

マルチレートフィルタと帯域制限補間を用いた ウェーブレット変換による楽音モーフィング

松石 登界 高橋 隆一 吉田 典可

広島市立大学 情報科学部 情報工学科

〒731-3194 広島市安佐南区大塚東 3-4-1

takakuni@lcl.ce.hiroshima-cu.ac.jp,

ryuichi@ce.hiroshima-cu.ac.jp, nyoshida@ce.hiroshima-cu.ac.jp

概要 従来コンピュータには困難とされてきた人間の主観を研究対象とした感性情報処理と呼ばれる研究が活発化している。感性情報処理の研究対象の一つに音楽がある。楽音モーフィングはその研究テーマの一つである。楽音モーフィングには倍音成分の解析が必須である。倍音成分の解析にはウェーブレット変換が有力だと考えられるが、スケールが2では全てを解析できない。ウェーブレット変換にマルチレートフィルタを用いることで、スケールを任意の有理数にすることができる。離散的な数値を扱うため、補間法も重要である。本稿では補間法に帯域制限補間を導入してウェーブレット変換に適用し、マルチレートフィルタを用いたモーフィングを行う手法を提案する。

キーワード ウェーブレット変換, マルチレートフィルタ, 帯域制限補間, 楽音モーフィング

Musical Tone Morphing using Wavelet Analysis with Multi-rate Filters and Band-limited Interpolation

Takakuni MATSUISHI Ryuichi TAKAHASHI Noriyoshi YOSHIDA

Faculty of Information Sciences, Hiroshima City University

3-4-1 Ozuka-higashi, Asaminami-ku, Hiroshima City, 731-3194 Japan

abstract Researches on "Sensibility(KANSEI) Information Processing" which deal with human feeling as a subject of information processing became active. Music and musical tone morphing are research topics in the sensibility information processing. For the musical tone morphing, analysis of harmonic-overtone components is necessary. Wavelet analysis seems to be effective to analyze the components. However, the conventional two scale wavelet analysis can not give an entire solution. By using multi-rate filters, the analysis of the scale of arbitrary rational number can be attained by interpolating wave data using the filters. In this paper, we propose a new wavelet analysis with multi-rate filters that introduces band-limited interpolation for musical tone morphing.

keywords wavelet analysis, multi-rate filters, band-limited interpolation, sound morphing

1 はじめに

近年、コンピュータには困難とされてきた人間の主観的を情報処理の対象として機械と人間のより良い関係を探ることを目的とした感性情報処理と呼ばれる分野の研究が活発化している [6][7][8]。感性情報処理の研究対象の一つとして音楽があげられる。特にコンピュータ音楽は芸術と音響技術、情報科学が融合する分野で発展を遂げている。

コンピュータ音楽、CG、アニメーションなどで必要とされる音響や画像処理研究の一つにモーフィングがある。画像処理のモーフィングでは、人の顔を徐々

に別の形へと変形するデモンストレーション等がある。この処理を楽音に施したものを本稿では楽音モーフィングという。楽音モーフィングとは2つの異なる楽音の間を連続と感じられるように変化させることである。音色や音声については、すでに多くの研究が行われてきた [1] [4][9][10]。しかしモーフィングの定義は曖昧であり、明確な問題設定がなされていない。

筆者らは楽音の性質が基音とその整数倍の周波数である倍音によって構成されていることに注目し、時間周波数解析に優れたウェーブレット変換による楽音のモーフィングを試みた。スケールには通常整数

値の2が用いられるが、これでは2の中乗の周波数しか扱えず、倍音の解析は困難である。

本稿ではマルチレートフィルタを用いたウェーブレット変換 [2][3] による倍音の加工、再構成を提案する。マルチレートフィルタを用いることで信号を任意のサンプリング周波数に変換することが可能となり、ウェーブレット変換に適用することで倍音の解析が可能となる。本研究ではさらに、補間法 [5][11] を工夫して雑音の少ないモーフィングを試みた。モーフィングの結果は、時間周波数領域で適当な処理を施した結果を逆ウェーブレット変換することで得られる。入力信号には管楽器の楽音を用いた。

