウェーブレット変換を計算する。ここで、ウェーブレット関数はスケーリング関数の一次微分となるように選ぶ。各スケールでの2進ウェーブレットのローカルピークのうち、統計量を用いて適応的に設定した閾値を越えた時点を生門閉鎖点の候補とする。これらの候補の中で複数のスケールにおいてその時点が一定の許容偏差内（0.3 ms 程度）に収まっている場合、小さいスケール（時間分解能が高いスケール）での生門閉鎖点の候補を生門閉鎖点とする。このように、複数のスケールで信号を分析し、多数決で生門閉鎖点を決定できる点が、抽出精度を高めている理由であると思われる。図3に男性が/びゅ/と発声したときの音声波形と2進ウェーブレット変換出力および検出された生門閉鎖点（縦線）を示す。表3は、男性話者1名による50文の自由発声に対する、本方式による生門閉鎖点の抽出実験の結果である。

ところで、抽出された生門閉鎖点が原音声の生門閉鎖点に一致していることが望ましい。合成音では、一致することが確かめられている²⁵⁾が、実音声では、実際の生門閉鎖点よりわずかに遅れている。

一方、河原らは複素ウェーブレット変換の位相から瞬時周波数を求め、ピッチ周波数を抽出する方法を考案している²⁸⁾。この方法では、「基本波らしさ」なる指標を導入し、この「基本波らしさ」が最大となるスケールとその隣り合うスケールの位相から求められる瞬時周波数の加重平均として基本周波数を推定している。

9. 音響信号の Classification

音響信号のクラス分けやセグメンテーションは、ウェーブレット変換による重要なアプリケーションの一つである。難聴者には、子音の区別がつきにくい場合が多い。補聴器の高性能化には、音響信号をクラス分けし、それぞれのクラスに対して選択的に音声強調を施すのが良い。Tan ら²⁹⁾は、この目的のために音声信号のウェーブレット変換を求め、各スケールごとの標準偏差を求め、最大尤度比検定を行い音声/非音声のクラス分けをし、その後、音声信号を母音、摩擦性子

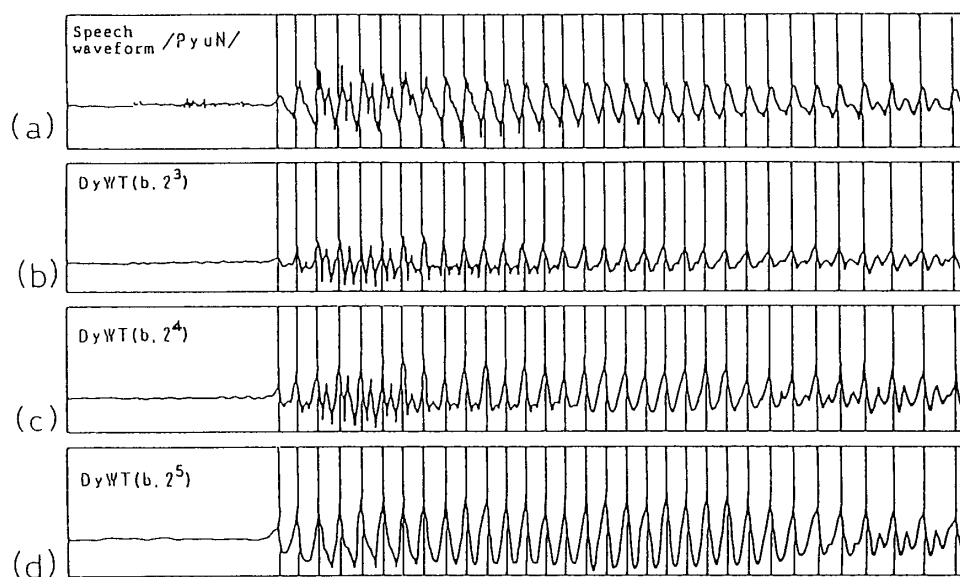


図3 男性による発声“ピュン”の音声波形とウェーブレット変換による1ピッチ区分化(上から音声波形, スケール3でのウェーブレット変換出力, スケール4でのウェーブレット変換出力, スケール5でのウェーブレット変換出力, 図中の縦線が検出された声門閉鎖点)

表3 ピッチ抽出結果²⁶⁾

ピッチ総数	脱落誤り	挿入誤り
20,103	415(2.1%)	155(0.8%)

音, 破裂性音, 無音の自動クラス分けとセグメンテーションを行っている。また, 比屋根ら³⁰⁾は Gabor ウェーブレットを用いて描いたスカログラムのテンプレートマッチングによって, 音響イベントを抽出する方法を提案している。

10. ま と め

本稿では, ウェーブレット変換による音響信号の時間一周波数分析の例, 聴覚フィルタとウェーブレット変換の関連, 雑音除去, ピッチ抽出, 音声認識のフロントエンドとしての適用, Classification などの適用例について解説した。これまでのところ, ウェーブレット変換は, 音響信号処理の分野では, 不連続点の抽出と聴覚フィルタなど非直交フィルタによる分析合成系の構築に威力を発揮しているようである。音響処理の分野では, ウェーブレット変換で劇的な成果は期待できないと思われるが, その特長を生かしたアプリケーションに適用することによって, 良好な結果が得られる。実際, 本稿で, そのような適用例のいくつかを紹介した。

