平成 23 年度　数学・数理科学と諸科学・産業との連携研究ワークショップ

ウェーブレット理論と工学への応用

プロシーディングス

Proceedings of the MEXT & OKU 2011 Workshop on

Wavelet Theory and its Applications to Engineering

主催：文部科学省，大阪教育大学

場所：大阪教育大学 天王寺キャンパス

日程：平成 23 年 9 月 12 日（月）13:00 - 18:00

平成 23 年 9 月 13 日（火） 9:00 - 12:30

開催趣旨

このワークショップでは，広い意味でウェーブレット解析によって解決できるかもしれないと期待できるトピックスに関して，講演者の方々に理論と工学的応用の現状，さらに解決すべき問題を解説していただき，その問題提起を端緒として参加者がディスカッションする形で，ウェーブレット解析が実際にどのように応用されているかをより深く理解することによって，新しい理論と応用への道が開かれることを目指します.

ウェーブレット理論と工学への応用プログラム

大阪教育大学 天王寺キャンパス

平成 23 年 9 月 12 日（月）13:00 – 18:00

13:00–14:00　五反田 博（近畿大学 産業理工学部）

独立成分分析に基づく音源分離について ・・・・・ **5**

従来の信号処理技術では解決困難な信号分離問題に対するアプローチとして、独立成分分析（ICA：Independent Component Analysis）がある。ここでは、ICA の基礎から実用化に向けた施策までを体系的に述べて、ICA を実際に適用する上で不可欠なスケールや成分置換の不定性問題の解消法について説明する。更に、音源分離を例に、ICA に既存技術を併用して、より高精度な結果を得るための施策や、実環境に ICA を適用する際の留意事項について言及するとともに、さらなる高性能化に向けて解決すべき問題を提起する。

14:15–15:15　章 忠（豊橋技術科学大学 機械工学系 計測システム）

寄生的離散ウェーブレット変換とその問題点 ・・・・・ **21**

解析信号の中の異常信号を検出、分離するために寄生的離散ウェーブレット変換を提案した。これは異常信号に対応してバンド通過フィルタと阻止フィルタを設計し、それらを従来の離散ウェーブレット変換に適用するものである。しかしバンド通過フィルタの通過領域は従来のハイパスとローパスフィルタの境界と重ねる際に、信号のエネルギー損失が発生しその克服は課題である。

15:45–16:45　三ケ田 均（京都大学 工学研究科 社会基盤工学）

地下構造探査における時系列処理の現状と問題点 ・・・・・ **51**

地下構造の探査において，自己回帰モデルが多用される。このモデルの応用において使われる赤池モデル，バーグモデルの差異とウェーブレットとの関係についての問題を提起し，議論する。

17:00–18:00　新井 康平（佐賀大学大学院 工学系研究科 知能情報システム部門）

ウェーブレットによるコンテンツ保護 ・・・・・ **61**

情報セキュリティの重要性が増してきている。ウェーブレットによるコンテンツ保護への期待も増大している。流通するコンテンツに秘密鍵を埋め込み保護する場合、(1) 流通コンテンツにおける秘密鍵の秘匿性、(2) 流通コンテンツの改竄、処理加工されたとしても秘密鍵が抽出できるかが課題である。これら課題を克服する方法を議論する。

平成 23 年 9 月 13 日（火）9:00 – 12:30

9:00–10:00　守本 晃（大阪教育大学 情報科学）

連続マルチウェーブレット変換に基づく画像分離 ・・・・・ **87**

本講演では，いくつかの画像が混合された複数の観測画像から，元画像を分離する問題を考える．一般に，自然な画像のエッジは区分的に連続な曲線であり，いくつかの元画像のエッジの交わりは点であることが期待できるので，混合画像のエッジを元画像のエッジに分離することができる．この結果から，未知の混合行列を推定し元画像を分離する．そこで，本講演では画像のエッジ抽出を行う変換として，連続マルチウェーブレット変換を提案し，混合画像の分離問題に応用する際の解決すべき問題を提起する．

10:15–11:15　入野 俊夫（和歌山大学 システム工学部）

音声からの声道長推定における聴覚的ウェーブレット変換について ・・・・・ **107**

音声を一声聞くだけで、大人か子供かすぐわかる。同時に話者の寸法に無関係に言語情報も獲得できる。このことから、人間は、寸法（スケール）と声道形状（音韻性）を分離抽出する機構を持っているものと考えている。この聴覚理論として、安定化ウェーブレット-メリン変換を提案してきた。この理論を受けて、最近、２８人の話者間の総当たりで、声道長比を求める実験を行った。１組ずつ数種類の「聴覚的スペクトル」上でスケール変形を行い、最もスペクトル距離が小さい場合を声道長比とした。この時の「聴覚的スペクトル」で、推定精度が最も良かったのは、実際に心理実験から求めた聴覚フィルタを用いた場合であった。これはオーバーコンプリートネスが高い。線形系よりも聴覚的な制約のある非線形性がある方がさらに良いこともわかった。これらの背景と結果を紹介し、聴覚的非線形性も含めた理論的枠組みをぜひ議論していただきたい。

11:30–12:30　井上 勝裕（九州工業大学 情報工学研究院）

脳波によるヒトの状態推定のためのウェーブレット手法の応用 ・・・・・ **131**

ここでは，ウェーブレット変換手法や，その非線形バージョンとも考えることができるモルフォロジカルフィルタを用いた多重解像度解析手法を終夜睡眠脳波の解析や，BCI (Brain Computer Interface) における事象関連電位や誘発電位の抽出に応用した事例を紹介する．なお，モルフォロジカル・ウェーブレットとしての理論体系は確立できておらず，その点に関する問題提起も行いたいと考えている．

大阪教育大学 天王寺キャンパス

〒 543-0054　大阪市天王寺区南河堀町 4-88　電話番号 (06)6775-6611

JR 天王寺駅、地下鉄天王寺駅、近鉄大阪阿部野橋駅下車、徒歩約 10 分。

JR 寺田町駅下車、徒歩 5 分。

独立成分分析に基づく音源分離について

五反田博 *
* 近畿大学　産業理工学部

概要． 従来技術では解決困難な信号分離問題に対する方法として独立成分分析（ICA）がある．ここでは，ICA の基礎から実用化に向けたアプローチを系統的に述べる．また，ICA の応用として雑音除去問題を考え，ICA に既存技術を併用してより良好な除去結果を得るためのアプローチや，実環境に ICA を適用する際の留意事項について言及するとともに，さらなる高性能化に向けて解決すべき問題を提起する．

Sound source separation based on independent component analysis

Hiromu Gotanda*
*Faculty of Humanity-Oriented Science and Engineering, Kinki University

Abstract. Independent component analysis (ICA) attracts much attention as a useful method for signal separation problems having been unsolved by the conventional technologies. In this report, the foundation of ICA is explained and several approaches for its application are systematically described. As an example of ICA application, we consider a noise cancellation problem; some approaches together with conventional technologies are presented for a more highly efficient cancellation, and useful comments are given on the real environmental applications. We also point out several problems to be solved for robust noise cancellation.

1. はじめに

我々の耳には必要な音以外に多数の不必要な音（雑音）も混じって入るが，我々は入り混じった混合音の中から所望の音のみを取り出すことができる．例えば，車の走行音，通行人の話声，店頭から流れる音楽や宣伝アナウンスなどの様々な音が入り混じる喧騒とした街頭で，サイレンを鳴らしながら近づいてくる救急車を見やったり，「チリン」という背後からの音に思わず振り向いて自転車をよける行為は，耳に入る複数の音の中から特定の音（本質的な情報）を取り出していることを如実に示している．このように複数の音の中から特定の声や音を聞き分ける能力（音源分離能力）はカクテルパーティ効果として知られている．

ブラインド信号分離（BSS: Blind Source Separation）は，このように様々な信号が入り混じって観測されたデータから元の信号を人工的に分離する技術の総称で，従来の信号処理技術では解決困難な問題に対する新たな方法として，音源分離，脳波解析，通信路推定，

画像処理，振動解析，時系列予測などの広範な分野で注目されている．BSS のアプローチには，独立成分分析法（ICA: Independent Component Analysis）[1–3]，時間周波数マスキング法（TFM: Time Frequency Masking）[4, 5]，スパースコーディング [6]，非負値行列分解 [7]，ウェーブレット変換に基づく方法 [8] などがある．

ここでは，ICA の基礎から実用化に向けた施策までを体系的に述べて，ICA を実際に適用する上で不可欠なスケールや成分置換の不定性問題の解消法について説明する．また，音源分離を例に，ICA に既存技術を併用してより高精度な結果を得るための施策や，実環境に ICA を適用する際の留意事項について言及するとともに，さらなる高性能化に向けて解決すべき問題を提起する[*1]．

2. ICA に基づくブランド信号分離

独立成分分析法（ICA）は，複数の信号源が統計的に独立であることを前提に，信号源からセンサーまでの伝達特性が未知のもとで，センサーでの観測データのみを用いて元の信号源を推定する統計的方法である．ICA は定常信号だけでなく話者音声や音楽などの非定常信号も分離できるという従来技術にない特徴を持っており，信号源の推定だけでなく，観測データの背後に潜む構造や特徴の抽出にも利用されている．

2.1 混合モデルと分離モデル

統計的に独立な $N(\geq 2)$ 個の信号源 $\boldsymbol{s}(t)=[s_1(t), s_2(t), \cdots, s_N(t)]^T$ から出た信号が N 個のセンサーで

$$\boldsymbol{x}(t) = A\boldsymbol{s}(t) \tag{2.1}$$

と観測される場合（混合モデル）を考える．ここに，$\boldsymbol{x}(t) = [x_1(t), x_2(t), \cdots, x_N]^T$ は観測データ（混合信号）で，T は転置記号を表す．また，A は a_{mn} を要素とする混合行列 $(N \times N)$，a_{mn} は n 番目の信号源から出た信号が m 番目のセンサーに到達するまでの伝達特性を表す未知の定数である．このとき，瞬時混合モデルに対して，分離モデルを

$$\boldsymbol{u}(t) = W\boldsymbol{x}(t) \tag{2.2}$$

と考える．ここに，$\boldsymbol{u}(t) = [u_1(t), u_2(t), \cdots, u_N(t)]^T$ は分離信号，W は w_{nm} を要素とする分離行列 $(N \times N)$ である．

以上のもとで，混合信号 $\{\boldsymbol{x}(t) \mid t = 1, 2, \cdots\}$ だけをデータとして使用して，分離信号が統計的に独立となるように分離行列 W を逐次更新しながら，分離信号 $\{u_n(t) \mid t = 1, 2, \cdots\}$ を生成しようと云うのが ICA アルゴリズムである．

[*1] 本稿は「数理解析研究所講究録 No. 1743」に掲載された内容を一部割愛修正し第 5 節を追加したものである．

2.2 Kullback-Leibler 情報量

信号源 $\boldsymbol{s}$ の同時分布を $p(s_1,\cdots,s_N)$（$=p(\boldsymbol{s})$），周辺分布を $p(s_n)$ と表記する[*2]．このとき，信号源が統計的に独立となるための必要十分条件は，$p(s_1,\cdots,s_N)=\prod_{n=1}^{N}p(s_n)$ と表される．したがって，分離信号 $\boldsymbol{u}$ を統計的に独立となるようにするには，$p(u_1,\cdots,u_N)=\prod_{n=1}^{N}p(u_n)$ と分離信号の同時分布が周辺分布の積と等しくなるように，式 (2.2) の W を更新していけば良いことになる．そこで，分布 $p(\boldsymbol{u})$ と $q(\boldsymbol{u})$ の差を量る尺度して知られている KL（Kullback-Leibler）情報量

$$KL(\boldsymbol{u})=\int p(\boldsymbol{u})\log\frac{p(\boldsymbol{u})}{q(\boldsymbol{u})}d\boldsymbol{u} \tag{2.3}$$

において，$q(\boldsymbol{u})$ を周辺分布の積 $\prod_{n=1}^{N}p(u_n)$ で置き換えて次のように得られる評価式で分離信号 $\boldsymbol{u}$ の独立性を測ることにする．

$$KL(\boldsymbol{u})=\int p(\boldsymbol{u})\log\frac{p(\boldsymbol{u})}{\prod_{n=1}^{N}p(u_n)}d\boldsymbol{u} \tag{2.4}$$

KL 情報量は，$KL(\boldsymbol{u})\geq 0$ と非負の値をとり，$\boldsymbol{u}$ が $p(u_1,\cdots,u_N)=\prod_{n=1}^{N}p(u_n)$ と独立のとき $KL(\boldsymbol{u})=0$ となって最小の値をとる．それゆえ，分離信号が統計的に独立であるか否かは，式 (2.4) の KL 情報量により判断できる．

上述のことより，統計的に独立な分離信号を生成するには，KL 情報量を W に関して最小化すれば良い．そこで以下では，KL 情報量を W の陽な関数として定式化する．式 (2.4) の KL 情報量は，エントロピーを用いて $KL(\boldsymbol{u})=\sum_{n=1}^{N}\mathcal{H}(u_n)-\mathcal{H}(\boldsymbol{u})$ と書き改められる．ここに，$\mathcal{H}(\boldsymbol{u})=-\int p(\boldsymbol{u})\log p(\boldsymbol{u})d\boldsymbol{u}$ と $\mathcal{H}(u_n)=-\int p(u_n)\log p(u_n)du_n$ はそれぞれ分離信号の同時エントロピーと周辺エントロピーである．また，式 (2.2) に基づいて分布の変換を行うと，$p(\boldsymbol{x})=p(\boldsymbol{u})/|W|$ なる関係が得られる．ここに，$|\cdot|$ は行列式を表す記号である．この関係を同時エントロピーの式に代入すると，$\mathcal{H}(\boldsymbol{u})=\mathcal{H}(\boldsymbol{x})+\log|W|$ となって，最終的に，KL 情報量が W の陽な関数として

$$KL(W)=\sum_{n=1}^{N}\mathcal{H}(u_n)-\log|W|-\mathcal{H}(\boldsymbol{x}) \tag{2.5}$$

と表現されることになる．

2.3 自然勾配法

$KL(W)$ を評価として，勾配法（最急降下法）を適用すれば，分離行列 W の更新式が

$$W\leftarrow W-\eta\mathrm{E}[\varphi(\boldsymbol{u})\boldsymbol{u}^T-I]W^{-T} \tag{2.6}$$

[*2] 以降，信号を時系列とみるとき $\boldsymbol{s}(t)$，確率変数とみるとき $\boldsymbol{s}$ のように表記する．

と得られる．ここに，η は探索ステップ幅，I は単位行列，$E[\dot{]}$ は期待値オペレータ，$\varphi(\boldsymbol{u})=[\varphi(u_1),\varphi(u_2),\cdots,\varphi(u_N)]^T$ で，その要素 $\varphi(u_n)$ はスコア関数と呼ばれる非２次的関数である．

ところで，勾配法は，元来，ユークリッド空間（ピタゴラスの定理が成立する空間）における探索法である．しかし，分離行列 W の N^2 個の要素 w_{nm} の張る空間はユークリッド空間ではない．つまり，各要素 w_{nm} のなす軸は互いに直交するユークリッド空間ではなく，曲がった（リーマン）空間である．そこで Amari は [9]，勾配の概念をリーマン空間に拡張し，そこでの W の更新を自然勾配（NG: Natural Gradient）アルゴリズムとして

$$W \leftarrow W - \eta \mathrm{E}[\varphi(\boldsymbol{u})\boldsymbol{u}^T - I]W \tag{2.7}$$

と定式化した．したがって，更新式 (2.7) により $KL(\boldsymbol{u})$ を最小にする解 W を求め，それを式 (2.2) に代入することで，統計的に独立な分離信号 $\boldsymbol{u}(t)$ が生成されることになる．

3. FastICA 法

KL 情報量の式 (2.5)，つまり，$KL(W)=\sum_{n=1}^{N}\mathcal{H}(u_n)-\log|W|-\mathcal{H}(\boldsymbol{x})$ の中で，混合信号のエントロピー $\mathcal{H}(\boldsymbol{x})$ は $KL(W)$ の最小化に寄与しない．また，混合モデルの自由度ゆえに，$KL(W)$ を最小にする分離行列 W は 1 つでなく多数（厳密には無限個）存在する．したがって，W を行列式が $|W|=1$ となる直交行列（$W^TW=I$）のクラスに限定しても差し支えない．この場合，KL 情報量は

$$KL(W) \approx \sum_{n=1}^{N}\mathcal{H}(u_n) \tag{3.1}$$

のように個々の分離信号 u_n の周辺エントロピーの和で近似できる．言い換えると，分離行列を直交行列に絞り込んだ場合，KL 情報量を最小化することは，分離信号の個々のエントロピーを最小化することと等価になる．Hyvärinen は，この考えを発展させて，高速な ICA アルゴリズムとして知られている FastICA 法を以下のように導いた [10] [11].

3.1 混合信号の中心化と白色化

分離行列 W を上述のように直交行列に絞り込めば，個々の分離信号のエントロピーを最小化することで，統計的に独立な分離信号を生成できる．ただし，W を直交行列に絞り込むには，あらかじめ混合信号に対して前処理（中心化と白色化）を行う必要がある．中心化とは $\acute{\boldsymbol{x}}=\boldsymbol{x}-\mathrm{E}[\boldsymbol{x}]$ のように混合信号の平均（中心）を原点に移動させる処理のことである．また，白色化とは，$\mathrm{E}[\acute{\boldsymbol{x}}\acute{\boldsymbol{x}}^T]$ の固有値と固有ベクトルをそれぞれ λ_n と $\boldsymbol{c}_n$ として定義される $\Lambda=\mathrm{diag}[\lambda_1,\cdots,\lambda_n]$ と $\Gamma=[\boldsymbol{c}_1,\cdots,\boldsymbol{c}_n]$ に基づいて，$\tilde{\boldsymbol{x}}=\Lambda^{-1/2}\Gamma^T\acute{\boldsymbol{x}}$ と変換し，$\mathrm{E}[\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T]=I$ となるように規格化する処理のことである．

このとき，原信号 s_m の平均をゼロ，分散を 1 と仮定して，白色化後の混合信号 $\tilde{\boldsymbol{x}}$ と原信号 $\boldsymbol{s}$ の関係を見てみると，

$$\tilde{\boldsymbol{x}} = \Lambda^{-1/2}\Gamma^T A\boldsymbol{s} \tag{3.2}$$

となる．したがって，白色化後の混合行列は，$\tilde{A} = \Lambda^{-1/2}\Gamma^T A$ と表現され，$\tilde{A}\tilde{A}^T = I$ を満たすことから，結果的に直交行列となる．このことは，式 (2.2) の分離モデルを式 (2.1) の混合モデルの逆変換過程とみれば，分離行列 $W = [\boldsymbol{w}_1, \cdots, \boldsymbol{w}_N]^T$ が直交行列のクラスに絞り込めることを示唆している．また，W が直交行列の場合，$\|\boldsymbol{w}_n\|^2 = 1$ となって，探索空間は超曲面に絞り込まれるため，探索アルゴリズム（ICA アルゴリズム）の収束は容易になると考えらる．以上のことから，混合信号に対する前処理は，分離行列 W を直交行列に絞り込むための準備，と位置づけられる．

3.2 FastICA アルゴリズム

中心化と白色化の前処理を行えば，式 (3.1) が成立することから，分離信号の統計的独立性は，個々の分離信号

$$u_n = \boldsymbol{w}_n^T \tilde{\boldsymbol{x}} \tag{3.3}$$

のエントロピー $\mathcal{H}(u_n)$ を最小化することで達成されることになる．しかし，$\boldsymbol{w}_n$ を更新する度に，分離信号の分布を推定して $\mathcal{H}(u_n)$ を求めることや，W が直交行列であることを制約条件に取り込んで最小化を図ることは容易でない．そこで，Hyvärinen は [2]，エントロピー $\mathcal{H}(u_n)$ を最小化する代わりに，

$$J(u_n) = \mathcal{H}(\nu) - \mathcal{H}(u_n) \geq 0 \tag{3.4}$$

と定義されるネゲントロピーを最大化することで，分離信号を統計的に独立させるアプローチをとった．ここに，ν は平均が 0 で分散が 1 のガウス分布に従う確率変数である．ネゲントロピーは，$J(u_n) \geq 0$ と非負の値をとり，u_n がガウス分布のとき最小の 0 となって，u_n の分布がガウス分布から遠ざかるほど大きくなることから，非ガウス性の尺度として利用できる．

ネゲントロピー $J(u_n)$ を近似して，制約条件 $\|\boldsymbol{w}_n\|^2 = 1$ を取り込むと，条件付き評価関数が

$$L(\boldsymbol{w}_n) = \{\mathrm{E}[G(u_n)] - \mathrm{E}[G(\nu)]\}^2 - \beta\{\|\boldsymbol{w}_n\|^2 - 1\} \tag{3.5}$$