以下、2節ではツースケールウェーブレット変換について概説する。3節ではマルチレートフィルタを用いたウェーブレット変換と補間法について説明する。4節では楽音モーフィングのアルゴリズムについて説明する。5節では適用例として実際の楽音の倍音抽出、及びモーフィングした波形の例を示す。

2 ツースケールウェーブレット変換

2.1 ウェーブレット

信号 $f(x)$ を時間と周波数の両面から同時にとらえる時間周波数解析において、ウェーブレットは信号の一部分を切り出す単位となる。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ウェーブレットを積分変換の核とする。ウェーブレット変換はウェーブレットにスケール変換とシフト変換を作用させて得られる関数の組を基底関数として用いる。

$$\phi(x) = \sum_{k \in \mathbf{Z}} p_k \phi(2x - k) \quad (1)$$

は2をスケールとするツースケール関係である。ウェーブレットで定まるツースケール数列 p_k に対して、(1)式をスケール関数という。スケール関数を用いてマザー（基本）ウェーブレット $\psi(x)$ を次式で定義できる。

$$\psi(x) = \sum_{k \in \mathbf{Z}} q_k \phi(2x - k)$$

q_k は数列 p_k から定めることができるツースケール数列である。マザーウェーブレット $\psi(x)$ による関数 $f(x)$ のウェーブレット変換は次のように定義される。

$$(W_\psi f)(b, a) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right) f(x) dx \quad (2)$$

a はウェーブレットのスケール、 b はシフト量であり、それぞれ周波数、時間に対応している。

逆にウェーブレット変換から元の信号 $f(x)$ を復元する逆ウェーブレット変換が存在し、次式で与えられる。

$$f(x) = \frac{1}{C_\psi} \int \int_{\mathbf{R}^2} (W_\psi f)(b, a) \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2}$$

ウェーブレット変換 $(W_\psi f)(2^{-j}k, 2^{-j})$ を d_k^j と書き、離散化を行なうことにより式 (2) から次の離散ウェーブレット変換が導かれる。

$$d_k^j = 2^j \int_{-\infty}^{\infty} \psi(2^j x - k) f(x) dx$$

離散逆ウェーブレット変換は次式で与えられる。

$$f(x) \sim \sum_j \sum_k d_k^j \psi(2^j x - k)$$

2.2 時間周波数解析

時間周波数解析を行なう手法としてウェーブレット変換の他に短時間フーリエ変換がある。ウェーブレット変換が短時間フーリエ変換に比べ有利であるのは次の2点である。

- (1) 短時間フーリエ変換の積分核 $\omega(x-b)e^{-i\omega x}$ は信号平面上に一定の面積を占めるが、窓の幅は窓関数のみに依存して解析する信号の位置 b と周波数 ω によらず一定である。ウェーブレット変換でも信号平面上に一定の面積を占めるが、窓の幅が変化する。高い周波数の時は時間解像度を上げることができるので無駄のない解析ができる。
- (2) 短時間フーリエ変換の積分核は基底関数とならない。このことは、分解して得られた信号から元の信号を完全に再構成することができないことを意味する。

2.3 多重解像度解析

スケール関数 ϕ が与えられると、それぞれのレベル $j \in \mathbf{Z}$ について $\{\phi(2^j x - k)\}_{k \in \mathbf{Z}}$ の張る L^2 の部分空間 V_j が決まる。任意の関数 $f_j \in V_j$ は係数 c_k^j を用いて次の形に表すことができる。

$$f_j(x) = \sum_k c_k^j \phi(2^j x - k)$$

ツースケール関係から $V_j \subset V_{j+1}$ である。 ϕ に対応するウェーブレット ψ が与えられると $\psi(2^j x - k)$ の張る L^2 の部分空間 W_j が決まる。任意の関数 $g_j \in W_j$ は係数 d_k^j を用いて次の形に表すことができる。

$$g_j(x) = \sum_k d_k^j \psi(2^j x - k)$$

f_j は一意的に分解でき、再帰的に繰り返すことにより f_j を g_j の和で表すことができる。

$$f_j(x) = g_{j-1}(x) + \dots + g_{j-n}(x) + f_{j-n}(x)$$

レベル j の関数 f_j の解像度は 2^j である。分解によってレベルを下げると、解像度は半分になる。これを繰り返せば f_j のレベルを1つずつ下げるたびに、解像度は半分になる。このように $\{f_j\}_{j \in \mathbf{Z}}$ が解像度の階層構造をもつことを多重解像度解析という。