参 考 文 献

- 1) J. M. Combes, A. Grossmann, Ph. Tchamitchian: WAVELETS, Springer-Verlag, Berlin (1989)
- 2) 榊原: ウェーブレット・ビギナズガイド, 東京電機大学 (1996)
- 3) J. J. ベネデット, M. W. フレージャー編, 山口, 山田監訳: ウェーブレット, シュプリンガー・フェアラーク東京 (1995)
- 4) 山田: 講座「ウェーブレット」, 日本シミュレーション学会誌 (1995-1996)
- 5) 山田: ウェーブレット応用(2), ウェーブレットセミナー, 日本応用数理学会 (1991)
- 6) 山口, 守本: ウェーブレットとその応用, 応用数学, 1-3, 202/213 (1990)
- 7) 佐藤: ウェーブレット理論の数学的基礎, 日本音響学会誌, 47-6, 405/423, (1991)
- 8) 入野: Wavelet 変換による音声信号処理, 電子情報通信学会技術研究報告, SP95-6, 39/45, (1995)
- 9) 古井: デジタル音声処理, 東海大学出版会 (1985)
- 10) M. Kobayashi, M. Sakamoto: Wavelet and Speech Signal Analysis, IBM Research Report, RT0085 (1993)
- 11) M. Kobayashi, M. Sakamoto: Wavelet Analysis of Speech Events, 電子情報通信学会技術研究報告, SP93-83, 47/52 (1993)
- 12) 章, 川畑: 連続ウェーブレット変換による時間一周波数解析, 第10回日本計算機統計学会シンポジウム論文集, 24/27 (1995)
- 13) D. Greenwood: A cochlear frequency-position function for several species-29 years later, J. Acoust. Soc. Am., 87-6, 2592/2605 (1990)
- 14) 入野, 河原: 聴覚末梢系表現からの信号再構成, 日本音響学会聴覚研究会資料, H-91-44, 1/8 (1991)

- 15) 河原：ウェーブレット解析の聴覚研究への応用，日本音響学会誌，47-6，424/429，(1991)
- 16) 鵜木，赤木：雑音が付加された波形からの信号波形の一抽出法，電子情報通信学会論文誌，J80-A-3，444/453 (1997)
- 17) K. Nishi, S. Ando, S. Aida: Optimum Harmonics Tracing Filter for Auditory Scene Analysis. Proc. ICASSP, 573/576 (1996)
- 18) 入野俊夫：聴覚末梢系の計算理論，電子情報通信学会技術研究報告，SP95-40，23/30，(1995)
- 19) T. Irino, R. Paterson: A time domain, level-dependent auditory filter: The gammachirp, J. Acoust. Soc. Am, 101-1, 412/419 (1997)
- 20) 西村，浅野，鈴木，曾根：Wavelet 変換を用いた Spectral Subtraction による音声強調，電子情報通信学会誌，J79-A，1986/1993 (1996)
- 21) 阪本，西村：Wavelet 変換を用いた音声認識，日本音響学会講演論文集，175/176，(1994)
- 22) 木村，荒井，城戸：Wavelet 変換の時間一周波数空間における多重解像度制御 HMM，日本音響学会講演論文集，175/176 (1994)
- 23) S. G. Mallat, S. Zhong: Characterization of signal from multiscale edges, IEEE Trans. Pattern Analysis and Machine intelligence, 14-7, 710/732 (1992)
- 24) S. Kadambe, G. F. Boudreaux-Bartels: A comparison of Wavelet function for pitch detection of speech analysis, ICASSP, 449/452 (1991)
- 25) S. Kadambe, G. F. Boudreaux-Bartels: Application of the wavelet transform for pitch detection of speech signals, IEEE transactions of information theory, 38-2, 917/924 (1992)
- 26) 阪本，齊藤，鈴木，小林：波形重畳法を用いた日本語テキスト音声合成システムについて，電子情報通信学会技術研究報告，SP95-6，39/45 (1995)
- 27) 西村，阪本，大嶋，齊藤，鈴木：日本語文音声認識システムを利用した音声圧縮・編集方式の提案，情報処理学会全国大会論文集 (1994)
- 28) 河原，A. Cheveigne：原理的に抽出誤りの存在しないピッチ抽出方法とその評価について，電子情報通信学会技術研究報告，SP96-96，9/18 (1996)
- 29) B. T. Tan, R. Lang, H. Schroder, A. Spray, P. Dermody: Applying wavelet analysis to speech segmentation and classification, SPIE94, Wavelet Applications (1994)
- 30) K. Hiyane, N. Sawabe: Sound recognition using wavelet transform and genetic algorithm, '95 RWC Symposium, 1995または比屋根，澤部：多次元特徴と遺伝アルゴリズムによる信号波形認識，神経回路学会第5回大会，34/35 (1994)
- 31) I. Daubechies: Orthonormal bases of compactly supported wavelets, Comm. Pure Appl. Math., XLI, 909/996 (1988)
- 32) C. K. Chui: Wavelets: A Mathematical Tool for Signal Analysis, SIAM, Philadelphia. (1997)