のように導かれる．ここに，$G(\cdot)$ はコントラスト関数と呼ばれる非 2 次的関数で，β はラグランジェの未定定数である．そして，式 (3.5) の不動点（$\partial L(\boldsymbol{w}_n)/\partial \boldsymbol{w}_n = 0$ となる点）における関係を求めると，

$$\mathrm{E}[\tilde{\boldsymbol{x}} g(\boldsymbol{w}_n^T \tilde{\boldsymbol{x}})] - \beta \boldsymbol{w}_n = 0 \tag{3.6}$$

なる恒等式が得られる．ここに，$g(\cdot)$ は $G(\cdot)$ の導関数で，式 (2.7) の自然勾配アルゴリズムにおける $\varphi(\cdot)$ に相当するスコア関数である．さらに，この恒等式にニュートン法を適用することで，最終的に，分離荷重 $\boldsymbol{w}_n$ を

$$\boldsymbol{w}_n^+ = E[\tilde{\boldsymbol{x}} g(\boldsymbol{w}_n^T \tilde{\boldsymbol{x}})] - E[g'(\boldsymbol{w}_n^T \tilde{\boldsymbol{x}})]\boldsymbol{w}_n \tag{3.7}$$

$$\boldsymbol{w}_n = \frac{\boldsymbol{w}_n^+}{\|\boldsymbol{w}_n^+\|} \tag{3.8}$$

のように更新する FastICA アルゴリズムが導かれる．この更新式 (3.7)(3.8) は，

$$|\boldsymbol{w}_{n,\mathrm{old}}^T \boldsymbol{w}_{n,\mathrm{new}}| \simeq 1 \tag{3.9}$$

のように更新前後の $\boldsymbol{w}_n$ の方向が一致したとき，収束したと判定される．ここに，添え字の old と new はそれぞれ更新の前後を指す．

上述の手順で，最初 ($n = 1$) に得られる分離荷重 $\boldsymbol{w}_1$ をもとに，$u_1 = \boldsymbol{w}_1^T \tilde{\boldsymbol{x}}$ と生成される分離信号は，N 個の信号源の中で非ガウス性が最大の信号源 s_m を分離したものとなる[*3]．そして，$n = 2$ として得られる分離荷重 $\boldsymbol{w}_2$ をもとに生成される分離信号は，非ガウス性が２番目に大きい信号源を分離したものとなる．以下，信号源は非ガウス性の大きいものから順に分離されることになる．

ただし，$n \geq 2$ の手順においては，分離荷重 $\boldsymbol{w}_n$ が先に推定された分離荷重 $\boldsymbol{w}_i (i \leq n-1)$ と等しくなるのを避けるため，グラムシュミットの方法で

$$\boldsymbol{w}_n = \boldsymbol{w}_n - \sum_{i=1}^{n-1} \boldsymbol{w}_i^T \boldsymbol{w}_n \boldsymbol{w}_n \tag{3.10}$$

のように直交化させて，式 (3.8) で規格化する必要がある．

FastICA アルゴリズムには，以上のように $\boldsymbol{w}_n$ を 1 つずつ更新するアルゴリズム（Deflationary FastICA）の他に，$\boldsymbol{w}_n$ ($n = 1, 2, \cdots, N$) を同時にまとめた行列 $W = [\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_N]^T$ を一括して更新するアルゴリズム（Symmetric FastICA）もある．その詳細については [2] を参照されたい．

4. 周波数領域ＩＣＡと分割スペクトル

本節では音響信号を対象に議論を進める．そのため，信号源は音源，センサーはマイクと読み替える．この場合，2 節で述べた式 (2.1) の瞬時混合モデルは，各音源から出た音波が個々のマイクに同時に到達すると云う非現実的なモデルとなってしまう．現実の環境

[*3] 式 (2.1) の混合モデルには自由度があるため，ICA 解には後述するスケールの不定性や成分置換の問題がある．この成分置換に起因して，最初に得られた分離信号 u_1 は必ずしも信号源 s_1 を反映した信号とはならない．

では，個々の音源からの直接波に加えて壁や天井からの反射波もマイクに入るため，

$$x_m(t) = \sum_{n=1}^{N} a_{mn}(t) * s_n(t) = \sum_{n=1}^{N} \sum_{t'=0}^{T'-1} a_{mn}(t') s_n(t-t') \tag{4.1}$$

のように観測される．ここに，$s_n(t)$ $(n = 1, 2, \cdots, N)$ は音源，$\{x_m(t)|t = 0, 1, 2, \cdots\}$ $(m = 1, 2, \cdots, N)$ は m 番目のマイクでの観測信号（混合信号），$a_{mn}(t')$ は n 番目の音源から m 番目のマイクまでのインパルス応答，t' は遅れ時間，T' はインパルス応答長，$*$ は畳込みを表す．式 (4.1) は時空間混合モデルあるいは時間領域畳込み混合モデルと呼ばれる．

4.1 時間領域 ICA

時間領域畳込み混合モデルに対して，その分離モデルを

$$u_m(t) = \sum_{m=1}^{N} w_{nm}(t) * x_m(t) = \sum_{m=1}^{N} \sum_{t''=0}^{T''-1} w_{nm}(t'')\, x_m(t-t'') \tag{4.2}$$

のようにタップ長が T'' の分離フィルター $w_{nm}(t)$ で構成する．このとき，混合信号のデータ $\{x_m(t) \mid t = 0, 1, 2, \cdots\}$ $(m = 1, 2, \cdots, M)$) から元の音源を式 (4.2) のように分離・復元する方法を時間領域 ICA（TDICA: Time Domain ICA) と云う．

TDICA は，反射が弱く残響時間 T_{60}[sec] [*4]が小さい場合や，T_{60}[sec] が大きくてもマイクから音源までの距離が近い場合，良好に機能する．しかし，マイクが音源から数十センチメートルも離れると，反射の影響を強く受けるため，良好な結果を得るのが難しくなる．これは，例えば，残響時間が T_{60}=50[msec] と小さい場合でも，8KHz サンプリングのときのインパルス応答長は $T' = 400$ となって[*5]，これと同程度のタップ長の分離フィルターを推定しなければならないからである．つまり，逆フィルターのタップ長を単純に $T'' = T'$ と考えたとしても，1 個の逆フィルターにつき 400 個と極めて多数のパラメータ $\{w_{nm}(t'')|t'' = 0, 1, 2, \cdots, T''-1\}$ を推定しなければならず，残響時間が大きくなるほど，その個数は増えることになって，収束が難しくなる[*6]．

4.2 周波数領域 ICA

上述のことから，実環境下では，式 (4.1) を短時間フーリエ変換して得られる周波数領域混合モデル

$$\boldsymbol{x}(\omega_l, k) = A(\omega_l)\boldsymbol{s}(\omega_l, k) \tag{4.3}$$

[*4] 音源の発音が止まってから，残響音が 60dB 減衰するまでの時間.

[*5] インパルス応答長 T' は残響時間 T_{60}[sec] とサンプリング周波数 f_s[Hz] の積で $T' \approx T_{60} f_s$ と近似できる [12].

[*6] インパルス応答は必ずしも最小位相推移とは限らないので，安定な逆フィルタが得られるかと云うことも問題になる.

を考えるのが一般的である．ここに，ω_l と k はそれぞれ後述する l 番目の規格化周波数とフレーム時刻，$A(\omega_l)$ は $a_{mn}(t')$ のフーリエ変換 $a_{mn}(\omega_l)$ を要素とする未知の周波数伝達関数行列，$\boldsymbol{s}_n(\omega_l, k)$ は未知の音源の短時間スペクトルである．また，$\boldsymbol{x}_m(\omega_l, k)$ は，次式で実際に計算して得られる混合信号の短時間スペクトル（以下，混合スペクトルと呼ぶ.）である．すなわち，混合信号 $\{x_m(t) \mid t = 0, 1, 2, \cdots\}$ を L 個ずつ切り出して，$\{x_m(l + kR) \mid l = 0, 1, 2, \cdots, L-1\}$ と得られる k フレーム目のデータを窓がけして，短時間フーリエ変換により，

$$x_m(\omega_l, k) = \sum_{l=0}^{L-1} x_m(l + kR)\, h(l)\, e^{-j\frac{2\pi}{L}lk} \qquad l = 0, 1, 2, \cdots, L-1 \tag{4.4}$$

と計算して混合スペクトル $x_m(\omega_l, k)$ を求める．ここに，$\omega_l = 2\pi l/L$ $(l = 0, 1, 2, \cdots, L-1)$ は規格化周波数，R はフレーム周期 (シフト幅)，$h(l)$ は窓関数，$j = \sqrt{-1}$ である．

また, 式 (4.3) の周波数領域混合モデルに対し, 式 (4.2) を $u_n(\omega_l, k) = \sum_{m=1}^{N} w_{nm}(\omega_l) x_m(\omega_l, k)$ と短時間フーリエ変換して得られる式

$$\boldsymbol{u}(\omega_l, k) = W(\omega_l)\boldsymbol{x}(\omega_l, k) \tag{4.5}$$

を周波数領域分離モデルと定義する．ここに，$\boldsymbol{u}(\omega_l, k)$ は分離信号の短時間スペクトル（以下，分離スペクトルと呼ぶ.），$W(\omega_l)$ は分離行列である．このとき，個々の周波数 ω_l において，混合スペクトル $\{\boldsymbol{x}_m(\omega_l, k) \mid k = 0, 1, 2, \cdots, K\}$ から分離行列 $W(\omega_l)$ を推定して，分離スペクトル $\boldsymbol{u}(\omega_l, k)$ を求め，それを逆短時間フーリエ変換することにより，元の音源に対応する分離信号 $u_n(t)$ $(n = 1, 2, \cdots, N)$ を生成しよう，と云うのが周波数領域 ICA（FDICA: Frequency Domain ICA）である．FDICA の場合，データとして用いる混合スペクトル $\boldsymbol{x}_m(\omega_l, k) \mid k = 0, 1, 2, \cdots, K\}$ や推定すべき分離行列 $W(\omega_l)$ の要素は複素数であるから，2 節で述べた実数版の ICA アルゴリズムは複素数版に変更する必要がある．

まず，周波数領域自然勾配（NG）アルゴリズムの場合，式 (2.7) の複素数版は

$$W(\omega_l) \leftarrow W(\omega_l) - \eta \mathrm{E}[\, \varphi(\boldsymbol{u}(\omega_l, k))\boldsymbol{u}(\omega_l, k)^H - \mathrm{diag}(\, \mathrm{E}[\, \varphi(\boldsymbol{u}(\omega_l, k))\boldsymbol{u}(\omega_l, k)^H)\,]W(\omega_l) \tag{4.6}$$

と与えられる [13]．ここに，$\varphi(\boldsymbol{u}(\omega_l, k)) = \varphi(\, \mathfrak{R}(\boldsymbol{u}(\omega_l, k)) + j\mathfrak{I}(\boldsymbol{u}(\omega_l, k)))$ で，$\mathfrak{R}$ と $\mathfrak{I}$ はそれぞれ実数部と虚数部を表し，H はエルミート転置記号である．

また，周波数領域 FastICA アルゴリズムの場合，式 (3.7)(3.8) の複素数版は

$$\begin{aligned} \boldsymbol{w}_n^+(\omega_l, k) \leftarrow\ & E[\tilde{\boldsymbol{x}}(\omega_l, k)\bar{u}_n(\omega_l, k) g(|u_n(\omega_l, k)|^2)] \\ & -E[g(|u_n(\omega_l, k)|^2) + |u_n(\omega_l, k)|^2 g'(|u_n(\omega_l, k)|^2)]\boldsymbol{w}_n(\omega_l, k) \end{aligned} \tag{4.7}$$

$$\boldsymbol{w}_n(\omega_l, k) \leftarrow \frac{\boldsymbol{w}_n^+(\omega_l, k)}{\|\boldsymbol{w}_n^+(\omega_l, k)\|} \tag{4.8}$$

と与えられる [14]．ここに，$u_n(\omega_l, k) = \boldsymbol{w}_n^H(\omega_l, k)\tilde{\boldsymbol{x}}(\omega_l, k)$ で，$\tilde{\boldsymbol{x}}(\omega_l, k)$ は各周波数ビン ω_l で $\{\boldsymbol{x}(\omega_l, k) \mid k = 1, 2, \cdots, K\}$ に対して中心化と白色化の前処理を行った後の混合スペクト

ルである．また，“$\bar{\ }$”は複素共役を表す．式 (4.6) の $\varphi(\cdot)$ と式 (4.7) の $g(\cdot)$ および $g'(\cdot)$ の比較から分かるように，スコア関数は周波数領域 NG アルゴリズムでは複素数値をとるのに対し，周波数領域 FastICA アルゴリズムでは実数値をとることに注意されたい．

4.3 分割スペクトル

周波数領域 NG や周波数領域 FastICA などの周波数領域 ICA（FDICA）アルゴリズムで分離行列 $W(\omega_l)$ を推定した場合，式 (4.3) に自由度があるため，

$$W(\omega_l)A(\omega_l) = P(\omega_l)D(\omega_l) \tag{4.9}$$

のようにスケールの不定性と成分置換の問題が残る．ここに，$D(\omega_l)$ は対角行列，$P(\omega_l)$ は置換行列（各行と各列において，1 の値をとる一個の要素を除いて，その他の要素はすべてゼロとなる行列）である．言い換えると，$W(\omega_l)$ の推定値を式 (4.5) に代入して $\boldsymbol{u}(\omega_l, k) = W(\omega_l)\boldsymbol{x}(\omega_l, k)$ と生成される分離スペクトルは，必ずしも音源のスペクトルと等しくならない．つまり，必ずしも $u_n(\omega_l, k) = s_n(\omega_l, k)$ とはならず，n 番目の分離スペクトルは $u_n(\omega_l, k)$=$d_i(\omega_l)s_i(\omega_l, k)$ のように n 番目の音源ではなく $i(\neq n)$ 番目の音源を $d_i(\omega_l)$ 倍した値となる．このように分離スペクトルの順番 n と音源のスペクトル順番 i が一致しないことを成分置換という．また，スケール $d_i(\omega_l)$ は周波数ビン l 毎に異なる．これをスケールの不定性という．

したがって，スケールの不定性と成分置換が解消されなければ，式 (4.5) から得られる分離スペクトル $\{\boldsymbol{u}(\omega_l, k) \mid l = 0, 1, \cdots, L-1\ k = 1, 2, \cdots, K\}$ を逆短時間フーリエ変換して時間領域に戻しても，$u_n(t)$ の期待する $s_n(t)$ は復元できない．

ここでは，スケールの不定性と成分置換の問題の本質を明らかにするため，周波数ビン ω_l やフレーム番号 k は外して議論する．また，分離スペクトルは成分が置換されていることを明示するため，分離スペクトル $\boldsymbol{u}$ の成分は $u_{\tilde{n}}$ のように成分番号を示す添字の n に“~”を付けて，$\boldsymbol{u} = [u_{\tilde{1}}, u_{\tilde{2}}, \cdots, u_{\widetilde{N}}]^T$ と表記し直す．補足すると，分離スペクトル $\boldsymbol{u}$ の第 $\tilde{n}$ 番目の成分 $u_{\tilde{n}}$ は N 個の音源 $\{s_1, s_2, \cdots, s_N\}$ のどれか 1 つと排他的に対応するが，どれと対応するか定かでない．つまり，集合 $\{u_{\tilde{1}}, u_{\tilde{2}}, \cdots, u_{\widetilde{N}}\}$ は集合 $\{s_1, s_2, \cdots, s_N\}$ と 1 対 1 の関係にあるが，どれがどれに対応するか，具体的な対応は不明である．

以上の準備のもとで，以下では，Murata ら [15] により

$$\boldsymbol{\xi}_{\tilde{n}} = \boldsymbol{W}^{-1}\,[0,\ \cdots,\ 0,\ u_{\tilde{n}},\ 0,\ \cdots,\ 0]^T \tag{4.10}$$

と定義されるスペクトル $\boldsymbol{\xi}_{\tilde{n}} = \left[\xi_{1\tilde{n}}, \xi_{2\tilde{n}}, \cdots, \xi_{j\tilde{n}}, \cdots, \xi_{N,\tilde{n}}\right]^T$ を活用することで，スケールの不定性や成分置換が解消できることを述べる．便宜のため，このスペクトルを分割スペクトル (Decomposed Spectrum；元の成分に戻されたスペクトル) と呼ぶ．この分割スペクトルについては，第 $\tilde{n}$ 番目の分割スペクトル $\boldsymbol{\xi}_{\tilde{n}}$ の第 m 要素 $\xi_{m\tilde{n}}$ と第 n 番目の音源 s_n との間に次の関係が成り立つ [16] [17] [18].

$$\xi_{m\tilde{n}} = a_{mn}s_n \quad m = 1,\ 2,\ \cdots,\ N \tag{4.11}$$

ここで，$u_{\tilde{n}}$ は，第 n 番目の音源 s_n が分離スペクトル $\boldsymbol{u}$ では順番が入れ替わって第 $\tilde{n}$ 番目の成分として算出されたもので，$\tilde{n}$ は未知であることに注意されたい．つまり，s_n が $\{u_{\tilde{1}}, u_{\tilde{2}}, \cdots, u_{\tilde{N}}\}$ のどれと対応するか判然としないが，とりあえず $u_{\tilde{n}}$ と表記しているにすぎない．したがって，$u_{\tilde{n}}$ から式 (4.10) のように誘導される第 $\tilde{n}$ 番目の分割シンボル $\boldsymbol{\xi}_{\tilde{n}}$ についても，それが具体的に何番目の音源に対応しているか分からない．

式 (4.11) は，たとえ音源の順番 n と分割スペクトルの順番 $\tilde{n}$ の対応が未知でも，分割スペクトル $\boldsymbol{\xi}_{\tilde{n}}$ には以下の性質があることを主張している [17].

1. 分割スペクトルの要素 $\xi_{m\tilde{n}}$ は，音源 s_n から各マイク $(m = 1, 2, \cdots, N)$ への入力分を表している．
2. $\xi_{m\tilde{n}}$ と a_{mn} は，第 2 添字が $\tilde{n}$ と n のように異なるとしても，第 1 添字は同じ m をとる．つまり，成分置換があったとしても，$\xi_{m\tilde{n}}$ の第 1 添字 m は a_{mn} の第 1 添字 m の順番をそのまま継承する．
3. 分離スペクトル $u_{\tilde{n}}$ から分割スペクトル $\xi_{m\tilde{n}}(m = 1, 2, \cdots, N)$ が生成される際のスケール (a_{mn}) は，音源 s_n からマイク x_m $(m = 1, 2, \cdots, N)$ までの伝達特性 (a_{mn}) に等しい．

性質 2 と 3 をまとめると，n 番目の音源から各マイク $(m = 1, 2, \cdots, N)$ への伝達メカニズムは，$u_{\tilde{n}}$ から分割スペクトル $\xi_{m\tilde{n}}(m = 1, 2, \cdots, N)$ を生成するメカニズムに継承される，と云える．言い換えると，例え成分置換が起きていたとしても，分割スペクトルの要素を調べることによって，信号源からセンサーへの（未知の）伝達メカニズム（混合過程）に関する情報が得られる，と云うことになる．また，このことは ICA アルゴリズムの種類に依存しない．

4.4 スケールの不定性と成分置換の是正

スケールの不定性については，分割スペクトルを導入することで以下のように自ずと解決される．式 (4.11) を周波数ビン ω_l やフレーム番号 k を復活して

$$\xi_{m\tilde{n}}(\omega_l, k) = a_{mn}(\omega_l) s_n(\omega_l, k) \quad m = 1,\ 2,\ \cdots,\ N \tag{4.12}$$

と表示すれば分かるように，分割スペクトル $\xi_{m\tilde{n}}(\omega_l, k)$ は音源を $a_{mn}(\omega_l)$ 倍した値となる．厳密に云うと，分割スペクトル $\xi_{m\tilde{n}}(\omega_l, k)$ は，第 n 番目の音源のみを活性させその他の音源を不活性にした状況で，第 m 番目のマイクで観測される値である．この場合，$a_{mn}(\omega_l)$ は，本来，音源とマイク間の周波数特性であることから，各周波数 ω_l でのスケール（倍率）は，音場の周波数特性で規定された値となることに注意されたい．以上のことから，分割スペクトル $\xi_{m\tilde{n}}(\omega_l, k))$ にはスケールの不定性はないと結論づけられる [16] [19].