3 マルチレートフィルタを用いたウェーブレット変換

3.1 マルチレートフィルタ

アナログ信号 $x_a(t)$ を周期 T でサンプルした信号 $x(mT)$ のサンプリング周波数を L 倍に増加させるシステムをインタポレータという。ここで m は整数, L は自然数である。以下ではサンプリング周波数を F_s で表す。

信号 $x(mT)$ と新しいサンプリング間隔 $T' = T/L$ でサンプリングして得られる信号 $y(mT')$ は次式で表される。

$$y(mT') = x(mT/L)$$

アップサンプリングされた信号はフィルタ $H_I(z)$ によりイメージ成分を除去する。

逆にアナログ信号 $x_a(t)$ を周期 T でサンプリングした離散信号 $x(nT)$ のサンプリング周波数を $1/M$ 倍に減少させるシステムをデシメータという。ここで n は整数, M は自然数である。

$x_a(t)$ を新しい周期 $T' = MT$ でサンプリングした信号を $y(nT')$ とすると, $x(nT)$ との関係は次式で表される。

$$y(nT') = x(nMT)$$

エイリアジンを回避するために信号はダウンサンプリングされる前にフィルタ $H_D(z)$ により帯域制限される。

L 倍のインタポレータと $1/M$ 倍のデシメータを組み合わせることで, サンプリング周波数を L/M 倍にすることができる。それぞれのフィルタ $H_I(z), H_D(z)$ から構成されるマルチレートフィルタ $H(z)$ は次のようになる。

$$H(z) = H_I(z)H_D(z) \quad (3)$$

ここで $L > M$ の場合エイリアジンは生じないためインタポレータのフィルタのみで実現できる。

3.2 ウェーブレット変換への適用

マルチレートフィルタを用いてウェーブレット変換を行うと任意の有理数のスケールでの解析が可能となる。ここでは離散ウェーブレット変換のスケールを変える手法について詳述する。

ツースケールウェーブレット変換は図1のように構成される。 $H_L(z)$ は低域通過フィルタで離散信号 $f_j(x)$ から, 低周波成分 $f_{j-1}(x)$ を抽出する。 $H_H(z)$ は高域通過フィルタで離散信号 $f_j(x)$ から, 高周波成分 $g_{j-1}(x)$ を抽出する。

マルチレートフィルタを用いることによりスケールに有理数 v をもつ v スケールウェーブレット変換を実現できる。 v スケールウェーブレット変換は図2のように構成される。この時, マルチレートフィルタはアッ

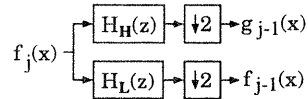


図1: ツースケールウェーブレット変換

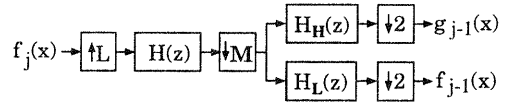


図2: v スケールウェーブレット変換

プサンプリングに用いる。 $v = L/M$ は変換前のサンプリング周波数 F_a と変換後のサンプリング周波数 F_b の比になり, 次式で表される。

$$v = \frac{L}{M} = \frac{F_b}{F_a} \quad (4)$$

任意の有理数 v に対して (4) 式の関係を満たす L, M を適宜選ぶことによって離散ウェーブレット変換の周波数軸におけるスケールを v とすることができる。

逆ウェーブレット変換では図3のように逆の手順を用いる。減算フィルタ $F_S(z)$ にアップサンプリングした高周波成分 $g_j(x)$ を通し, 加算フィルタ $F_A(z)$ にアップサンプリングした低周波成分 $f_j(x)$ を通して $f_{j+1}(x)$ を再構成する。サンプリング周波数 F_a と変