成分置換についても式 (4.12) に基づいて解決できる．まず，音源とマイクの位置関係が先験的に与えられる場合について述べる．すなわち，簡単のため，2 個の音源と 2 個のマイクが対向して並んでおり，$n = 1$ 番目の音源は $m = 1$ 番目のマイクに近く，$n = 2$ 番目

の音源は $m = 2$ 番目のマイクに近い，と云う先験情報がある場合を考える．この場合，伝達関数 $a_{mn}(\omega_l)$ のゲインと位相について

$$|a_{nn}(\omega_l)| > |a_{mn}(\omega_l)|,\ \angle a_{nn}(\omega_l) > \angle a_{mn}(\omega_l) \quad for\ m \neq n \tag{4.13}$$

なる不等式が成り立ち，これを式 (4.12) に反映させると，

$$|\xi_{11}(\omega_l, k)| > |\xi_{21}(\omega_l, k)|,\ |\xi_{22}(\omega_l, k)| > |\xi_{12}(\omega_l, k)| \tag{4.14}$$

$$\angle\xi_{11}(\omega_l, k) > \angle\xi_{21}(\omega_l, k),\ \angle\xi_{22}(\omega_l, k) > \angle\xi_{12}(\omega_l, k) \tag{4.15}$$

なる関係が得られる [16] [20]．したがって，生成された分割スペクトル $\boldsymbol{\xi}_{\tilde{n}}$ が式 (4.14) のゲイン条件を満たすとき成分置換はないと判定され，満たさないとき成分置換が起きていると判定できる．また，式 (4.15) の位相条件からも同様な判定が行える．ゲイン条件と位相条件は理論上は全く等価であるが，実際の応用では音響信号の伝達特性が周波数帯域で異なるため，帯域毎に使い分けることで精度の高い成分置換の修正が可能となる [20].

次に，音源が音声と雑音の場合，前者の分布は非ガウス的で，後者の分布はガウス分布に近いことが知られている．これをエントロピーの観点から焼き直して分割スペクトルのエントロピーの大小関係として得られる先験情報に基づいて，成分置換を解決する方法も提案されている [21]．この方法は音源とマイクの配置に依存しない点に特徴があり，単一話者の音声の抽出を目的とする雑音除去法としては実用的である．

また，2 つのマイクの中心から見た音源 $s_1(t)$ と $s_2(t)$ の到来方向をそれぞれ $\theta_1(\omega_l)$ と $\theta_2(\omega_l)$ とするとき，これらの推定値は分離行列 $W(\omega_l)$ の逆行列をもとに，

$$\hat{\theta}_1(\omega_l) = \cos^{-1}\left(\frac{c(\angle[W^{-1}(\omega_l)]_{21} - \angle[W^{-1}(\omega_l)]_{11})}{2dF_s\omega_l}\right) \tag{4.16}$$

$$\hat{\theta}_2(\omega_l) = \cos^{-1}\left(\frac{c(\angle[W^{-1}(\omega_l)]_{22} - \angle[W^{-1}(\omega_l)]_{12})}{2dF_s\omega_l}\right) \tag{4.17}$$

と与えられることが式 (4.12) から導かれる．ここに，c は音速，d はマイク間距離，F_s はサンプリング周波数で，$\angle[W^{-1}(\omega_l)]_{nm}$ は $W(\omega_l)$ の逆行列 $W(\omega_l)^{-1}$ の (n, m) 要素の位相である．したがって，到来方向の推定値 $\hat{\theta}_1(\omega_l)$ と $\hat{\theta}_2(\omega_l)$ を用いて，周波数毎に成分置換を是正できる．この方法は先験情報を必要としない点に特徴があり，クラスタリングを適切に行うことで 2 個以上の音源に対して拡張できる．

5. 実環境下での FDICA 適用における留意事項と課題

ここでは，実環境下で FDICA を適用する際，留意すべき事項として，マイク間隔の決定，スコア関数（コントラスト関数）の選定，FDICA 後に残る残留歪の除去等について述べる．また，さらなる高性能化に向けて，今後解決すべき課題（高残響下での適用，移動音源に対する適用，リアルタイム処理化等）について述べる．

5.1 マイク間隔と空間的エリアシング

マイクに入る音（連続信号）をサンプリングして離散信号になおす際，サンプリング周波数 F_s を連続信号の最高周波数 F_{max}) の 2 倍以上にする必要があり，それを満たさない場合，時間的エリアシング（折り返し歪み）が起きることはよく知られている．2 個以上のマイクを使う場合，時間的エリアシングだけでなく，空間的エリアシングも起きる可能性がある．これを回避するには，$d < c/(2F_{max})$ を満たすようにマイク間隔 d を設定する必要がある．ここに，c[m/sec] は音速を表す．例えば，c=340[m/sec] として，最高周波数が F_{max})=4KHz の場合，マイク間隔 d は 4.25cm 未満にする必要がある．

5.2 スコア関数の選定

前述のように，スコア関数としては $\varphi(u) \approx -d \log p(u)/du$ が望ましいが，原信号 s は未知であることから，その復元信号 u の分布 $p(u)$ を正確に求めるのは難しい．そこで，音声を対象とする場合，先験的にもしくは最終的に想定される分布 $p(u)$ がスーパーガウシアンのとき $\varphi(u) = \tanh(u)$，サブガウシアンのとき $\varphi(u) = u^3$ を用いるのが一般的である．しかし，適用分野によって，スコア関数をどのように選ぶかで，ICA の分離能力は大きく異なる．例えば，通信分野のベースバンド信号を対象とする場合，スコア関数は $\varphi(u) = |u|^q sig(u)$ と近似した方が良好な結果が得られる [24].

以上のことから，スコア関数や原信号の確率密度関数自体も含めて推定する Peason ICA や Kerner ICA がある [25] [26]．Peason ICA では，スコア関数を $\varphi(u) = (u - a)/(b_2u^2 + b_1u + B_0)$ とモデル化して，そのパラメータ a, b_0, b_1, b_2 を $-nb_0\mu_(n-1) - (n+1)b_1\mu_n - (n+2)b_2\mu_(n+1)$=$\mu_(n+1) - a\mu_n$ の関係から逐次推定する．ここに，μ_n は n 次のモーメントである．このようにスコア関数の推定や分布の推定を含めた ICA アプローチについても発展が望まれる．

5.3 高残響対策

実環境下での音響信号を対象にする場合，1 フレームを数十 [msec] として，混合信号の短時間スペクトルを求めることが一般的である．この場合，インパルス応答長が 1 フレーム以内に収まれば，式 (4.3) は近似モデルとして十分に意味をなす．しかし，インパルス応答長がフレーム長を越えると，式 (4.3) の近似は崩れる．つまり，周波数領域で瞬時混合モデルとして定式化することが難しくなる．そのため，FDICA 後に成分置換やスケーリングを是正して時間領域信号に戻してもクロストークや残留歪みが残る．暗騒音（方向性のない雑音）や残響の影響が比較的軽い場合，これらのクロストークや残留歪みは，スペクトル差分法やウィーナーフィルタ等の後処理 [27] [28] を施すことにより軽減できる．しかし，残響時間が数百 msec と長くなって，インパルス応答長がフレーム長を大幅に越

える場合，周波数領域での瞬時混合モデルに対して定式化された FDICA は殆ど無力となる．これを改善するため，残響時間が長い場合，周波数領域での畳込みモデル [29] を発展させた方法 [23] や，ICA とは別の観点からのアプローチが種々試みられているが [30]，実用化に向けて解決すべき課題は少なくない．

5.4 移動音源とリアルタイム化

ブラインド信号分離を含む既存の雑音除去技術は，端的に云うと，事前にデータを溜め込んで探索もしくはトレーニングを行うことにより，目的音声や雑音の方向を推定し，目的音声方向に対する指向性を強めるとともに，雑音方向への指向性を弱めることを基本原理に雑音の除去を行っている．そのため，音環境に変化が無く，音源が固定されている場合，数秒間ため込んだマイク収音データをもとにトレーニング（FDICA）を行うことにより，擬似的なリアルタイム化を行うことは可能である．しかし，音源が移動したり雑音源の数が増減したりするなど，音環境が変化する場合，リアルタイム化は難しくなる．したがって，より少ない収音データをもとに音源をロバストに分離する方法の開発が望まれる．

6. 終わりに

本稿では，独立成分分析（ICA）の代表的手法である自然勾配法と FastICA 法について，瞬時混合モデルの枠内で導出原理を述べるとともに，FastICA 法の収束が自然勾配法に比べて速い理由を明らかにした．実環境下で ICA を適用する場合，時間領域畳込みモデルに基づく時間領域 ICA（TDICA）と，周波数領域瞬時混合モデルに基づく周波数領域 ICA（FDICA）の２つのアプローチが考えられるが，アルゴリズムの収束は前者に比べて後者が有利となることを明らかにした．また，ICA 特有の問題として知られるスケールの不定性と成分置換問題に言及し，これらの問題が分割スペクトルを導入することで解決できることを示した．さらに，実環境下で FDICA を適用する際，留意すべき事項として，マイク間隔の決定，スコア関数（コントラスト関数）の選定，FDICA 後に残る残留歪の除去等について述べるるとともに，さらなる高性能化に向けて，今後解決すべき課題（高残響下での適用，移動音源に対する適用，リアルタイム処理化等）について述べた．

参考文献

[1] A. Cichocki and S. Amari: Adaptive blind signal and image processing, learning algorithm and applications; *John Wiley & Sons* (2002).

[2] A. Hyvärinen, J. Karhunen and E. Oja: Independent component analysis; *John Wiley & Sons* (2001).

[3] T. W. Lee: Independent component analysis; *Kluwer Academic* (1998).

[4] O. Yilmaz and S. Rickard: Blind separation of speech mixtures via time-frequency masking; *IEEE Trans. Signal Processing*, Vol. 52, No. 7, pp. 1830-1847 (2004).

[5] K. Fujita: Remarks on a method of blind source separation; *Information*, Vol. 13, No. 3(B), pp. 829-834 (2010).

[6] B. A. Olshausen and D. J. Field: Sparse coding of sensory inputs; *Current Opinion in Neurobiology*, Vol. 14, pp. 481-487 (2004).

[7] A. Cichocki, R. Zdunek and S. Amari: Nonnegative matrix and tensor factorization; *IEEE Signal Processing Magazine* Vol. 25, No. 1, pp. 142-145 (2008).

[8] R. Ashino, T. Mandai and A. Morimoto: Blind source separation of spatio-temporal mixed signals using phase information of analytic wavelet transform; *Int. J. Wavelets Multiresolut. Inf. Process.*, Vol. 8, No. 4 pp. 575-594 (2010).

[9] S. Amari: Natural Gradient Works Efficiently in Learning; *Neural Computation*, Vol. 10, No. 2, pp. 251-276 (1998).

[10] A. Hyvärinen and E. Oja: Independent component analysis: algorithms and applications; *Neural Networks*, Vol. 13, No. 4-5, pp. 411-430 (2000).

[11] A. Hyvärinen: Fast and Robust Fixed-Point Algorithms for Independent Component Analysis; *IEEE Trans. Neural Networks*, Vol. 10, No. 3, pp. 626-634 (1999).

[12] E. A. P. Habets, S. Gannot, I. Cohen and P. C. W. Sommen: Joint Dereverberation and Residual Echo Suppression of Speech Signals in Noisy Environments; *IEEE Trans. Audio Speech and Language Processing*, Vol. 16, No. 8, pp. 1433-1451 (2008).

[13] P. Smaragdis: Blind separation of convolved mixtures in the frequency domain; *Neurocomputing*, Vol. 22, pp. 21-34 (1998).

[14] E. Bingham and A. Hyvärinen: A fast fixed-point algorithm for independent component analysis for complex valued signals; *Int. J. Neural Systems*, Vol. 10, No. 1, pp. 1-8 (2000).

[15] N. Murata, S. Ikeda and A. Ziehe: An approach to blind source separation based on temporal structure of speech signals; *Neurocomputing*, Vol. 41, Issue 1-4, pp. 1-24 (2001).

[16] H. Gotanda, K. Nobu, T. Koya, K. Kaneda, T. Ishibashi, N. Haratani: Permutation correction and speech extraction based on split spectrum through FastICA; *Proc. ICA2003*, pp. 379-384 (2003).

[17] 中河史成，高瀬成史，白土浩，五反田博：ICA による OFDM 周波数オフセットの推定とシンボル復元; 電子情報通信学会論文誌 A, Vol. J91-A, No. 4, pp. 448-457 (2008).

[18] 五反田、石橋孝昭、岩崎宣生、井上勝裕：独立成分分析の基礎と応用；数理解析研究所講究録, No. 1743, pp. 6-20(2011).

[19] K. Nobu, T. Koya, K. Kaneda, N. Haratani and H. Gotanda: Noise Reduction Using Locational Information on Target Sound Source; *J. Robotics and Mechatronics*, Vol. 15, No. 1, pp. 15-23 (2003).

[20] 石橋孝昭, 井上勝裕, 五反田博, 熊丸耕介: 実環境下での伝達特性を利用した周波数領域 ICA の成分置換問題の解決; システム制御情報学会論文誌, Vol. 19, No. 12, pp. 471-478 (2006).

[21] 金田圭市，古屋武志，五反田博：分割スペクトルのエントロピーに基づく成分置換解消法；電子情報通信学会論文誌 A, Vol. J87-A, No. 7, pp. 1065-1069 (2004).

[22] サイビシット ヴィタヤ，木村 哲也，中河 史成，白土 浩，原谷 直実，五反田博：QAM-OFDM における周波数オフセットと伝送路のブラインド推定；電子情報通信学会論文誌 A, Vol. J92-A, No. 3, pp. 141-149 (2009).

[23] 古屋武志, 石橋孝昭, 白土浩, 五反田博: 周波数領域畳込みモデルに基づく高残響環境下での音源分離; システム制御情報学会論文誌, Vol. 22, No. 8, pp. 287-294 (2009).

[24] S. Amari, S. C. Douglas, A. Cichoki, and H. H. Yang: Multichannel blind deconvolution and equalization using the natural gradient; *Proc. on Signal Processing Advance in Wireless Communication Workshop*, pp.101-104, Paris(1997).

[25] J. Karvanen and V. Koivunen: Blind separation methods based on Pearson system and its extensions; *Signal Processing*, Vol. 82, No. 4, pp. 663-673 (2002).

[26] F. R. Bach and M.I. Jordan: Kernel independent component analysis; *J. Machine Learning Research*, Vol. 3, pp. 1-48 (2002).

[27] R. Mukai , S. Araki , H. Sawada and S. Makino: Removal of residual crosstalk components in blind source separation using time-delayed spectral subtraction; *Proc. ICASSP2002*, vol. II, pp. 1789-1792(2002).

[28] K.S. Park, J.S. Son and H.T. Kim: Postprocessing With Wiener Filtering Technique for Reducing Residual Crosstalk in Blind Source Separation; *IEEE Signal Processing Letters*, Vol. 13, No. 12, pp. 749-751(2006).

[29] C. Servière: Separation of speech signals with segmentation of the impulse responses under reverberant conditions; *ICA2003*, pp. 511-516 (2003).

[30] P.A. Naylor and N.D. Gaubitch: Speech Dereverberation; *Springer* (2010).

寄生的離散ウェーブレット変換とその問題点

章　忠 *　大滝　仁 †　今村　孝 *
三宅　哲夫 *　戸田　浩 *
* 豊橋技術科学大学工学部　† 豊橋技術科学大学大学院

概要.　解析信号の中の異常信号を検出、分離するために寄生的離散ウェーブレット変換を提案した．これは異常信号に対応してバンド通過フィルタと阻止フィルタを設計し，それらを従来の離散ウェーブレット変換に適用するものである．しかしバンド通過フィルタの通過領域は従来のハイパスとローパスフィルタの境界と重ねる際に信号のエネルギー損失が発生しその克服は課題である．

Parasitic Discrete Wavelet Transform and its Problems

Zhong Zhang*　Jin Ohtaki†　Takashi Imamura*
Tetsuo Miyake*　Hiroshi Toda*
*Toyohashi University of Technology
†Graduate School of Toyohashi University of Technology

Abstract.　Parasitic Discrete Wavelet Transform (P-DWT) has parasitic filters that are effective for the extraction and the de-noising of abnormal signals has been developed. The P-DWT is a method of using the parasitic filter, which has a band pass filter and a band rejected filter approximated to effective Real-signal Mother Wavelet (RMW) and added to traditional Discrete Wavelet Transform (DWT), for the detecting of abnormal signal. However, the P-DWT has matters loss of precision because it comes up energy loss of frequency element that exists in a specific frequency domain and how improve this drawback becomes problem.

1.　はじめに

従来，非定常信号を解析する手法として，フーリエ変換とウェブレット変換などの時間・周波数解析手法が多く用いられてきた [1]．フーリエ変換は周波数特性の情報を得る代わりに時間情報を失うため，定常的な信号に対しては有効であるが，非定常な信号に対しては有効ではない．そこで，短時間フーリエ変換などの時間・周波数解析が提案された．短時間フーリエ変換は窓関数により信号の一部を切り出してフーリエ変換することにより，時間情報を失わずに周波数解析を行うことができる．しかし，短時間フーリエ変換は窓関数の窓幅が固定されているため，全周波数領域での解析に対して適切な時間・周波数解析を行うことが困難である．解析対象に合わせて窓幅を選択する必要がある他，広範囲の周波数を解析するのには不向きである．これに対して，ウェーブレット変換 (Wavelet

Transform：WT) はマザーウェーブレット (Mother Wavelet：MW) と呼ばれる単一の関数 $\psi(t)$ を，拡大 a と平行移動 b の変数により変形させた関数組 $\psi_{a,b}(t)$ により，各周波数帯域に適した解析を行う [2]．よって WT は短時間フーリエ変換と異なり，窓幅の影響がなく，広範囲の周波数を持つ信号に対しても各周波数領域に適した時間・周波数解析を有効に行うことが可能となった．またこの WT は，変数 a と b を連続変数とする連続ウェーブレット変換 (Continuous wavelet transform, CWT) と，離散変数とする離散ウェーブレット変換 (Discrete wavelet transform, DWT) に分けることができる．CWT においては，MW はアドミッシブル条件 (Admissibility condition) を満たしていればどのような関数でも使える．しかし MW は数多く存在し，その選択により異なる解析結果を得る．そのため，複数の周波数の強度などを評価する場合，MW の選択が困難になる．また DWT においては，MW は信号の再構成を保証するため (双) 直交条件を満たさなければならない [3]．よって DWT に適用する MW は CWT と比較すると数が少なく，その構成も難しい．

一方，WT は一種の相似変換と考え，既知の異常信号により作成した $\psi(t)$ から定義される $\psi_{a,b}(t)$ をスケールとして解析信号 $f(t)$ を計り，$f(t)$ と $\psi_{a,b}(t)$ の相似性を CWT の係数 $w(a,b)$ で数値化できる [4]．著者らはこの特徴を生かし，異常信号による実信号マザーウェーブレット (Real-signal mother wavelet, RMW) の構成法を提案し，それを用いた CWT から得られたスケール $a = 1$ のウェーブレット係数の絶対値をウェーブレット瞬時相関 (Wavelet instantaneous correlation, WIC) と定義した [5, 6]．そしてこれを離散ウェーブレット変換に適用する寄生的離散ウェーブレット変換 (Parasitic Discrete Wavelet：P-DWT) も提案し，ウェーブレット瞬時相関の高速化を実現した [7]．さらに抽出対象の周波数帯を持つ帯域通過型実信号マザーウェーブレット (Band Pass-RMW：BP-RMW) とそれ以外の周波数帯を持つ帯域除去型実信号マザーウェーブレット (Band Rejection-RMW：BP-RMW) を設計し，それを P-DWT に適用することにより異常信号の検出に強く，再構成が可能な寄生的離散ウェーブレット変換が可能となった [8]．

ところが，P-DWT に BP-RWM や BR-RMW を適用したことによって特定の周波数領域に存在する周波数成分のエネルギー損失が発生した．これにより信号抽出の精度が低下し，正確な結果を得ることができないという問題が生じている．本稿ではこれまでに P-DWT の研究成果を紹介し，エネルギー損失の問題に対してその原因と改善策を検討して今後の課題を述べる．

2. 異常信号検出のための RMW とそれを用いるウェーブレット瞬時相関

有限なエネルギーを持つ時系列信号 $f(t)$ の CWT は

$$w(a,b) = a^{-1/2} \int_{-\infty}^{\infty} f(t)\overline{\psi(\frac{t-b}{a})}dt \tag{2.1}$$

と定義される [2]．ここで a ($a > 0$) はスケール，$1/a$ は周波数に対応し，b は時刻のパラメータである．なお，$\overline{\psi(t)}$ は $\psi(t)$ の複素共役である．関数 $\psi(t)$ は MW と呼ばれ，通常アドミッシブル条件 (admissibility condition) を満たさなければならない．しかし $\psi(t)$ が遠方で充分速く零になる場合には，次のような条件を満足すれば，実用上，問題ないことがわかっている [5,6].

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \tag{2.2}$$

この意味で，MW の選択範囲は広く，その構成も DWT の場合より簡単である．

また，式 (2.1) からわかるように，CWT は時間 b と周波数 $1/a$ の 2 次元の平面上で非定常信号の特徴が解析できる．ところが，バンドパスフィルタの特性を有する通常の MW を用いた異常信号の検出と定量評価には，様々な困難が伴う．著者らは，これらの困難を克服するために異常信号から実信号マザーウェーブレット RMW を構成し，それを用いた CWT より得られた，スケール $a = 1$ における $|w(1, b)|$ を，ウェーブレット瞬時相関 (WIC) $R(b)$ と定義した [5,6].

$$R(b) = |w(1, b)| \tag{2.3}$$

なお RMW の構成手順は次のとおりである．

(1) 異常信号から特徴的な部分を切り出し，遠方で充分速く零になるような窓関数を掛け，次に平均値を差し引くことにより DC 成分を取り除き，実数型 RMW $\psi^R(t)$ を構成する．

(2) ノルム $||\psi^R||$ が 1 となるように正規化する．

$$||\psi^R|| = \left[\int_{-\infty}^{\infty} |\psi^R(t)|^2 dt\right]^{1/2} = 1 \tag{2.4}$$

(3) $\psi^R(t)$ をフーリエ変換し周波数スペクトル $\hat{\psi}^R(f)$ を得る．

(4) 周波数領域 $f \leq 0$ において $\hat{\psi}^R(f)$ をゼロとし，また $f > 0$ においては $\hat{\psi}^R(f) = 2\hat{\psi}^R(f)$ とする．

(5) ここで $\hat{\psi}(f)$ の実数部を $\hat{\psi}_r(f)$，虚数部を $\hat{\psi}_i(f)$ で表すことにして，

$$\hat{\psi}_r(f) = \sqrt{(\hat{\psi}_r^R(f))^2 + (\hat{\psi}_i^R(f))^2} \tag{2.5}$$

また $\hat{\psi}_i(f)$ はゼロとし，全周波数成分の位相情報を削除する処理を行う．

(6) 逆フーリエ変換により $\psi(t)= \psi_r(t) + i\psi_i(t)$ が得られる．このように構成された RMW の実数部 $\psi_r(t)$，虚数部 $\psi_i(t)$ は対称性を有する複素数型 RMW であり，対称的複素数型 RMW(Symmetric complex real-signal mother wavelet, SC-RMW) と呼ぶ.

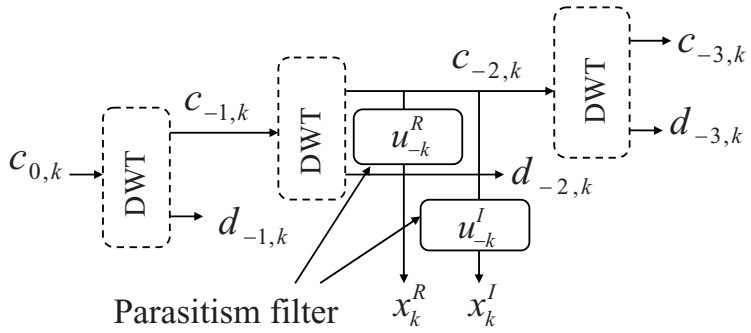


Fig. 1. Decomposition tree of the parasitic discrete wavelet transform (P-DWT) and the parasitic filters

(7) さらに 2 つ，またはそれ以上の SC-RMW を加算し，正規化したものを平均的 RMW(A-RMW) と呼ぶ.

文献 [6] では上述の RMW の詳細を紹介し，さらに RMW を用いた WIC が異常信号検出に有用であることを確認している.

3. 寄生的離散ウェーブレット変換によるウェーブレット瞬時相関の高速化

3.1 寄生的離散ウェーブレット変換と寄生フィルタ

DWT に適用する MW は信号の再構成を保証するため (双) 直交条件を満たさなければならない．しかし異常信号から構成された RMW はこの条件を満たさず，そのまま DWT に適用できない．著者らは異常信号の検出と評価を目的とした，異常信号から構成された RMW に近似したフィルタを，通常の DWT に付与させる新たな離散ウェーブレット変換を提案した [7, 8]．この RMW に近似したフィルタを寄生フィルタ，寄生フィルタが付けられた DWT を寄生的離散ウェーブレット変換 (P-DWT) と呼ぶことにする.

Fig. 1 には提案した P-DWT のツリー構造を示す．図中の破線に囲まれた部分は従来の (双) 直交の MW を用いる多重解像度解析による DWT，または著者らから提案した完全シフト不変複素数離散ウェーブレット変換（Perfect Translation Innariance Complex Discrete Wavelet Transform：PTI-CDWT) である [10]．ここで今後の議論を簡単にするために，前者の DWT と後者の PTI-CDWT を区別せず，DWT と呼ぶことにする．また RMW と区別するため，DWT に用いた MW をベース MW と呼ぶことにする．そして $c_{j,k}$，$d_{j,k}$ は DWT により得られたウェーブレット係数 (高周波成分) とスケーリング係数 (低周波成分)，x^R_k, x^I_k は寄生フィルタにより検出される異常信号の実部と虚部であり，$\{u^R_k\}, \{u^I_k\}$ はそれぞれ RMW に近似した寄生フィルタの実数部，虚数部である．また寄生フィルタが付与されるレベルを寄生レベルと呼ぶことにする．Fig. 1 の例では寄生レベルは -2 であるが，異常信号の特性に応じて他のレベルに，また複数の寄生フィルタを付与させること

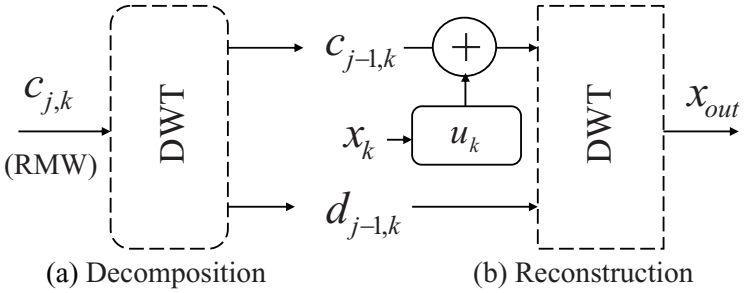


Fig. 2. Design method of the parasitic filters

も可能である．

また Fig. 1 からわかるように，P-DWT のツリー構造は寄生フィルタがなければ，従来の DWT のツリー構造とまったく同じ構成となる．すなわち，P-DWT は従来 DWT に対し，寄生フィルタを付与させて異常信号を検出するので，寄生フィルタは DWT の計算には影響を与えない．また寄生フィルタにより検出された異常信号を再構成する必要はないので，信号の再構成を保証するための (双) 直交条件を満たす必要はない．従って寄生フィルタの設計も簡単となる．

寄生フィルタは Fig. 2 に示すツリー構造を利用して設計を行う．設計の手順は以下のとおりである．

(1) 第 2 章で示した手法で構成した RMW を DWT に入力し寄生レベルまで分解する．
(2) 寄生レベルで得られた係数 $c_{j-1,k}$ を寄生フィルタの初期係数 $\{u_k\}$ にセットする．
(3) Fig. 2 の再構成部を利用し，入力信号を $c_{j-1,k}$=0，$d_{j-1,k}$ =0，$x_k = \delta_{k,0}$ (Kronecker delta) として逆変換を行い再構成信号 x_{out} を求める．
(4) 評価関数 $\mathrm{argmin}\|x_{out} - RMW\|$ が最小になるように最適化手法により $\{u_k\}$ を更新し，RMW に最適近似する寄生フィルタ $\{u_k\}$ が得られる．
(5) 複素数 RMW $\psi(t) = \psi_r(t) + i\psi_i(t)$ の場合には，RMW の実数部および虚数部に対応する寄生フィルタ $\{u_k^R\}, \{u_k^I\}$ を以上の手順に従って，それぞれ設計する必要がある．

ここで，例として式 (3.1) に示した 50[Hz]，100[Hz]，200[Hz] の 3 つの成分を有するモデル信号を用い，

$$f(t) = \sin(100\pi t) + 0.7\sin(200\pi t) + 0.7\sin(400\pi t) \tag{3.1}$$

サンプリング周波数を 3500[Hz]，RMW の長さを 512 点にして第 2 章で示した手法で SC-RMW を構成した．そしてベース MW を Symlet5 として，この RMW を寄生レベル −2 まで分解し，上述の手順で寄生フィルタの設計を行った．Fig. 3(a) は RMW の虚数部を元にして設計した寄生フィルタ $\{u_k^I\}$ の例である．また比較のため，RMW の虚数部の例を Fig. 3(b) に示す． さらに RMW $\psi(t)$ の周波数特性 $|\hat{\psi}(f)|$，$x_{out,k}^R$ と $x_{out,k}^I$ の周波数特性 $|X_{out}^R(f) + iX_{out}^I(f)|$ を求め，Fig. 4 に示す．また Fig. 4 には寄生レベルにおける $c_{j,k}$ を得るための分解フィルタ（ベース MW に対応）の周波数特性 $|A(f)|$ も示している．Fig. 4 によ

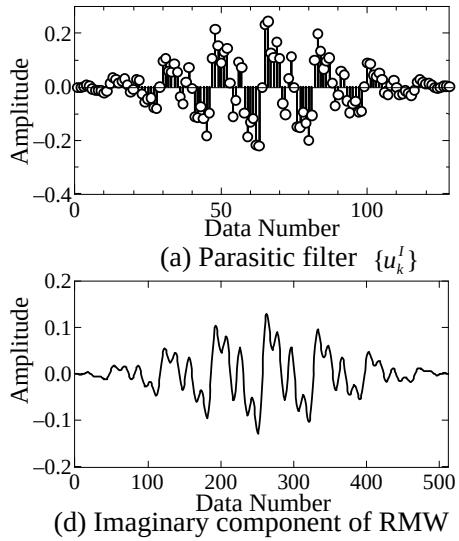


Fig. 3. Example of the designed parasitic filters

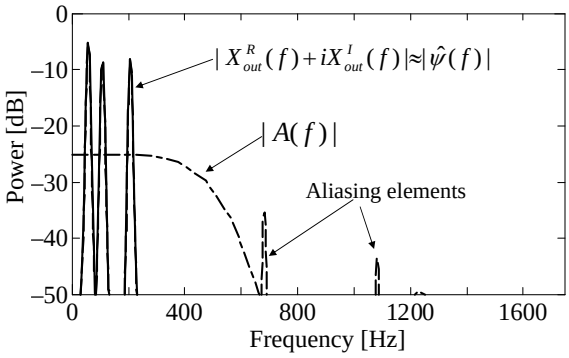


Fig. 4. Example of the frequency characteristics of the RMW $|\hat{\psi}(f)|$ and $|X_{out}^R(f) + iX_{out}^I(f)|$.

うに，$|\hat{\psi}(f)|$ と $|X_{out}^R(f) + iX_{out}^I(f)|$ はほぼ完全に重なっており，寄生フィルタが RMW に良く近似していることが分かる．また $|A(f)|$ の周波数領域には $|X^R(f) + iX^I(f)|$ の周波数成分が含まれているため，その中から寄生フィルタ $\{u_k^R\}$ と $\{u_k^I\}$ により異常信号の検出は可能であることがわかる．ここで，Fig. 4 において 650[Hz] と 1100[Hz] 付近に見られる成分は，寄生フィルタの近似誤差の影響で生じたエイリアシング成分であるが，その値は 50〜200[Hz] の主成分に比較して −20[dB] 以下であり，問題ないレベルである．

4. 寄生的離散ウェーブレット変換による高速ウェーブレット瞬時相関

ここで, 異常信号の検出を定量化するために, 高速ウェーブレット瞬時相関 (Fast wavelet instantaneous correlation, F-WIC) を次式のように定義する．

$$R(k) = \sqrt{(x_k^R)^2 + (x_k^I)^2} \tag{4.1}$$

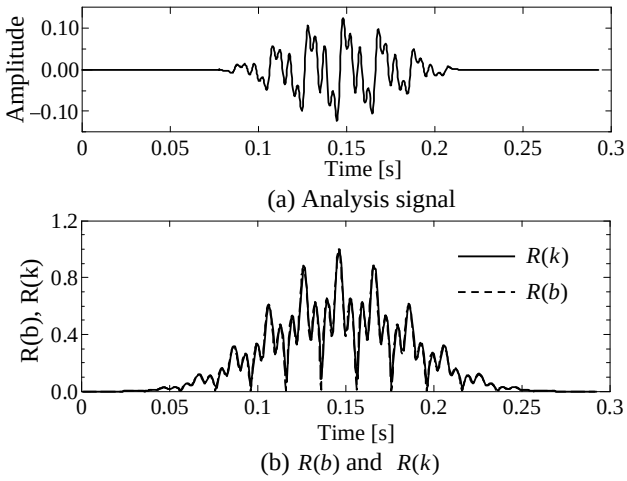


Fig. 5. Example of $R(k)$ obtained from the P-DWT and $R(b)$ obtained from the CWT

ただし，k は離散化時間であり，離散化間隔 Δt は寄生レベルによって異なり，$\Delta t = \Delta\tau/2^j$ となる．また $\Delta\tau$ はサンプリング間隔である．すなわち $R(k)$ の時間間隔 Δt は寄生レベル j が深くなるにつれ，CWT を用いるウェーブレット瞬時相関 $R(b)$ の時間間隔 $\Delta\tau$ より大きくなる．これに対応して $R(k)$ の計算量が減り高速化が実現される．また実部 x_k^R と虚部 x_k^I を用いることで，$R(k)$ は解析信号と寄生フィルタの間に生じる位相の影響を抑えることができる．

通常，寄生フィルタが付与している寄生レベルがある程度深くなると，ダウンサンプリングにより計算速度は速くなるが，一方で寄生フィルタの係数の数は少なくなり，寄生フィルタに対応する RMW の形状が崩れて検出精度が落ちることもある．ここで寄生レベルを定める際に，寄生フィルタでは表されない RMW の部分を，RMW のエネルギー損失 L_e として次式 (4.2) により定義する．

$$L_e = 10\log_{10}\left(\frac{\sum_j \sum_k (d_{j,k})^2}{\|\psi(t)\|^2}\right) \tag{4.2}$$

ただし式 (4.2) の $d_{j,k}$ は，3.1 節の寄生フィルタ設計手順 (1) を実行した場合のウェーブレット係数 $d_{j,k}$ であり，これらは寄生レベルの係数 $c_{j-1,k}$ には反映されない損失成分と考えることができる．

ここで，Fig. 3(a) に示した寄生フィルタ $\{u_k^I\}$ と $\{u_k^R\}$ を用い，Fig. 3(b) に示した RMW の虚数部を解析信号として上述の手順より $R(k)$ を求めた．この時の寄生フィルタによる RMW のエネルギー損失 L_e は −32[dB] である．また比較のために RMW の実数部と虚数部（Fig. 3(b) には虚数部のみを表示）を用いて同様の解析信号を解析して，CWT による $R(b)$ を求めた．それらの結果を Fig. 5 に示す．ただし Fig. 5(a) は解析信号で，(b) は求められた $R(k)$ と $R(b)$ である．Fig. 5(b) に示した $R(k)$ と $R(b)$ を比較してわかるように，$R(k)$ と $R(b)$ の振幅の大きさと振幅の時刻はよく一致している．ここで $R(k)$ と $R(b)$ の差を評価するために，$\sum |R(k) - R(b)|^2/N$ で表される平均二乗誤差 (MSE) を定義して求めた

Table 1. Number of multiplication

	RMW length				
level	64	128	256	512	1024
-1	78	142	270	526	1038
-2	66	98	162	290	546
-3	90	106	138	202	330
-4	162	170	186	218	282
CWT	138	260	516	1028	2052

ところ，−69.4251[dB] となった．すなわち本研究で提案した P-DWT による F-WIC は，寄生フィルタ $\{u_k^R\}$ と $\{u_k^I\}$ の設計が良好であれば，元の RMW を用いた CWT による WIC と，ほぼ同様の検出精度を持つと言える．

F-WIC による異常信号検出は，実時間で行うことを想定している．ここで $R(k)$ を 1 サンプル得るための乗算回数を計算量の目安として次式で定義する．

$$Q_j = 10\sum_{i=p}^{-1} 2^{i+1} + 2^{p+1}L + 4 \tag{4.3}$$

ただし p ($p < 0$：整数) は寄生レベル，L は RMW の長さである．また式 (4.3) の第 1 項目は DWT，2 項目は寄生フィルタ，3 項目は F-WIC による乗算回数である．また比較のために，CWT による WIC の計算量も次に示す．

$$Q_j = 2L + 4 \tag{4.4}$$

表 2 には P-DWT による F-WIC の計算量と，CWT による WIC の計算量を示す．表 2 が示すように，RMW の長さが長く，寄生レベルが深くなるほど，WIC に対して F-WIC の計算量がより小さくなる傾向にある．これは P-DWT がダウンサンプリングを有効に利用し，計算量を減少させる方向に働くためである．しかし寄生レベルが下げすぎると P-DWT の計算量が再び増加していく現象が見られる．これは，寄生レベルを下げたことによる寄生フィルタの計算量の減少よりも，それによって付加される DWT の計算量のほうが大きくなるためである．従って寄生レベルの決定には，計算量を最小にする工夫を施している．

5. 信号抽出を実現した寄生的離散ウェーブレット変換

これまでに Fig.1 に示すようなツリー構造を用いて寄生的フィルタによる信号検出の高速化を実現した [8]．しかし，P-DWT による検出は DWT による分解のみを利用した手法

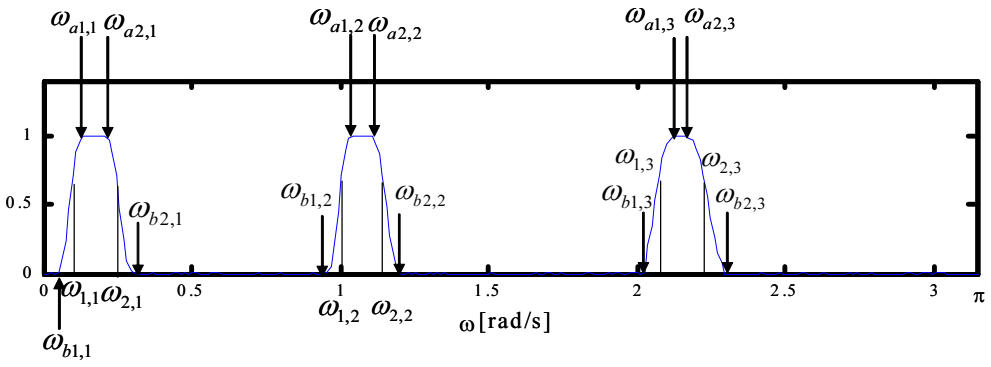


Fig. 6. Example of frequency characteristic of the BP-RMW

であり，DWT による再構成を利用していない．そのため，P-DWT による応用範囲が限られてしまうのが現状である．ここでは，3 章に示した P-DWT を拡張し，DWT の再構成を用いた P-DWT として，P-DWT による信号抽出手法を紹介する．

5.1 帯域通過型実信号マザーウェーブレット (BP-RMW)

3.1 節に示すように，異常信号により作成した RMW から寄生フィルタを構成して P-DWT を利用する場合，異常信号を抽出することはできない．これは，RMW の各周波数のパワーは異常信号のもつ周波数によって異なるため，RMW を持つ信号のみを再構成により抽出した際に，抽出した信号の各周波数の振幅は増幅や減衰が生じてしまう．そこで，RMW のもつ周波数成分をバンドパスフィルタのように置き換えた RMW を提案した．この RMW を帯域通過型実信号マザーウェーブレット (Band Pass-RMW,BP-RMW) と呼ぶ [8]．Fig.6 に設計した BP-RMW の周波数特性を示し，Fig.6 を参考に BP-RMW の設計手順に以下に示す. ただし，Fig.6 の横軸は角周波数 ω[rad/s] である．

(1) 異常信号成分から RMW を作成し，その周波数特性から通過させる周波数領域 (通過領域)$\omega_{a1,l}, \omega_{a2,l}(\omega_{a2,l} \leq \omega_{a1,l})$ を決定する．l は通過させる周波数領域の数を表しており，$l \leq 2$ に関しては $\omega_{a1,l+1} > \omega_{a2,l}$ が成立する．Fig.6 では $l = 1, 2, 3$ となり 3 つの通過領域をもつことになる．

(2) 過渡・減衰曲線部分の開始・終了周波数 $\omega_{b1,l}$，$\omega_{b2,l}$ を $\Delta_{1,l}$，$\Delta_{2,l}$ を用いて式 (5.1) より決定する．

(5.1) $$\omega_{b1,l} = \omega_{a1,l} - 2\Delta_{1,l}, \omega_{b2,l} = \omega_{a2,l} + 2\Delta_{2,l}$$

(3) 式 (5.2) より過渡減衰領域の中心となる周波数 ω_1，ω_2 を求める．

(5.2) $$\omega_{1,l} = \frac{\omega_{a1} + \omega_{b1}}{2}, \omega_{2,l} = \frac{\omega_{a2} + \omega_{b2}}{2}$$