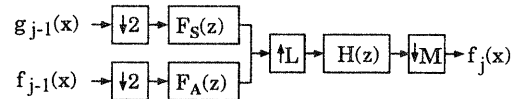


図3: v スケール逆ウェーブレット変換

換後のサンプリング周波数 F_b の関係は (4) 式の逆数になりダウンサンプリングされる。この時 $L < M$ の関係になるが, 再構成後の高周波成分 $g_j(x)$, 低周波成分 $f_j(x)$ に含まれている周波数は変換後のナイキスト周波数より低くエイリアジが生じない。

3.3 補間法

従来筆者らは, インタポレータに線形補間を用いていた [2]。このため, マルチレート信号処理時の信号に大きく誤差が生じていた。本研究では誤差を改善するために, ラグランジュ, スプライン, B スプライン, 帯域制限の各補間法によるインタポレータを用いウェーブレット変換に適用して比較した。

・ラグランジュ補間

一般に,

$$p(x_i) = y_i, \quad (i = k, k+1, \dots, k+l)$$

を満たす l 次多項式によって $f(x)$ を近似する補間法をラグランジュ補間という。具体的な計算方法は、

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_lx^l$$

であり、各係数を連立方程式から求めることで補間関数 $p(x)$ を求める。一般にラグランジュの補間公式、

$$p(x) = \sum_{i=k}^{k+l} y_i N_i(x)$$

$$N_i(x) = \prod_{j=k, j \neq i}^{k+l} \frac{x - x_j}{x_i - x_j}$$

が用いられる。ただし $j=i$ の項は除く。

・スプライン補間

連続性と滑らかさを目標に開発されたのがスプライン関数である。接点の座標 $x_0, x_1, x_2, \dots, x_{n-1}, x_n$ とその関数値 $y_0, y_1, y_2, \dots, y_{n-1}, y_n$ が与えられていると仮定する。また、補間する点 x は区間 $[x_k, x_{k+1}]$ の中にあるとする。この区間における補間式を

$$p_k(x) = a_k x^3 + b_k x^2 + c_k x + d_k$$

とすれば、

$$a_k x_k^3 + b_k x_k^2 + c_k x_k + d_k = y_k$$

$$a_k x_{k+1}^3 + b_k x_{k+1}^2 + c_k x_{k+1} + d_k = y_{k+1}$$

となる。この式から 1 階導関数、2 階導関数を算出し、これらを連立させて a_k, b_k, c_k, d_k について解く。

・B スプライン補間

スプライン補間は与えられた点を通ることが条件だが、B スプライン補間はその点を通らなくてもよい補間法である。データ $(x_i, y_i) (i = 0, 1, \dots, N)$ をあらかじめ決められた節点 $\xi_1, \xi_2, \dots, \xi_n$ をもつ m 階のスプライン関数 $S(x)$ を用いて補間する場合を考える。節点の順序付けは次のように仮定する。

$$\xi_1 < \xi_2 < \dots < \xi_n$$

スプライン関数 $S(x)$ の基底として m 階 ($m-1$ 次) の B スプライン関数を用いる。必要な基底を作るためには $2m$ 個の付加節点 $(\xi_{1-m}, \xi_{2-m}, \dots, \xi_0; \xi_{n+1}, \xi_{n+2}, \dots, \xi_{n+m})$ を導入しなければならない。 $\xi_0 \leq x \leq \xi_{n+1}$ において $S(x)$ は

$$S(x) = \sum_{i=1}^{n+m} c_i^* N_{mi}(x)$$

と表される。 $N_{mi}(x)$ は節点 $\xi_{i-m}, \xi_{i-m+1}, \dots, \xi_i$ に対して定義された m 階の B スプラインである。ここで用いたのは 3 次 B スプライン補間であり、補間式は、