(4) BP-RMW の周波数特性 $\hat{\psi}(\omega)$ を式 (5.3)，式 (5.4)，式 (5.5) を用いて求める．

(5.3)
$$\hat{\psi^{BP}}(\omega) = \begin{cases} 1 & (\omega_{a1} \leq \omega \leq \omega_{a2}) \\ \hat{\psi'^{BP}}(\omega) & (\omega_{b1,l} < \omega < \omega_{a1,l}, \omega'_{a2,l} < \omega < \omega_{b2,l}) \end{cases}$$

(5.4)
$$\hat{\psi'^{BP}}(\omega) = cos\left[\frac{\pi}{2}\left(\frac{1-\Delta_{1,l}}{2\Delta_{1,l}}\left\{\frac{|\omega'|}{(1-\Delta_{1,l})\pi} - 1\right\}\right)\right]$$
$$\omega' = \frac{\omega}{\omega_{1,l}}\pi \quad (\omega_{b1,l} < \omega < \omega_{a1,l})$$

(5.5)
$$\hat{\psi'^{BP}}(\omega) = cos\left[\frac{\pi}{2}\left(\frac{1-\Delta_{2,l}}{2\Delta_{2,l}}\left\{\frac{|\omega'|}{(1-\Delta_{2,l})\pi} - 1\right\}\right)\right]$$
$$\omega' = \frac{\omega}{\omega_{1,l}}\pi \quad (\omega_{b1,l} < \omega < \omega_{a1,l})$$

(5) 逆フーリエ変換を行い．設計した BP-RMW が得られる．

設計手順 (1) に関しては以下の 2 つの方法を挙げる．

(A) 作成した RMW の周波数特性において，通過させる周波数帯域の閾値を用い，閾値以上の周波数帯域の開始周波数と終了周波数を $\omega_{a1,l}$， $\omega_{a2,l}$ として決定する．

(B) あらかじめ，抽出したい周波数帯域と過渡・減衰曲線部分の開始・終了周波数を自ら設定し，それに応じて $\omega_{a1,l}$， $\omega_{a2,l}$ を決定する．

ただし，(B) の手順のように RMW から設計を行わなくても BP-RMW は作成可能である．また，設計手順 (1) における通過領域に関しては $l \leq 2$ の時，$\omega_{a1,l+1} > \omega_{a2,l}$ が成立しない場合，通過領域が重なる状態であるので，これらの領域を一つの通過領域として置き換える方が好ましい．設計手順 (2) においては，$l \leq 2$ の時，$\omega_{a2,l}$ と $\omega_{a1,l+1}$ との間隔が式 (5.1) の $\Delta_{2,l}$，$\Delta_{2,l+1}$ で埋まるのであれば, 次式を成立させる必要がある．

(5.6)
$$|\hat{\psi^{BP,l}}(\omega)|^2 + |\hat{\psi^{BP,l+1}}(\omega)|^2 = 1\ ,\ (\omega_{a2,l} < \omega < \omega_{a1,l+1})$$

ただし，式 (5.6) を成立させるには $\Delta_{2,l} = \Delta_{1,l+1}$ とする．また，$\omega_{a1,l+1} > \omega_{a2,l}$ が成立していたとしても，式 (5.1) により $\omega_{a1,l} < \omega_{b1,l+1} < \omega_{a2,l}$ または $\omega_{a1,l+1} < \omega_{b2,l} < \omega_{a2,l+1}$ となる．つまり，ある通過領域が他の通過領域のもつ過渡・減衰領域に干渉され，かつ Δ の調整でも対応できないのであれば，一つの通過領域とした方が良い．

ここで，式 (5.4)(5.5) は式 (5.7) を参考にしたものである．

(5.7)
$$\hat{\phi}(\omega) = cos\left[\frac{\pi}{2}\left(\frac{1-\Delta}{2\Delta}\left\{\frac{|\omega'|}{(1-\Delta)\pi} - 1\right\}\right)\right]$$
$$(1-\Delta)\pi < |\omega| < (1-\Delta)\ ,\ 0 < \Delta \leq 1/3$$

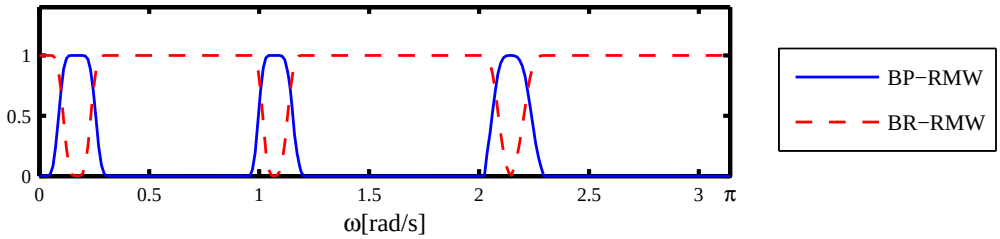


Fig. 7. Example of frequency characteristic of the BP-RMW

式 (5.7) は Daubechies が提案した Meyer ウェーブレットのスケーリング関数 [3] における曲線部分の範囲を式中の Δ によって可変するために著者らが提案したものである [9]. この式は角周波数 $\omega = \pi$ を曲線部分の中心周波数としているため，式 (5.4)(5.5) では，式 (5.7) 中の曲線部を構成する ω の区間を任意の区間となる ω' に置き換えている．Fig.7 に，作成した BP-RMW の例を示す．作成された BP-RMW は RMW の周波数領域を持ったバンドパスフィルタの集まりとなっている．

5.2 帯域除去型実信号マザーウェーブレット (BR-RMW)

ここで，BP-RMW に対し，BP-RMW の周波数帯域を除去する RMW を構成する．これを帯域除去型 RMW((Band Rejection-RMW,BR-RMW) と呼ぶ [8]．BR-RMW の周波数特性 $\hat{\psi}^{BR}(\omega)$ は Meyer ウェーブレットの曲線部分の関係式を利用した式 (5.8) により決定される．Fig.7 の例から作成された $\hat{\psi}^{BP}(\omega)$ は Fig.7 の破線部分に示される周波数特性を持つ．Fig.7 より $\hat{\psi}^{BP}(\omega)$ 以外の周波数領域で構成されていることが分かる．

$$\hat{\psi}^{BR}(\omega) = \sqrt{1 - |(\hat{\psi}^{BP})(\omega)|^2} \tag{5.8}$$

5.3 寄生フィルタの作成

作成した BP-RMW と BR-RMW から寄生フィルタ $u_k^{R,BP}$，$u_k^{I,BP}$ および $u_k^{R,BR}$，$u_k^{I,BR}$ を作成するが，作成手順は 3.1 節で述べたものとは異なる．その理由として以下の 2 点が挙げられる．

(1) P-DWT の抽出に用いるフィルタは DWT の分解数列，再構成数列に相当するため，DWT の分解アルゴリズムによって得た係数は，ダウンサンプリングを行っているため，そのまま寄生フィルタとして用いることができない．

(2) 高周波成分側の分解数列 b_n^R, b_n^I に該当する寄生フィルタを作成する上で，低周波成分側の分解数列 a_n^R，a_n^I によって分解された影響を受けてしまう. これにより，b_n^R, b_n^I のような通過領域が構成されず，信号の抽出精度が悪化する．

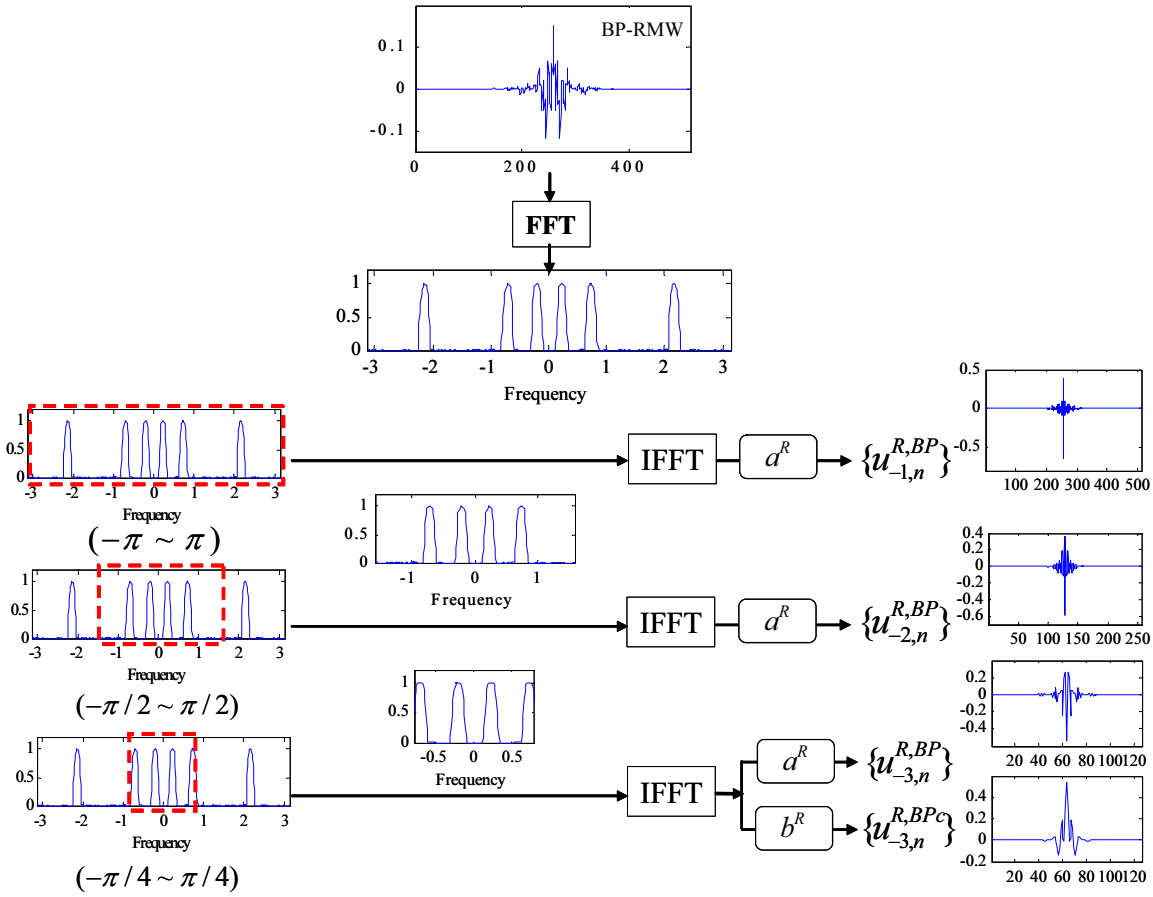


Fig. 8. Design method of Parasitic filters

これらの問題を改善するために，以下の手順を用いて寄生フィルタを作成する．分解レベル j の最小値を J(J は負の整数) とする．また，Fig.8 に作成手順の模式図を示す．ただし図中の Frequency と書かれた横軸は角周波数 ω[rad/s] である．

(1) BP-RMW $\psi^{BP}(t)$ のフーリエ変換結果 $\hat{\psi}^{BP}(\omega)$ を得る $(-\pi < \omega < \pi)$.

(2) フーリエ変換した BP-RMW を $(-\pi/2^{-j-1} < \omega < \pi/2^{-j-1})$ で切り出し，逆フーリエ変換を行い，これを $\psi_j^{BP}(t)$ とする．

(3) $\psi_j^{BP}(t)$ を b_n^R, b_n^I で分解し，これを寄生フィルタ $u_{j,k}^{R,BP}$，$u_{j,k}^{I,BP}$ とする．ただし j=J の場合には $\psi_j^{BP}(t)$ を a_n^R, a_n^I を行い，低周波成分側の分解する寄生フィルタ $u_{j,k}^{R,BPc}$, $u_{j,k}^{I,BPc}$ を求める．

(4) (1)–(3) を分解レベル j が J になるまで行う (j=-1，-2，…，J+1，J).

(5) BR-RMW から作成する寄生フィルタに関しても同様に (1)～(4) の手順を行い，$u_{j,k}^{R,BR}$，$u_{j,k}^{I,BR}$，$u_{j,k}^{R,BRc}$，$u_{j,k}^{I,BRc}$ を得る．

ここで，$u_{j,k}^{R,BP}$，$u_{j,k}^{R,BR}$ の j は分解レベル，R は寄生フィルタの実数部を表し，$u_{j,k}^{I,BP}$，$u_{j,k}^{I,BR}$ の I は寄生フィルタの虚数部を表している．以上の手順により，レベル J までの P-DWT

による分解に用いる以下の寄生フィルタを得ることができる.

・BP-RMW から作成された寄生フィルタ.

$$u_{-1,k}^{R,BP}, u_{-2,k}^{R,BP}, .., u_{J,k}^{R,BP}, u_{J,k}^{R,BPc}$$
$$u_{-1,k}^{I,BP}, u_{-2,k}^{I,BP}, .., u_{J,k}^{I,BP}, u_{J,k}^{I,BPc}$$

・BP-RMW から作成された寄生フィルタ.

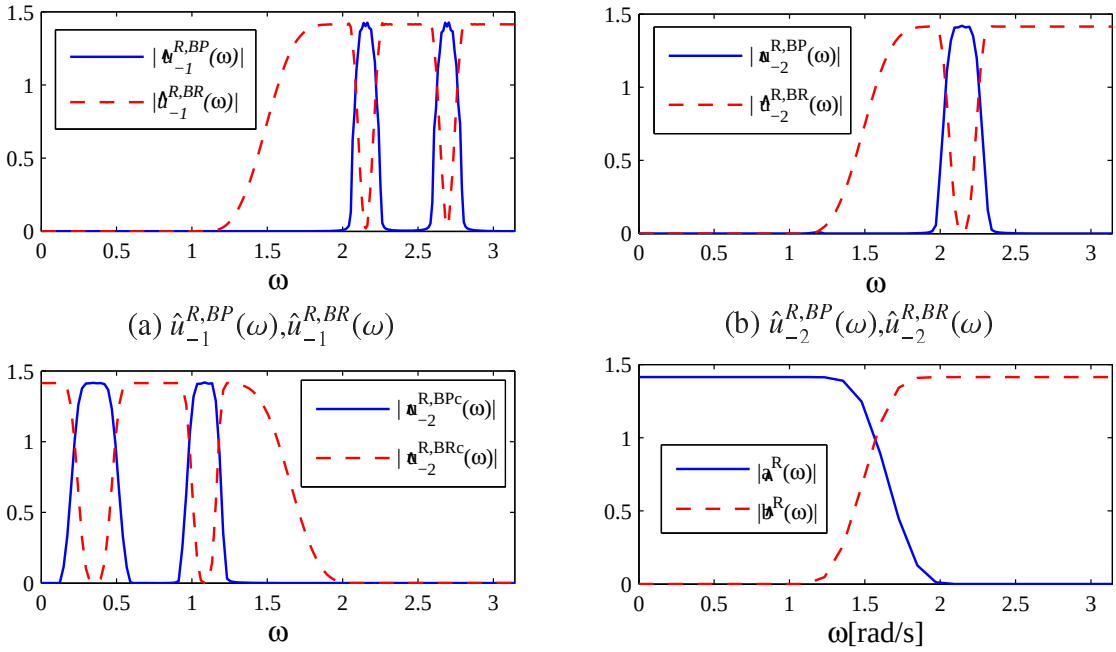
$$u_{-1,k}^{R,BR}, u_{-2,k}^{R,BR}, .., u_{J,k}^{R,BR}, u_{J,k}^{R,BRc}$$
$$u_{-1,k}^{I,BR}, u_{-2,k}^{I,BR}, .., u_{J,k}^{I,BR}, u_{J,k}^{I,BRc}$$

作成した寄生フィルタにおいてその両端が 0 に減衰している領域があると見なされれば，0 と見なした領域を使用せず寄生フィルタのデータ点数を少なくしても良い．寄生フィルタのデータ点数を少なくした場合，P-DWT における計算量を減少させることができる．これらの寄生フィルタ $u_{j,k}^{R,BP}$，$u_{j,k}^{R,BR}$ の周波数特性の例と分解数列 a_n^R，b_n^R の周波数特性を Fig.9 に示す．Fig. 9 の (a) はレベル-1 の高周波成分，(b) は-2 の高周波成分，(c) レベル-2 の低周波成分の分解を行う寄生フィルタにおける周波数特性を表している. 図中の横軸は各周波数 ω[rad/s] であり，その範囲が π〜π であるが，これは DWT の分解アルゴリズムによってダウンサンプリングされた高周波・低周波成分における各周波数を表しているためである．しかし，各レベルの寄生フィルタが分解する実際の正の周波数領域は，サンプリング周波数を f_s とすると，Fig.9 の (a) で $f_s/4$〜$f_s/2$，(b) で $f_s/8$〜$f_s/4$，(c) で 0〜$f_s/8$ となる. また，各レベルの寄生フィルタ $u_{j,k}^{R,BP}$，$u_{j,k}^{R,BR}$ の周波数特性を重ねると，Fig.9 の (d) に示す分解数列 b_n^R と a_n^R の周波数特性と一致する. これは，寄生フィルタ $u_{j,k}^{R,BP}$，$u_{j,k}^{R,BP}$ のそれぞれの分解によって，各寄生フィルタのもつ周波数成分を分解し，かつ寄生フィルタ全体で従来の分解アルゴリズムが対象とする周波数成分を全て分解していることを表している.

5.4 P-DWT による信号抽出の分解・再構成アルゴリズム

P-DWT の再構成は RMW の周波数成分を持つ係数 x_k または寄生対象の低周波成分・高周波成分の再構成が可能であるが，低周波・高周波成分から x_k を除いた成分を再構成することはできない．これは x_k を再構成する場合には寄生対象の低周波成分・高周波成分は 0 として再構成するためである．もし，x_k が除去対象の信号であれば，x_k 以外の周波数成分を再構成するうえで，x_k の持つ周波数周辺の成分を再構成することが困難となる．提案する P-DWT 信号抽出は Fig.10，Fig.11 に示すツリー構造にて行う．これらのツリー構造は従来の DWT の分解・再構成アルゴリズムをベースに作成されているが，PTI-CDWT の分解・再構成アルゴリズムをベースに作成されても構わない.

Fig.10 は信号抽出における P-DWT の分解アルゴリズムを表しており，高周波成分側の



(a) $\hat{u}_{-1}^{R,BP}(\omega),\hat{u}_{-1}^{R,BR}(\omega)$ (b) $\hat{u}_{-2}^{R,BP}(\omega),\hat{u}_{-2}^{R,BR}(\omega)$

(c) $\hat{u}_{-2}^{R,BPc}(\omega),\hat{u}_{-2}^{R,BRc}(\omega)$ (d) $\hat{a}^{R}(\omega),\hat{b}^{R}(\omega)$

Fig. 9. Example of Frequency characteristic of $u_{j,k}^{R,BP}$,$u_{j,k}^{R,BR}$,a^R,and b^R

分解数列 b_n^R，b_n^I の代わりに，寄生フィルタ $u_k^{R,BP}$,$u_k^{R,BR}$，$u_k^{I,BP}$,$u_k^{I,BR}$ が寄生するかたちで追加される. 分解レベル J においては高周波成分側だけではなく，低周波成分側の分解数列 a_n^R，a_n^I の代わりに $u_k^{R,BPc}$,$u_k^{R,BRc}$，$u_k^{I,BPc}$,$u_k^{I,BRc}$ 寄生するかたちで追加される. これらの寄生は従来の構造を残していた寄生にくらべて，完全に寄生した状態となっている. ただし，原信号 $f(t)$ から c_0^R,c_0^I への処理は文献 [10] に示した補間を用いる.

Fig.10 示される分解アルゴリズム中の高周波成分を求める計算は次式となる.

(5.9) $$d_{j,n}^{R,BP} = \sum_k u_{j,2n-k}^{R,BP} c_{j+1,k}^R \quad , \quad d_{j,n}^{I,BP} = \sum_k u_{j,2n-k}^{I,BP} c_{j+1,k}^R$$

(5.10) $$d_{j,n}^{R,BR} = \sum_k u_{j,2n-k}^{R,BR} c_{j+1,k}^R \quad , \quad d_{j,n}^{I,BR} = \sum_k u_{j,2n-k}^{I,BR} c_{j+1,k}^R$$

低周波成分 $c_{j,n}^R$,$c_{j,n}^I$ は完全シフト不変複素数離散ウェーブレット変換 (PTI-CDWT) により求める [10]. ただし，分解レベル j=J においては，低周波成分の分解も寄生フィルタを用いて行い，次式にて求める.