$$x = x_{k-2}q_0(s) + x_{k-1}q_1(s) + x_kq_2(s) + x_{k+1}q_3(s)$$

$$y = y_{k-2}q_0(s) + y_{k-1}q_1(s) + y_kq_2(s) + y_{k+1}q_3(s)$$

ただし、

$$q_0 = s^3/6$$

$$q_1 = (-3s^3 + 3s^2 + 3s + 1)/6$$

$$q_2 = (3s^3 - 6s^2 + 4)/6$$

$$q_3 = (-s^3 + 3s^2 - 3s + 1)/6$$

となる。

・帯域制限補間

周波数 F で帯域制限された連続時間信号 $x(t)$ は、ナイキスト間隔 $1/2F$ を超えない時間間隔 Δt でサンプリングして得た離散時間信号 $x[n]$ によって、

$$x(t) = \frac{1}{2\pi} \int_{-\pi/\Delta t}^{\pi/\Delta t} X_F(\omega) e^{j\omega t} d\omega$$

$$= \sum_{n=-\infty}^{\infty} x[n] \frac{\sin(\pi(t/\Delta t - n))}{\pi(t/\Delta t - n)}$$

のように表される。このサンプリング定理を利用して連続時間信号を復元し、その値から補間値を得る方法を帯域制限補間という。

それぞれの補間法の比較結果を図 4 に示す。楽音 $f(x)$ にはオーボエ (基音は 370Hz) を用いた。マルチレートフィルタを用いたウェーブレット変換によって第 10 倍音 ($g_{-12}(x), f_{-12}(x)$) まで分解した。そのデータをマルチレートフィルタを用いた逆ウェーブレット変換によって $f(x)$ に復元したものである。この図が

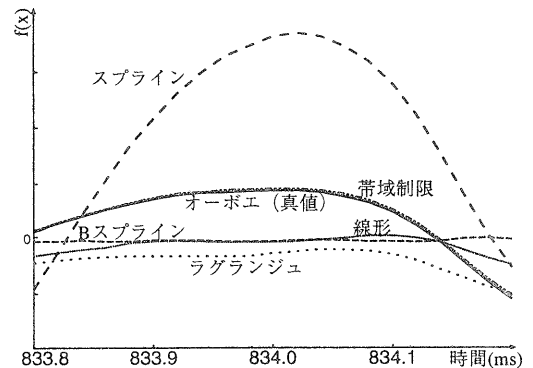


図 4: 各補間法を用いた再構成の結果

ら帯域制限補間が元の信号を最も忠実に復元できていることがわかる。そこで本研究では帯域制限補間を用いて楽音モーフィングを試みた。

4 楽音モーフィングへの適用

4.1 楽音の性質

発音体の振動は一般にいくつかの単純な音 (部分音) の集合だと考えられる。この場合、振動体の基本振動によって生まれる最も振動数の少ない部分音を基音、その他を上音という。基音の振動数の整数倍の振動数を持つ上音を調和倍音あるいは倍音とい

い、倍音によって構成された音を楽音という。非整数倍の上音をもつ場合は噪音という。例えばバイオリンの第4弦の開放弦 $g = 196\text{Hz}$ は、分析によって $392\text{Hz}, 588\text{Hz}, 784\text{Hz}$ などに成分をもつことが知られている。部分音の割合によって音色が左右される。

この性質は、2分の1の解像度をもつツースケールウェーブレット変換では捕らえることができない。しかしマルチレートフィルタを用いたウェーブレット変換では、スケールを任意の有理数にできるため抽出が可能となる。

4.2 モーフィングアルゴリズム

楽音における補間問題では、2つの音の間で対応する物理値を見出す必要がある。筆者らはウェーブレット変換によって楽音の特徴を抽出し、時間、周波数に依存したモーフィングを試みた。

楽音によって倍音成分が異なる。楽音 $g(x)$ と楽音 $h(x)$ のウェーブレット変換結果を、

$$(W_\psi g)(b, a) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right) g(x) dx$$

$$(W_\psi h)(b, a) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right) h(x) dx$$

と表す。 a はウェーブレット関数のスケールで b はシフトであり、それぞれ周波数、時間に対応している。従って $(W_\psi g)(b, a)$ 及び $(W_\psi h)(b, a)$ は、時間、周波数に依存した変数であり、それぞれの楽音の特徴を示した値となる。楽音 $g(x)$ と楽音 $h(x)$ 間のモーフィング音 $f(x)$ を各周波数の特徴を変化するだけでなく時間的にも変化を与えるために、時間、周波数に依存した特徴係数 $U(b, a)$ と、ウェーブレット変換後の結果 $(W_\psi g)(b, a)$ 、 $(W_\psi h)(b, a)$ を用いて次式のモーフィングを行う。

$$(W_\psi f)(b, a) = U(b, a)(W_\psi g)(b, a) + (1 - U(b, a))(W_\psi h)(b, a)$$

$U(b, a)$ の範囲は $0 \leq U(b, a) \leq 1$ とする。構成したいモーフィング音は次式の逆ウェーブレット変換によって生成する。

$$f(x) = \frac{1}{C_\psi} \iint_{\mathbb{R}^2} (W_\psi f)(b, a) \frac{1}{\sqrt{|a|}} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2}$$

離散ウェーブレット変換では、楽音 $g(x)$ の変換結果を g_k^j 、楽音 $h(x)$ の変換結果を h_k^j とすると、構成したいモーフィング音 $f(x)$ のもつ時間周波数成分 d_k^j は、

$$d_k^j = u_k^j g_k^j + (1 - u_k^j) h_k^j \quad (5)$$

となる。ここで、 u_k^j は特徴係数 $U(b, a)$ を離散化したものである。

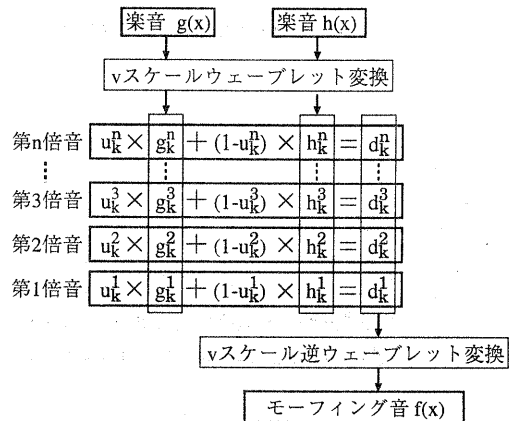


図5: 楽音モーフィング系の構成

5 適用例

5.1 倍音抽出

図5に楽音モーフィング系の構成を示す。2つの楽音（オーボエ、フレンチホルン）をマルチレートフィルタを用いたウェーブレット変換にかけ、第10倍音（3700Hz）まで解析した。補間法には帯域制限補間を用い、音階は $f\#$ （基音は370Hz）、サンプリング周波数は48kHzとした。図6に示すようにオーボエは高周波成分が多い。図7からフレンチホルンは低周波成分を多くもっていることがわかる。

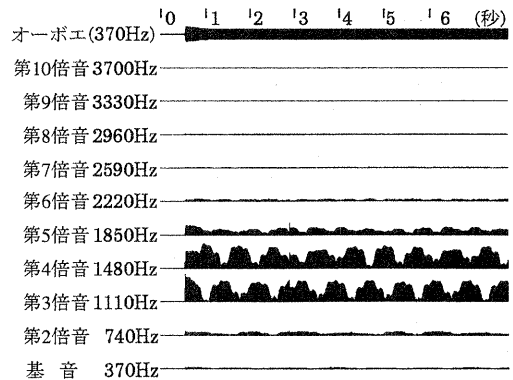


図6: オーボエの楽音成分

5.2 モーフィング音構成

式(5)を用いて、オーボエからフレンチホルンへのモーフィングを試みた。図8に特徴計数 u_k^j を示す。この例では線形に各周波数の割合を変化させた。図9に、オーボエからフレンチホルンへのモーフィング時の倍音成分の変化を、マルチレートフィルタを用