(5.11) $$c_{J,n}^{R,BP} = \sum_k u_{J,2n-k}^{R,BPc} c_{J+1,k}^R \ , \ c_{J,n}^{I,BP} = \sum_k u_{J,2n-k}^{I,BPc} c_{J+1,k}^R$$

(5.12) $$c_{J,n}^{R,BR} = \sum_k u_{J,2n-k}^{R,BRc} c_{J+1,k}^R \ , \ c_{J,n}^{I,BR} = \sum_k u_{J,2n-k}^{I,BRc} c_{J+1,k}^R$$

Fig.11 は信号抽出における P-DWT の再構成アルゴリズムを表しており，Fig.10 に示される分解アルゴリズムの同様に寄生フィルタが追加されている. Fig.11 示される再構成ア

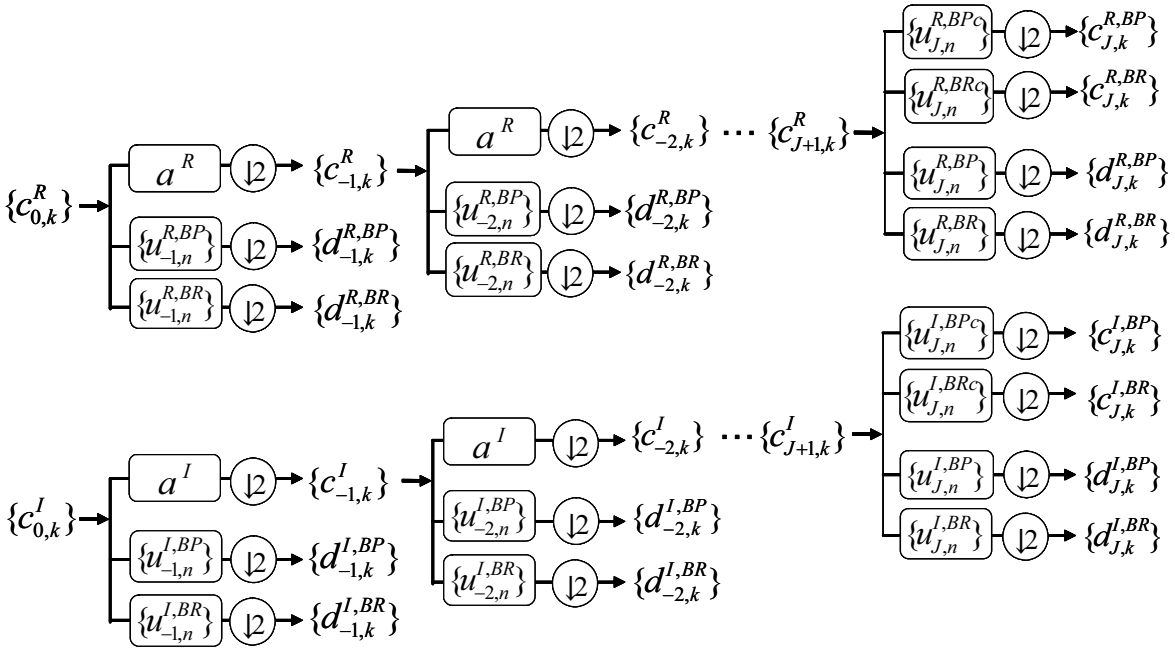


Fig. 10. Decomposition tree by P-DWT to signal extraction

ルゴリズムの計算方法について述べる. 式 (5.9)～(5.12) により得られた分解レベル j=J の高周波成分および低周波成分からレベル J+1 の低周波成分 $c^R_{J+1,n}, c^I_{J+1,n}$ を次式にて求める.

$$c^R_{J+1,n} = (\sum_k u'^{R,BPc}_{J,n-2k} c^{R,BP}_{J,k} + \sum_k u'^{R,BRc}_{J,n-2k} c^{R,BR}_{J,k}) + (\sum_k u'^{R,BP}_{J,n-2k} d^{R,BP}_{J,k} + \sum_k u'^{R,BR}_{n-2k} d^{R,BR}_{J,k}) \tag{5.13}$$

$$c^I_{J+1,n} = (\sum_k u'^{I,BPc}_{J,n-2k} c^{I,BP}_{J,k} + \sum_k u'^{I,BRc}_{J,n-2k} c^{I,BR}_{J,k}) + (\sum_k u'^{I,BP}_{J,n-2k} d^{I,BP}_{J,k} + \sum_k u'^{I,BR}_{n-2k} d^{I,BR}_{J,k}) \tag{5.14}$$

分解レベル $j(j = -1, -2, \ldots, J+1)$ の低周波成分 $c^R_{j,n}$, $c^I_{j,n}$, および式 (5.9) と (5.10) により得られた高周波成分からレベル +1 の低周波成分 $c^{R,BP}_{j+1,n}$, $c^{I,BP}_{j+1,n}$ を求める.

$$c_{j+1,n} = \sum_k g^R_{n-2k} c^R_{j,k} + (\sum_k u'^{R,BP}_{j,n-2k} d^{R,BP}_{j,k} + \sum_k u'^{R,BR}_{n-2k} d^{R,BR}_{j,k}) \tag{5.15}$$

$$c_{j+1,n} = \sum_k g^I_{n-2k} c^I_{j,k} + (\sum_k u'^{I,BP}_{j,n-2k} d^{I,BP}_{j,k} + \sum_k u'^{I,BR}_{n-2k} d^{I,BR}_{j,k}) \tag{5.16}$$

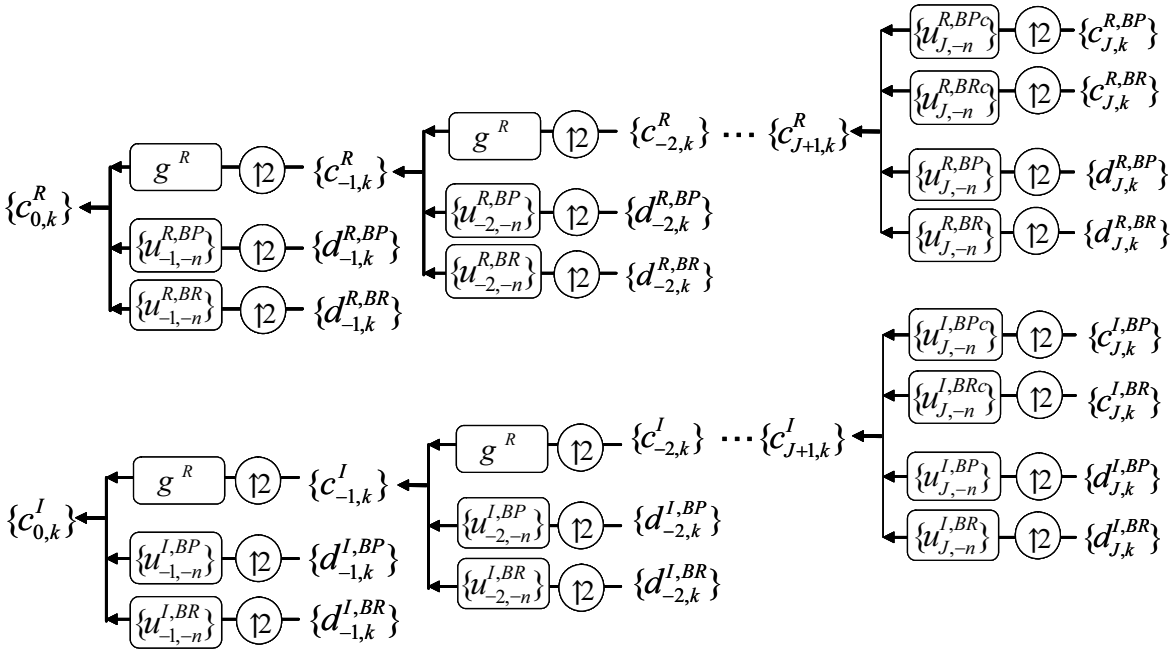


Fig. 11. Reconstitution tree by P-DWT to signal extraction

$$
\begin{aligned}
u'^{R,BP}_{j,k} &= u^{R,BP}_{j,-k} \quad , \quad u'^{R,BPc}_{J,k} = u^{R,BPc}_{J,-k} \\
u'^{R,BR}_{j,k} &= u^{R,BR}_{j,-k} \quad , \quad u'^{R,BRc}_{J,k} = u^{R,BRc}_{J,-k} \\
u'^{I,BP}_{j,k} &= u^{I,BP}_{j,-k} \quad , \quad u'^{I,BPc}_{J,k} = u^{I,BPc}_{J,-k} \\
u'^{I,BR}_{j,k} &= u^{I,BR}_{j,-k} \quad , \quad u'^{I,BRc}_{J,k} = u^{I,BRc}_{J,-k}
\end{aligned}
\tag{5.17}
$$

もし，ある分解レベルにおいては寄生レベルを用いず従来法の処理を行いたい場合には，該当する分解レベルの高周波成分における分解・再構成アルゴリズムを PTI-CDWT の分解・再構成アルゴリズム [10] で行えばよい.

5.5 P-DWT による信号抽出手順

P-DWT による信号抽出は，原信号を P-DWT によって分解し，抽出したい成分以外の値を全て 0 にし，再構成を行うことで実行される. 例えば，P-DWT の分解によって得た高周波成分のうち，d^{BP}_{k}，$c^{BP}_{J,k}$ 以外の成分を 0 にし，再構成アルゴリズムを実行すると，BP-RMW の持つ周波数成分のみを再構成することができる. 一方，d^{BP}_{k}，$c^{BP}_{J,k}$ を全て 0 にし，再構成アルゴリズムを実行した場合には BP-RMW の持つ周波数成分以外，つまり BR-RMW の持つ周波数成分のみを抽出することができる. また，$d^{BP}_{-2,k}$ と $c^{BR}_{J,k}$ を再構成すれば，BP-RMW の持つ周波数成分の一部と BR-RMW の波形を組み合わせた波形を抽出することが可能となる.　つまり，P-DWT による信号抽出は DWT よる周波数のオクター

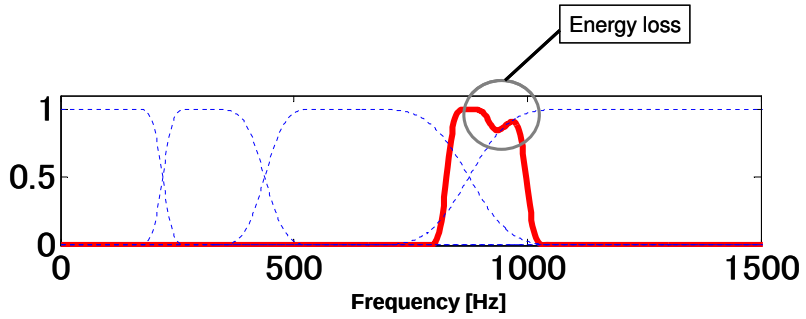


Fig. 12. Signal extraction that uses P-DWT

ブ分割と寄生フィルタによる周波数分割を組み合わせることで，異常信号の周波数帯域での成分を操作することができるため，柔軟性のある信号抽出が可能となる．

6. P-DWT を用いた信号抽出におけるエネルギー損失問題

前章で寄生的離散ウェーブレット変換 (P-DWT) を用いた信号抽出の理論について述べた．P-DWT による信号抽出は DWT による周波数のオクターブ分割と，寄生フィルタによる周波数分割を組み合わせることで，所望の信号（以下，目的信号とする）の周波数帯域での成分を操作することが可能なため，柔軟性のある信号抽出が可能になった．一方，P-DWT を用いた特定信号の再構成では，従来適用してきた異常信号の有無のみの確認と異なり，信号の再構成や，BP-RMW や BR-RMW の適用が必要となる．しかし，これらを用いた結果，新たに，ある周波数領域でエネルギー損失が起こるという問題が生じた．本章ではエネルギー損失について述べ，その原因と解決方法について述べる．

エネルギー損失はある周波数領域が持つパワー，つまり振幅が減少する現象である．この問題は P-DWT の分解によって生じた周波数バンド間である過渡領域で発生する．これにより過渡領域と抽出する周波数領域が重なると，抽出によって得た信号は一部の周波数領域が欠けた信号となる.Fig. 12 にエネルギー損失が生じた解析例を示す．ただし，破線は分解レベルごとの周波数バンドを表す．これは，インパルス信号に 900[Hz] 付近の周波数を抽出するための RMW を適用した時に生じたものである．理想的には，抽出したい周波数領域が Fig. 13 のように領域が欠けることなく抽出されることが望ましい．しかし，エネルギー損失が生じた結果，Fig. 12 のように領域が欠けてしまい，所望の信号周波数領域を正確に抽出できないことがわかる．このように一部の周波数領域のエネルギーが損失してしまうと，抽出精度を大きく低下させる原因となる．

6.1 エネルギー損失問題の原因

従来研究である異常信号検出では信号の再構成を行わなかったため，エネルギー損失問題は発生しなかった．また，P-DWT で用いられる複素数離散ウェーブレット変換

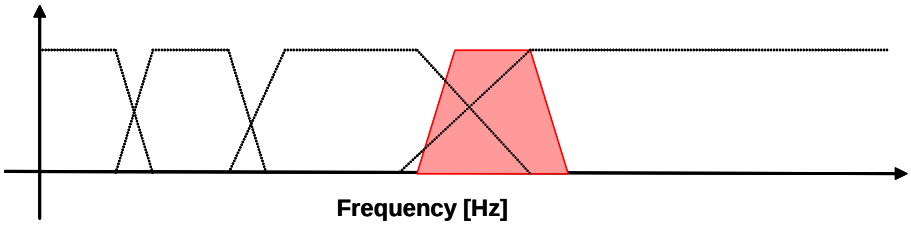


Fig. 13. Ideal signal extraction

(CDWT) のみの処理ではこの問題は発生していない．よって信号抽出にて新たに必要になった帯域通過型実信号マザーウェーブレット (BP-RMW) や帯域除去型実信号マザーウェーブレット (BR-RMW) を適用したことや，信号を再構成処理がエネルギー損失を発生させる一因であることが考えられる．

エネルギー損失がどの処理で発生しているかを模式的に Fig. 14 に示す．ただし，図の上部のフローチャートは P-DWT を用いた抽出手法を簡略的にまとめたものであり，下部は Fig. 12 のエネルギー損失の発生を処理ごとに縦軸パワー，横軸周波数 [Hz] で表したものである．また，分解レベルは-3 とし，サンプリング周波数は 3500[Hz] である．抽出対象の周波数領域である 900[Hz] 付近の領域は，P-DWT の分解によって生じる Level-1 と Level-2 の周波数バンド間の過渡領域にあたり，このためエネルギー損失が発生している．このエネルギー損失を P-DWT のフローチャートで見ると，Level-1 ではすべての処理において周波数領域は欠けず，エネルギー損失は見られないが，Level-2 の周波数バンドの領域において，分解・再構成時に周波数領域が小さくなっていることが確認できる．Level-1 の欠けていない周波数領域と，Level-2 の小さくなった周波数領域を足し合わせると　の一部が欠けた周波数領域となる．つまり，エネルギー損失は分解によって生じた Level-2 周波数バンド，つまり分解によって分けられた低周波側で発生している．

6.2 ダウンサンプリングによるナイキスト周波数のシフト

Fig. 14 にて P-DWT での分解時，すでに Level-2 のウェーブレットが持つ周波数領域は小さくなっていることが確認できる．小さくなる原因は，Level-1 と Level-2 の周波数バンド間が交わる点以降の周波数領域が欠如したことにある．この欠如には，ダウンサンプリングすることでナイキスト周波数がシフトすることが大きく関係している．標本化定理より，ナイキスト周波数 f_N とは信号の情報を失わない限界の周波数を言い，サンプリング周波数 f_s の $\frac{1}{2}$ の周波数である．一方，多くの DWT と P-DWT では分解時，データを二つに一つ抜き出すダウンサンプリングを行っている．これにより f_s は $\frac{1}{2}$ され，同時にナイキスト周波数も $\frac{1}{2}$ される．P-DWT での分解時のナイキスト周波数の移動を Fig. 15 に示す．ただし，$f(t)$ は抽出対象の信号，$\{a_k\}$ は分解数列，$RMW_j (j = 1, 2, 3)$ は BP-RMW などの RMW，$d_j (j = 1, 2, 3), c_3$ は分解によって得られた各レベルのウェーブレット係数とスケーリング係数である．図中の上部の流れ図と下部の周波数領域で表したナイキスト

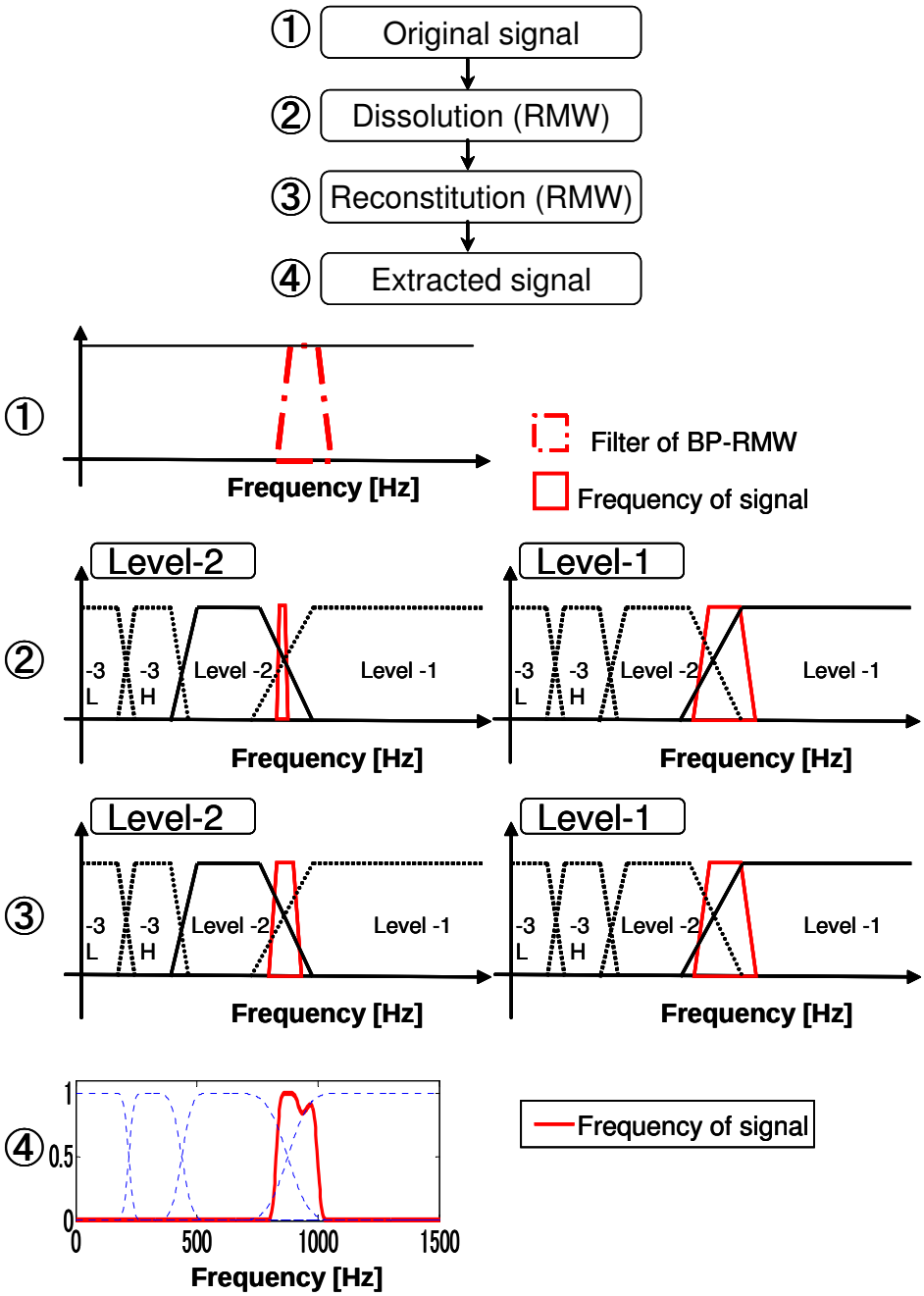


Fig. 14. Process of occurring energy loss

周波数の移動する様子は点線円の数字で対応している．

6.3 アップサンプリングによるミラー周波数の発生

上記に示したように，ナイキスト周波数以上の成分は失われる．Fig. 14 の Level-2 の の周波数域が小さくなったのはこれが原因である．一方，ダウンサンプリングを用い

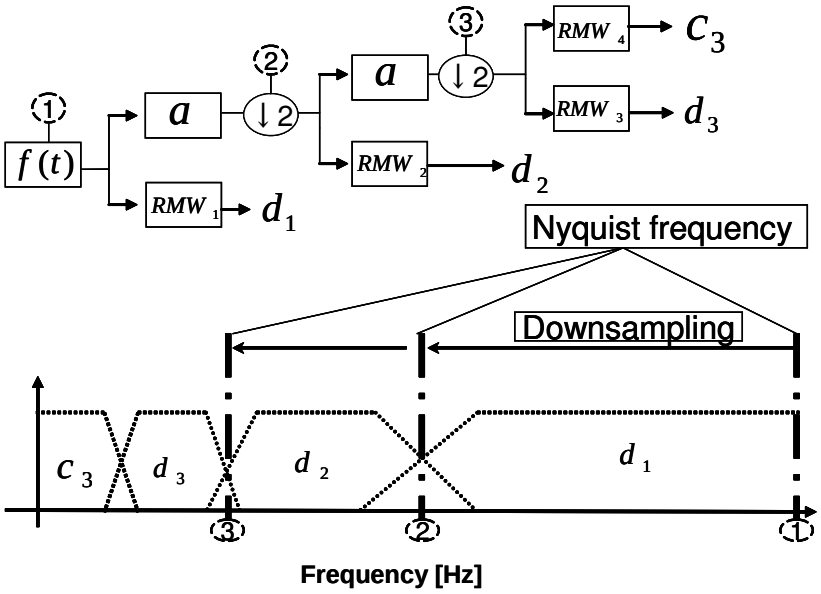


Fig. 15. Moving Nyquist frequency

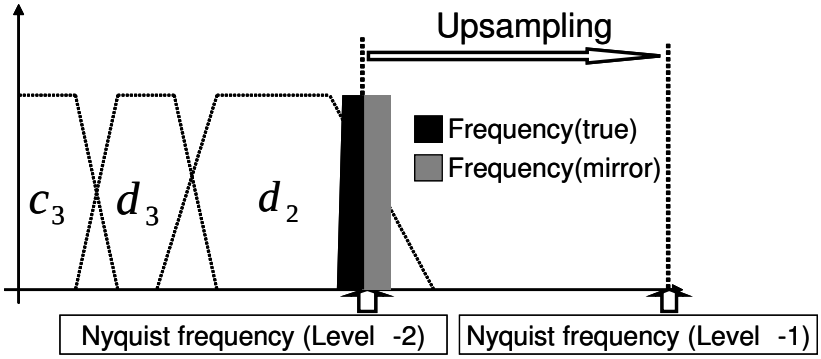


Fig. 16. Occurring mirror frequency

た DWT や P-DWT は再構成を行うためにアップサンプリングを行う．しかし，アップサンプリングを行っても一度失った周波数成分は復元されず，Fig. 16 のような同じ周波数データを持った周波数領域が，ミラーのようにナイキスト周波数を線対称として現れる．これにより，Fig. 14 の Level-2 の　の周波数域が生まれる．これはアップサンプリングが以下の式で表されるためである．

(6.1) $$F_U(\omega) = F(U\omega)$$

ただし，$F(\omega)$ は信号 $f(t)$ をフーリエ変換したものであり，$F_U(\omega)$ は $F(\omega)$ をアップサンプリングしたもの，U はアップサンプリングによって各データ間に入れる 0 の個数である．

今回 $U = 2$ であるため，アップサンプリングによってその周期性は 2 倍になる．この 2 倍されたことによって発生した周波数領域が元の周波数のデータを持ち，ミラーのように配置されることによって Fig. 16 のようになる．これと再構成数列や再構成のための RMW を畳み込み，再構成されることによってエネルギー損失のような形状になる．

アップサンプリングによって発生する周波数領域はミラーであることを示した．これはダウンサンプリングによって失われた周波数域は復元できないことを示すものであり，本研究で行う P-DWT の信号抽出ではこのような領域が存在するとエネルギー損失のような周波数領域が欠けた抽出結果となってしまう．ダウンサンプリングを行うことによって周波数領域のデータが失われるため，「ダウンサンプリングによるナイキスト周波数のシフト」が原因である．

7. エネルギー損失問題の解決法の一提案

前節でエネルギー損失問題の原因を述べた．ダウンサンプリングを行うことによってナイキスト周波数がシフトし，周波数成分を失う領域が発生することが原因であった．これを解決するためにダウンサンプリングを行わないウェーブレット変換である定常ウェーブレット変換 (Stationary Wavelet Transform：SWT) を P-DWT に適用することを提案した [11].

7.1 定常ウェーブレット変換 (SWT)

定常ウェーブレット変換 (SWT) [12] の特徴はダウンサンプリングを行わないことにある．これにより冗長性のあるウェーブレット変換が可能になる．従来の DWT はダウンサンプリングを行うことで信号のデータ数を変化させ，分解・再構成数列と呼ばれる同じフィルタを各分解レベルで用いて分解再・構成を行った．しかし，SWT はフィルタを次式により分解・再構成数列をアップサンプリングすることによりフィルタ特性の周期性を変化させ，信号のデータ数を変化させることなくフィルタの形状を変化させることにより分解・再構成を可能とする．しかし，ダウンサンプリングを行わないことによって計算量が増加し，処理速度が低下する問題点を抱える．

$$
\begin{aligned}
&a_j \to \uparrow 2 \to a_{j-1}, b_j \to \uparrow 2 \to b_{j-1} \\
&g_j \to \uparrow 2 \to g_{j-1}, h_j \to \uparrow 2 \to h_{j-1}
\end{aligned} \tag{7.1}
$$

ただし，$\{a_j\}, \{b_j\}$ は分解数列，$\{g_j\}, \{h_j\}$ は再構成数列，$j(j = -1, -2, ..., J+1, J)$ は分解レベル，$\uparrow 2$ はアップサンプリングを示す．そうして作られた分解数列などは Fig. 17 のように配置される．図中の上部は分解数列のフィルタ特性，下部は分解の流れ図を示す．ただし，$\{a_j\}, \{b_j\}, (j = -1, -2, -3)$ は式 (7.1) によってアップサンプリングされた分解数列であり，分解レベルは-3 である. また，再構成数列も上部と同様に配置され，下部の流れ図では，逆の処理により再構成を行う．Fig. 15 の上部と比較するとダウンサンプリングが行われていないことがわかる．

P-DWT における DWT と SWT でのナイキスト周波数の移動によって生じるエネルギー損失の違いについて Fig. 18 に示す．ただし，分解レベルは-3 とし，信号はインパルス信号とする．従来のように P-DWT に DWT を用いた場合，分解時周波数領域が失わ

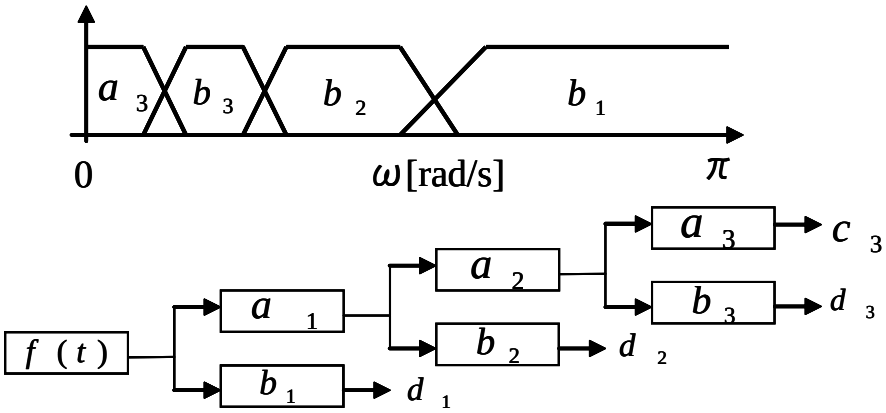


Fig. 17. Process of SWT

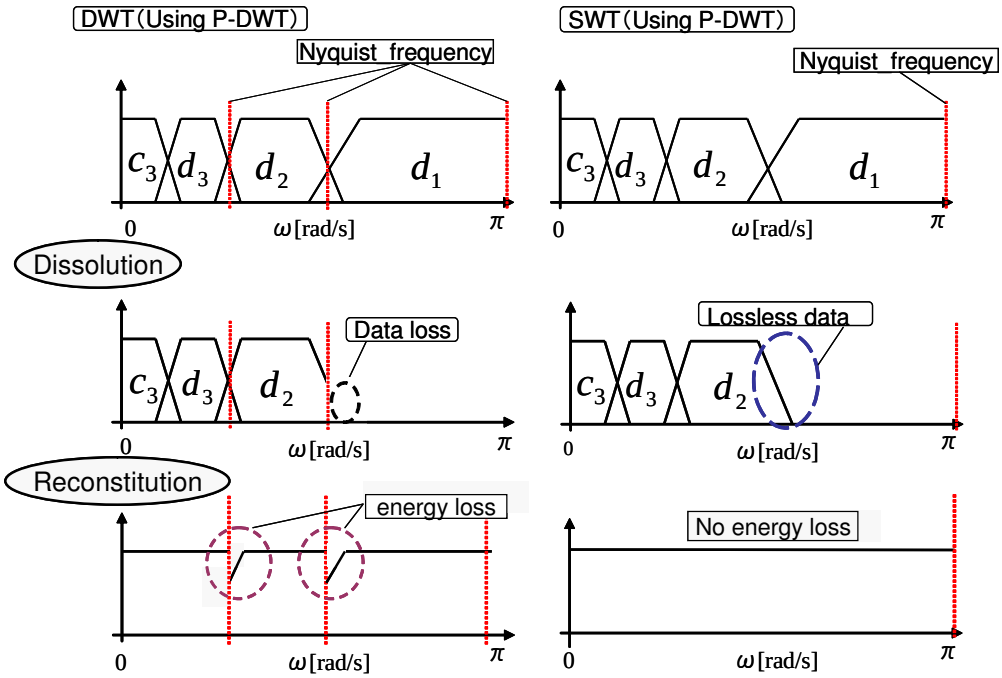


Fig. 18. Seting DWT against SWT

れ，再構成時にエネルギー損失が現れる．一方，SWT を P-DWT に適用するとナイキスト周波数のシフトが発生しないので，分解時にエネルギー損失が発生せず，再構成時にはエネルギー損失なく再構成が可能である．

7.2 冗長性のあるフィルタの設計

前小節にてエネルギー損失問題に対する SWT の有効性を述べた．しかし，実際に P-DWT に適用すると，SWT の特徴である分解・再構成数列のアップサンプリングによって各 RMW を作成する際にサンプリングのずれが発生した．これにより各レベルのウェーブレット係数，スケーリング係数の信号がずれ，完全に再構成できない実装上の問題が

新たに生じた．よって，アップサンプリングを行わずに分解・再構成数列の役割を果たすフィルタが必要になった．Fig. 17 に示した SWT の分解・再構成数列の周波数領域でのフィルタ特性の周期性を参考に，以下の式でフィルタを作成する [13].

$$\hat{\psi^M}(\omega) = \left(\sqrt{\left(\hat{\phi^M}(\frac{1}{2}\omega)\right)^2 - \left(\hat{\phi^M}(\omega)\right)^2}\right) e^{-i\frac{\omega}{2}} \tag{7.2}$$

$$\hat{\psi^R_{j,0}}(\omega) = \sqrt{2^{-j}}\hat{\psi^M}(2^{-j}\omega)e^{ij\omega} \tag{7.3}$$

$$H^R_j(\omega) = \frac{\hat{\psi^R_{j,0}}(\omega)}{\hat{\phi^M}(\omega)} \tag{7.4}$$

$$\hat{\phi^R_{j,0}}(\omega) = \sqrt{2^{-j}}\hat{\phi^M}(2^{-j}\omega) \tag{7.5}$$

$$G^R_j(\omega) = \frac{\hat{\phi^R_{j,0}}(\omega)}{\hat{\phi^M}(\omega)} \tag{7.6}$$

$$\hat{\phi^I_{0,0}}(\omega) = \hat{\phi^M}(\omega)e^{-i\frac{\omega}{2}} \tag{7.7}$$

$$\hat{\psi^I}(\omega) = \begin{cases} -\hat{\psi^M}(\omega)e^{-i\frac{\omega}{2}} & (\omega \geq 0) \\ \hat{\psi^M}(\omega)e^{-i\frac{\omega}{2}} & (\omega < 0) \end{cases} \tag{7.8}$$

$$\hat{\psi^I_{j,0}}(\omega) = \sqrt{2^{-j}}\hat{\psi^I}(2^{-j}\omega)e^{ij\omega} \tag{7.9}$$

$$H^I_j(\omega) = \frac{\hat{\psi^I_{j,0}}(\omega)}{\hat{\phi^I}(\omega)} \tag{7.10}$$

$$\hat{\phi^I_{j,0}}(\omega) = \sqrt{2^{-j}}\hat{\phi^I}(2^{-j}\omega) \tag{7.11}$$

$$G^I_j(\omega) = \frac{\hat{\phi^I_{j,0}}(\omega)}{\hat{\phi^I}(\omega)} \tag{7.12}$$

ただし，$\hat{\phi^M}(\omega)$ は Meyer のスケーリング関数，$\hat{\psi^M}(\omega)$ は Meyer のウェーブレット，$\hat{\phi}_{j,0}(\omega)(j = -1, -2, ..., J+1, J)$ は各レベルのスケーリング係数，$\hat{\psi}_{j,0}(\omega)$ は各レベルの

ウェーブレット，$G_j(\omega)$, $H_j(\omega)$ は各レベルで使用する再構成数列の役割を持つフィルタであり，R, I はそれぞれ実数部，虚数部を表す．また式 (7.4), (7.6), (7.10), (7.12) 中の $\phi^R_{j,0}(t)$, $\phi^I_{j,0}(t)$, $\psi^R_{j,0}(t)$, $\psi^I_{j,0}(t)$ は次式より求められた．

$$\phi^R_{j,0}(t) = \sum_k g^R_k \phi^M_k(t) \tag{7.13}$$

$$\phi^I_{j,0}(t) = \sum_k g^I_k \phi^I_k(t) \tag{7.14}$$

$$\psi^R_{j,0}(t) = \sum_k h^R_k \phi^M_k(t) \tag{7.15}$$

$$\psi^I_{j,0}(t) = \sum_k h^I_k \phi^I_k(t) \tag{7.16}$$

ただし，$\{g_k\}, \{h_k\}$ は再構成数列を示し，k はデータ数を示す．式 (7.13) から 7.16 は Meyer のスケーリング関数を再構成数列と畳み込みすることにより，分解レベル-j で分解した高・低周波部分のスケーリング係数やウェーブレットが得られることを示したものである．これらの式の両辺に対してフーリエ変換を行い，式変形することにより，式 (7.4), (7.6), (7.10), (7.12) が得られる．つまり，分解レベル $-j$ のウェーブレット係数やスケーリング係数をフーリエ変換した $\hat{\psi}_{j,0}(\omega)$,$\hat{\phi}_{j,0}(\omega)$ から Meyer のスケーリング関数をフーリエ変換した $\hat{\phi^M}(\omega)$ を割ることによって各レベルの再構成数列のフーリエ変換した $G_j(\omega)$, $H_j(\omega)$ が得られることを示している．式 (7.2) と (7.3) で $\hat{\psi}_{j,0}(\omega)$，(7.5) で $\hat{\phi}_{j,0}(\omega)$ を求めることで $G_j(\omega)$, $H_j(\omega)$ を導出した．

式 (7.2) から (7.12) で作られる再構成数列 $G_j(\omega)$, $H_j(\omega)$ は Fig. 19 のように配置される．ただし，分解レベルは-3 である ($j = -1, -2, -3$)．一方，求めた $G_j(\omega)$, $H_j(\omega)$ を逆フーリエ変換し再構成数列が得られ，以下の式により分解数列 $\{a_k\}, \{b_k\}$ が得られる．

$$a_k = g_{N-k} b_k = h_{N-k} \tag{7.17}$$

ただし，$(k = 1, 2, ..., N-1, N)$ である．これによりアップサンプリングを行わず SWT の分解・再構成数列が得られた．

7.3 SWT を適用した P-DWT の分解・再構成方法

前節で P-DWT に SWT の理論を組み込むために必要な分解・再構成数列の作成方法を述べた．SWT を P-DWT に適用するために変更する処理を以下に示す．

- 寄生フィルタ作成時に，従来は節 5.3 のように DWT を用いて，ダウンサンプリングの影響を受けないように寄生フィルタの作成を行っていた．一方，SWT の場合，

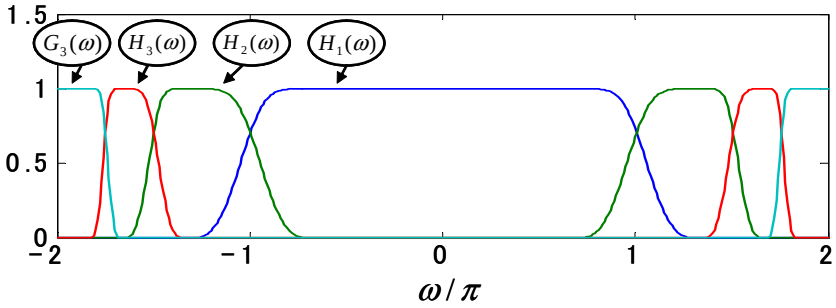


Fig. 19. Redundant filter

すでにアップサンプリングによる影響を受けない分解方法を確立しているため，この分解数列を用いて抽出対象の周波数域をもつ BP-RMW を分解し，それを寄生フィルタとする．

- 従来は DWT の理論により，分解・再構成を BP-RMW の周波数域を持つ寄生フィルタにより行っていた．SWT を適用した P-DWT ではこれを SWT の理論を応用したものによって行う．

SWT の分解・再構成方法はアップサンプリングを行わない分解・再構成数列を用いるため，節 7.1 で述べた分解・再構成方法と異なる．次にその分解・再構成方法を示す．

$$
\begin{aligned}
d_j^R &= \sum_k u_{j,k}^R c_{0,k}^R \\
c_J^R &= \sum_k u_{J+1,k}^R c_{0,k}^R \\
d_j^I &= \sum_k u_{j,k}^I c_{0,k}^I \\
c_J^I &= \sum_k u_{J+1,k}^I c_{0,k}^I
\end{aligned}
\tag{7.18}
$$

$$
\begin{aligned}
c_0^R &= \sum_j \sum_k \acute{u}_{j,k}^R c_{j,k}^R + \sum_k \acute{u}_{J+1,k}^R c_{J,k}^R \\
c_0^I &= \sum_j \sum_k \acute{u}_{j,k}^I c_{j,k}^I + \sum_k \acute{u}_{J+1,k}^I c_{J,k}^I
\end{aligned}
\tag{7.19}
$$

ただし,$d_j, (j = -1, -2, ..., J-1, J)$ は各レベルのウェーブレット係数，c_J は最低レベルのスーケーリング係数，$\{u_j\}, \{u_{J+1}\}$ は BP-RMW や BR-RMW をアップサンプリングしない分解数列にて分解して得た寄生フィルタ，$\{\acute{u}_j\}, \{\acute{u}_{J+1}\}$ は再構成のために $\{u_j\}, \{u_{J+1}\}$ を再構成数列のように変化させたもの，c_0 は補間によって得たレベル 0 のスケーリング関数を示す．その分解・再構成を模式的に示した図を Fig.20 に示す．ただし $f(t)$ は抽出対象となる信号，$\acute{f}(t)$ は抽出された信号を示す．これにより，ナイキスト周波数はシフトせず，エネルギー欠損をする周波数領域が発生しない P-DWT の信号抽出が可能になる．

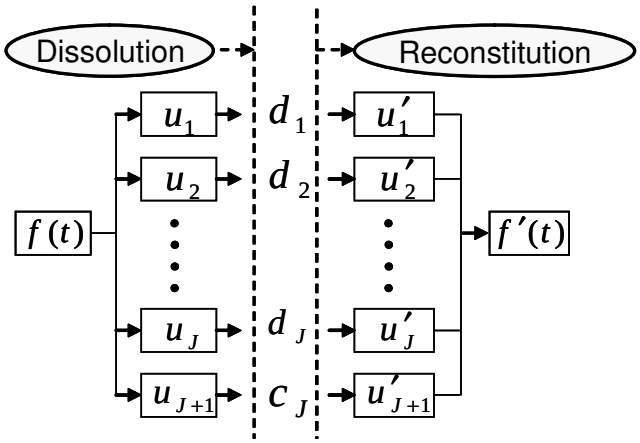


Fig. 20. Process of P-DWT(SWT)

7.4 SWT を適用した P-DWT の評価

従来の DWT を用いた P-DWT と提案した SWT を用いた P-DWT を用い，エネルギー損失の改善を評価する下記の数値実験を行う．

式 (7.21) の $f_1(t)$ はエネルギー損失が起こりやすい，各バンド間の過渡領域の周波数成分を持つ信号で，$f(t)$ は $f_1(t)$ を含むチャープ信号である．またサンプリング周波数は 4096[Hz]，分解レベルは-3 までである．

(7.20) $$f(t) = f_1(t) + sin(2\pi 700t)$$

(7.21) $$f_1(t) = sin(2\pi 1050t) + cos(2\pi 525t) + sin(2\pi 260t)$$

各 P-DWT で $f_1(t)$ を抽出するために $f(t)$ を分解再構成することで，抽出された信号である $\acute{f}_1(t)$ を得る．そして，$f_1(t)$ と $\acute{f}_1(t)$ を次式による平均二乗誤差を用いて比較することで抽出精度を評価する．

(7.22) $$error = 10log\left(\frac{\left(\sum_k \left(\acute{f}_{1,k}(t) - f_{1,k}(t)\right)\right)}{N}\right)$$