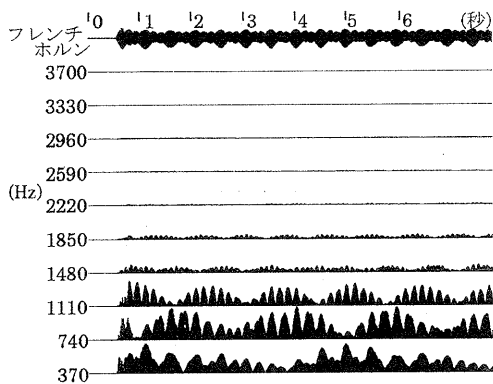


図 7: フレンチホルンの楽音成分

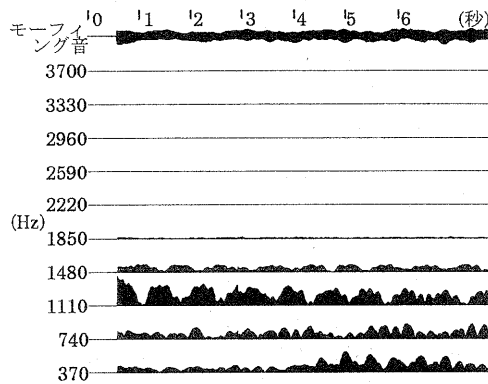


図 9: モーフィングの結果
(オーボエ→フレンチホルン)

いたウェーブレット変換によって追跡した結果を示す。時間を追うごとにフレンチホルンの低周波成分のみが徐々に現れているのがわかる。これは周波数ごとに時間的な変化をもたせた結果である。

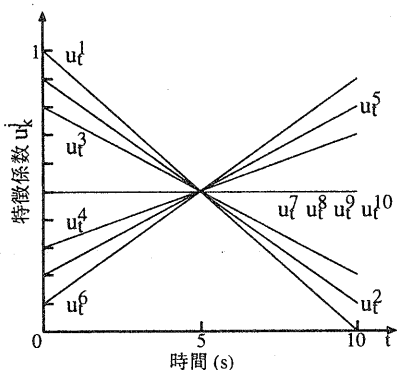


図 8: 特徴係数の変化

6 まとめ

マルチレートフィルタを用いたウェーブレット変換を用いることにより、倍音成分の割合を変化させる楽音モーフィングが行えることを示した。今回は時間的な変化の少ない楽音を対象としたが、ウェーブレット変換では時間周波数解析が可能であり、非定常的な楽音を対象とすることもできる。様々な楽音の解析、定量的評価、感性情報処理への対応付けなどが今後の課題である。

参考文献

[1] 阿部 匡伸: “基本周波数とスペクトルの漸次変形による音声モーフィング”, 音響学会 '95 秋期大会, 2-1-8, pp.259-260, 1995.

[2] 江島 崇, 高橋隆一, 吉田典可: “マルチレートフィルタを用いたウェーブレット変換による感性情報処理～民族音楽と生理指標解析のための一手法～”, 情処研報 音楽情報科学MUS 30-10, pp.45-50, 1999.

[3] 江島 崇, 塚田 健一, 高橋隆一, 吉田典可: “マルチレートフィルタを用いたウェーブレット変換による民族音楽の解析～イスラム圏の音楽における微分音の抽出～第 50 回 中国支部連大 071814 p.193,1999.

[4] 引地 孝文, 小坂 直敏: “打弦, 撥弦, および弾性体振動のモーフィングの検討”, 信学技報, SP96-111, pp.23-28, 1997.

[5] 今井 聖: “信号処理工学”, テレビジョン学会, 1993.

[6] 笠原 裕: “感性のあるコンピュータ”, コンピュートロール No.37, pp. 82-87 コロナ社, 1992.

[7] 方寄晴弘: “コンピュータミュージック” コンピュートロール No.37, pp. 88-93 コロナ社, 1992.

[8] 森安 紀之, 岩永 誠, 高橋 隆一, 吉田典可: “ウェーブレット変換を用いた心拍データの解析”～音楽鑑賞によるリラクセーションを求めて～情処研報 音楽情報科学 MUS25-4, pp21-27,1998.

[9] N.Osaka: “Timbre Interpolation of Sounds using a Sinusoidal Model,” Proc. of ICMC, pp.408-411, 1995.

[10] E.Tellman: “L.Haken and B.Holloway: Timbre morphing of sound with unequal numbers of features,” Journal of the Audio Engineering Society, Vol.43, No.9, 1995

[11] 戸川 隼人: “科学技術計算ハンドブック”, サイエンス社, 1992.