ただし，$k(k = 1, 2, ..., N-1, N)$ はデータ数，$erorr$ は平均二乗誤差である．

抽出対象の信号 $f(t)$ から抽出した信号 $f_1(t), \acute{f}_1(t)$ を引くことでエネルギー損失の結果を 21 に示す．ただし図 (a) には式 (7.21) に示した $f(t)$ と $f_1(t)$ との差，図 (b) には $f(t)$ と DWT により得られた $f_1(t)$ との差，図 (c) には $f(t)$ と SWT により得られた $f_1(t)$ とのさを示す．図示のように，従来手法で抽出した結果である図 (b) には 0-1000,2000-3000[sec] の間に振動が見られる．これがエネルギー損失によって生じた誤差である．一方，図 (c) ではこの誤差は発生しておらず，エネルギー損失は発生していないと言える．また，従来の P-DWT では平均二乗誤差は-30.28[dB] であったのに対し，SWT を適用した P-DWT は-117.06[dB] となった．これは，$f_1(t)$ と比較して，従来手法では全体で平均して 4.8×10^{-2}

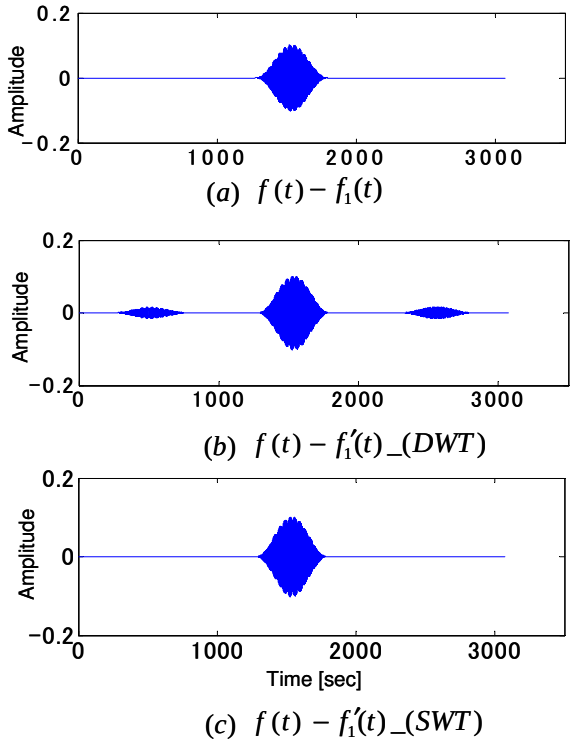


Fig. 21. Resulf of energy loss

の誤差が生じているのに対し，SWT を適用した P-DWT では 8.2×10^{-6} の誤差のみ生じていることを示しており，エネルギー損失が発生しないことで精度が向上したことが言えた．以上より，エネルギー損失問題を解決したことが言えた．

8. 終わりに

著者らは最初異常信号を検出すために，異常信号による実信号マザーウェーブレット (Real-signal mother wavelet, RMW) の構成法を提案し，それを用いた連続ウェーブレット変換から得られたスケール $a = 1$ のウェーブレット係数の絶対値をウェーブレット瞬時相関 (Wavelet instantaneous correlation, WIC) と定義した．そしてこれを離散ウェーブレット変換 (Discrete Wavelet Transform：DWT) に適用する寄生的離散ウェーブレット変換 (Parasitic Discrete Wavelet：P-DWT) も提案し，それによる高速ウェーブレット瞬時相関を実現した [7]．さらに P-DWT を異常信号の抽出に拡張し，抽出対象の周波数帯を持つ帯域通過型実信号マザーウェーブレット (Band Pass-RMW：BP-RMW) とそれ以外の周波数帯を持つ帯域除去型実信号マザーウェーブレット (Band Rejection-RMW：BP-RMW) を設計し P-DWT に適用し，異常信号の検出に強く，再構成が可能な寄生的離散ウェーブレット変換が可能となった．

ところで，P-DWT に BP-RWM や BR-RMW を適用したことによって特定の周波数領域に存在する周波数成分のエネルギー損失が発生した．これにより信号抽出の精度が低

下し，正確な結果を得ることができないという問題が生じている．この問題を解決するために，エネルギー損失の原因を解明し，定常ウェーブレット変換 (SWT) の理論を用いてナイキスト周波数がシフトしない P-DWT を提案し，この問題の改善を試みした．その結果，エネルギー損失問題を解決することができたが，ダウンサンプリングを行わない SWT の理論を用いたことによって計算量が増加した．よって P-DWT をリアルタイム性を求める処理に適用する場合には，これを解決しなければならない．今後この問題を解決し，P-DWT がもっと広く応用できるように目指している．

謝辞 本研究の一部は独立行政法人日本科学技術振興会 H22 年科学研究費補助金の助成を受けて行われた．ここで付記して，謝意を表します．

参考文献

[1] Allen, R.L. and Mills, D.W., Signal Analysis, *IEEE Press*, 2004.

[2] 榊原 進，“ウェーブレットビギナーズガイド”， 東京電機大学出版局,1998.

[3] I. Daubechies, Ten Lectures on Wavelets, *SIAM, Philadelphia, PA*, 1992.

[4] Ishimitsu, S., Kitagawa, H., Horihata, S., Active control of the room noise of ship by the correction analysis of measured wavelet, *Proc. of inter.noise*, Vol.3(2000), 1768–1771.

[5] 章 忠, 戸田浩, 川畑洋昭， 実信号マザーウェーブレットおよびその異常信号抽出への応用， 日本機械学会論文集 C 編, Vol.70, No.696(2004), 2360–2367,

[6] 章 忠, 池内宏樹, 石井秀明, 堀畑 聡, 今村 孝, 三宅哲夫，実信号マザーウェーブレットおよびその異常信号抽出への応用 (平均的複素数実信号マザーウェーブレットの設計とその応用)， 日本機械学会論文集 C 編, Vol.73, No.730(2007), 1676–1683,

[7] 章 忠，池内宏樹, 斎木典保, 今村 孝, 石井秀明, 戸田 浩, 三宅 哲夫，寄生的離散ウェーブレット変換およびその異常信号検出への応用，日本機械学会論文集 C 編， Vol.75, No.757(2009),pp.2529-2536.

[8] Z. Zhang, N. Saiki, H. Toda, T. Imamura and T. Miyake, Parasitic Decrease Wavelet Transform and its Application on denoising, Proc. of International Conference of Wavelet Analysis and Pattern Recognition(ICWAPR 2010), 345–350.

[9] H. Toda, Z. Zhang and T. Imamura, The Design of Complex Wavelet Packet Transforms Based on Perfect Translation Innariance Theorems, International Journal of Wavelets, Multiresolution and Information Processing, Vol.8(2010), Issue 4, 537–558.

[10] 戸田 浩, 章 忠, 完全シフト不変性を実現する複素数離散ウェーブレット変換, Journal of Signal Processing 「信号処理」，Vol.12，No.3(2008)，155–166.

[11] 大滝仁，章 忠，戸田浩，今村孝，三宅哲夫，石川康宏，寄生的離散ウェーブレット変換による異常信号抽出報に関する検討，日本機械学会東海支部第 60 期総会講演会 (2011.03)，109–110.

[12] 芦野 隆一, 馬場 則夫, 守本 晃, ウェーブレットの基礎理論とそのニューラルネットへの適用-動的環境における学習と適応と中心として (改訂版), 第 47 回システム制御情報学会研究発表講演会, 大阪教育大学, 2003.

[13] H. Toda, Z. Zhang, Variable-Density Complex Discrete Wavelet Transform Based on Perfect Translation Invariance, Proceedings of the 2008 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR 2008), 711–716.

章　忠 (豊橋技術科学大学工学部)
〒441-8580 豊橋市天伯町雲雀ケ丘 1-1
E-mail: zhang@is.me.tut.ac.jp

大滝　仁 (豊橋技術科学大学大学院)
〒441-8580 豊橋市天伯町雲雀ケ丘 1-1
E-mail: ohtaki@is.me.tut.ac.jp

今村 孝 (豊橋技術科学大学工学部)
〒441-8580 豊橋市天伯町雲雀ケ丘 1-1
E-mail: ima@is.me.tut.ac.jp

三宅 哲夫 (豊橋技術科学大学工学部)
〒441-8580 豊橋市天伯町雲雀ケ丘 1-1
E-mail: miyake@is.me.tut.ac.jp

戸田　浩 (豊橋技術科学大学工学部)
〒441-8580 豊橋市天伯町雲雀ケ丘 1-1
E-mail: pxt00134@nifty.com

地下構造探査における時系列処理の現状と問題点

三ケ田　均 *
* 京都大学大学院工学研究科

概要.　地下構造探査におけるウェーブレットは基本波形と呼ばれ，構造探査の根幹を成す概念である。この概念を活かし、地下構造の探査においては自己回帰モデルが多用される。このモデルの応用において用いられる手法には，赤池モデルおよびバーグモデルと呼ばれる２種類のモデルが存在する。両者の意義および２つのモデルの差異について概説し，今後考慮すべきウェーブレット処理について議論する。

Time Series Analysis in Subsurface Geophysical Exploration

Hitoshi Mikada*
*Graduate School of Engineering, Kyoto University

Abstract.　Wavelet is in general called as basic waveform and is a very important concept in geophysical structural exploration. This concept is used with autoregressive modeling in the exploration of subsurface structure. In the application of the concepts, there exist two kinds of models called the Akaike and the Burg models. The difference between the two models is outlined, and the wavelet processing that should be considered in the future is discussed in terms of wavelet processing.

1.　はじめに

本稿は，文部科学省の実施する平成 23 年度「数学・数理科学と諸科学・産業との連携による数学イノベーションの推進」により大阪教育大学で開催されるワークショップのプロシーディング論文として作成される。

反射法地震探査等では，多くの信号処理理論が用いられる。これは，地表で地震波を発生させ，地下からの反射波を捕らえるという条件下で，得られた地震波形が一次元の「たたみ込みモデル」として記載可能であるからである。即ち，たたみ込みモデルでは，地表で発生した地震波という時系列と地下の反射係数列という時系列がたたみ込みにより得られる，という仮定を置くことができるからである。更に 1960 年代より始まったディジタル革命は，それまでのアナログ信号を扱う状況を一変させ，信号がディジタル化されただけでなく，データ処理上で時間軸を反転させることも可能にした。この状況下，信号処理技術も大きく発展し，基本波形を他の波形に変換するウィーナ・フィルタの原理も進化した。このディジタルデータ処理の基本的な事項をまとめ，その問題点を議論することは今後の工学への応用数学の導入を考える際に非常に意味があると考えられる。

たたみ込みモデルをディジタル化する際に必要となったモデルは，グーピョ・モデル [1]

と呼ばれる。これはデータの離散化の際に，必ずしも時間関数にとり好ましい点のサンプルを得られないことから必要となる。工学に応用されるに当たり必要となった [2] の理論，そしとしてその意義を説明する。

2. 構造探査のたたみ込みモデルと限界

2.1 グーピョ・モデル

反射法地震探査などにより地表で観測される地震波は時系列として得られる。一般に，地下の媒質は物性に差異により深さとともに異なる地震波速度を持つため，反射波がどの深さの反射面で生じたかを知るためには，地震波速度を求める必要がある。他方，時間記録として得られる地震波の時間は，「時間深度（time-depth）」として表記されるように，時間そのものが深さを示していることも事実である。グーピョー・モデル（Goupillaud, 1961）は水平成層構造を仮定し，離散化された時系列のサンプル時間ごとに反射係数を有するとするモデルを提唱した（Fig.1）。離散化により，サンプル時間に一致しない時間にあるスパイク状の関数は，その前後のサンプル時間における複数のスパイクで近似できる（Fig.2）ことから，グーピョー・モデルは離散化されたモデルとして妥当であると考えることができる。この考え方により，地震波記録が地表に対して鉛直に入射する場合，時間深度に対する地震波処理は全て信号処理で代用できることとなる。換言すれば，離散化された時間に関係なく地層境界（反射面）ごとに生じるスパイク状の反射係数列を持つ時間関数を離散化することが可能である。この操作により，離散化というスペクトルの帯域を狭める操作を導入しても，反射係数列を表現することが可能となる。従って反射法地震探査記録の畳み込みモデルの妥当性が確立されただけでなく，各種の信号処理方法の適用が行なわれることとなった。即ち，入力となる地震波，反射係数列，反射地震波記録をそれぞれ z 変換 ($z = e^{i\omega}$) を用いて，$x(z)$，$r(z)$，$s(z)$ とすると，以下の式が成り立つ。

$$s(z) = x(z)r(z)$$

この式の右辺は，たたみ込みを示している。構造探査の対象となる地下構造は $r(z)$ であり，この式は，入力地震波 $x(z)$ が既知であれば，出力 $s(z)$ から，得たい $r(z)$ を推定可能であることを示している。

この基本的な理論は，上述グーピョ・モデルを基にしており，最近進化の著しい Daylight Seismic Processing として知られる理論の基礎となる Kunetz-Claerbout 方程式 [3] の導入にも用いられている。

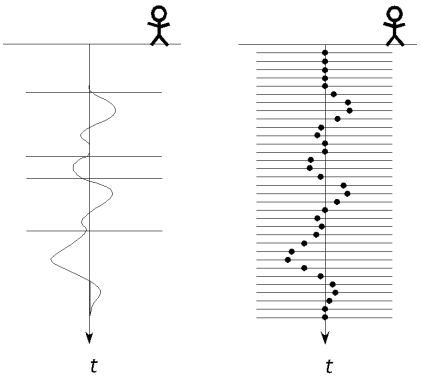


Fig. 1. Typical reflection seismogram that is composed of the basic waveform convolved with the underground reflectivity series (Left). Right figure depicts that a descretized time series is obtained after sampling by equispaced time.

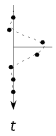


Fig. 2. Samples denoted by solid circles indicate discretized waveform. Any impulsive signal could be expressed by a group of samples that give equivalent spectrum in the range of frequencies inside alias frequency.

2.2 信号処理方法

グーピョ・モデルの導入により地下構造を推定することが可能であることは示されたが，実は，帯域制限という問題が残っている。Rayleigh の原理によれば，こうした解析の解像力を決定するのは，使用する周波数（上述の ω）の帯域の広さである。残念ながら，離散化による高周波カットという帯域制限，そして測定機器の応答，更に地球の媒質の加える高周波成分の減衰，といった一連の物理過程は，上述の $w(z)$ および $s(z)$ 双方を狭帯域信号に変化させてしまう。このため，構造探査で得られた記録から真の $r(z)$ を再現するためには，各種信号処理手法が必要となる。この処理手法に最も多用される方法が，ウィーナ・フィルタである。

このウィーナ・フィルタでは，上述の $x(z)$ を変換するフィルタ $h(z)$ を，ある理想的な波形 $d(z)$ に最小二乗法的に推定する。即ち，次式を用いる。

$$d(z) = h(z)x(z)$$

そして，このフィルタを $s(z)$ に適用することで，ある理想的な入力に対する出力波形 $\tilde{s}(z)$ を推定する。

$$\tilde{s}(z) = h(z)x(z)r(z) \tag{2.1}$$

$$= (h(z)x(z))r(z) \tag{2.2}$$

$$= d(z)r(z) \tag{2.3}$$

2.3 ウィーナ・フィルタ

上述のウィーナ・フィルタは，最小二乗フィルタとして認識される。簡単のため，$x(z)$ を次式の数列で記載する。その他の $d(z)$ や $h(z)$ も同様である。

$$x(z) = \sum_j x_j z^j = \{x_j\}$$

この時，たたみ込みを用いて次式が成り立つ。

$$d_i = \sum_j h_j x_{i-j}$$

実際にフィルタを適用した際には，推定誤差 e_i を生じるので，次式となる。

$$d_i = \sum_j h_j x_{i-j} + e_i$$

最小二乗の考え方を導入すれば，このフィルタ h_j を推定するには，誤差の二乗和 $\sum_j e_i^2$ を細小にすれば良い。従って，次式を解けば良い。

$$\frac{\partial}{\partial h_k} \sum_j e_i^2 = 0$$

但し，ここで k は，h_j の次数を m とした時に，$1 < k \leq m$ である。次数が m とした時に h_j を $h_j^{(m)}$ として，整理すると次式が導かれる。

$$\sum_{j=0}^{m} h_j^{(m)} \sum_i x_{i-j} x_{i-k} = \sum_i d_i x_{i-k}$$

従って，整理すると次式となる。

$$\sum_{j=0}^{m} h_j^{(m)} \phi_{k-j} = c_k$$

但し，ϕ および c は，それぞれ次式で示される自己相関および相互相関係数列である。

$$\phi_k = \sum_i x_i x_{i+k}$$

$$c_k = \sum_i x_i d_{i+k}$$

行列を用いると，次式となる。

(2.4)
$$\begin{pmatrix} \phi_0 & \phi_1 & \cdots & \phi_m \\ \phi_1 & \phi_0 & \cdots & \phi_{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_m & \phi_{m-1} & \cdots & \phi_0 \end{pmatrix} \begin{pmatrix} h_0^{(m)} \\ h_1^{(m)} \\ \vdots \\ h_m^{(m)} \end{pmatrix} = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_m \end{pmatrix}$$

左辺の行列は，次数の高い場合直接解くことが困難な行列であり，Toeplitz の行列と呼ばれる。

3. Levinson 漸化式による Toeplitz 行列の解法

3.1 自己回帰モデル

Levinson 漸化式は，自己相関モデルを用いて解くことができる。Fig.3 のような時間的に定常な時系列 x_i を考えると，自己相関モデルでは過去の時系列サンプルから $t = 0$ という現在のサンプルを予測することができるとする。

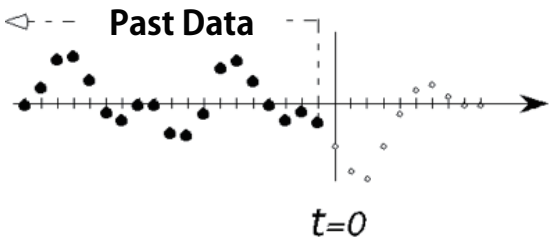


Fig. 3. Time series and autoregressive prediction of sample at $t = 0$. solid and open circles depicts samples of the known past and unknown future, respectively.

今，m 次のフィルタ $a_i^{(m)}$（自己回帰係数と呼ばれる）を用いて，過去のデータから現在のデータを予測できるとする。即ち，次式が成り立つ。

$$\begin{aligned} x_n &= a_1^{(m)} x_{n-1} + a_2^{(m)} x_{n-2} + a_3^{(m)} x_{n-3} + \ldots \\ &= \sum_{i=1}^{m} a_i^{(m)} x_{n-i} + e_n^{(m)} \end{aligned}$$

但し，$e_n^{(m)}$ の自己回帰係数を用いた時の x_n の推定誤差である。この $e_n^{(m)}$ の二乗和を P_m とすると，次式が成り立つ。

(3.1)
$$P_m = \sum_n \left(e_n^{(m)}\right)^2$$

前節と同様に，この P_m を最小化する $a_i^{(m)}$ を最小二乗法的に推定する式を導出すると，次式となる。

$$\begin{pmatrix} \phi_0 & \phi_1 & \cdots & \phi_{m-1} \\ \phi_1 & \phi_0 & \cdots & \phi_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m-1} & \phi_{m-2} & \cdots & \phi_0 \end{pmatrix} \begin{pmatrix} a_1^{(m)} \\ a_2^{(m)} \\ \vdots \\ a_m^{(m)} \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_m \end{pmatrix} \tag{3.2}$$

この式を解く方法が Levinson 漸化式である。先ず，自己相関係数の代わりに，予測誤差係数 $\alpha_i^{(m)}$ を導入する。この時，次式を用いる。

$$\begin{aligned} \alpha_0^{(m)} &= 1 \\ \alpha_i^{(m)} &= -a_i^{(m)} \quad (i = 0, m) \end{aligned}$$

この時，行列の式は以下に変形できる。

$$\begin{pmatrix} \phi_1 & \phi_0 & \phi_1 & \cdots & \phi_{m-1} \\ \phi_2 & \phi_1 & \phi_0 & \cdots & \phi_{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_m & \phi_{m-1} & \phi_{m-2} & \cdots & \phi_0 \end{pmatrix} \begin{pmatrix} \alpha_0^{(m)} \\ \alpha_1^{(m)} \\ \vdots \\ \alpha_m^{(m)} \end{pmatrix} = 0 \tag{3.3}$$

この式の第一列および第 m 列を右辺に移動し，更に第 m 行を削除すると次式となる。

$$\begin{pmatrix} \phi_0 & \cdots & \phi_{m-2} \\ \vdots & \ddots & \vdots \\ \phi_{m-2} & \cdots & \phi_0 \end{pmatrix} \begin{pmatrix} \alpha_1^{(m)} \\ \alpha_2^{(m)} \\ \vdots \\ \alpha_{m-1}^{(m)} \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{m-1} \end{pmatrix} + \alpha_m^{(m)} \begin{pmatrix} \phi_{m-1} \\ \phi_{m-2} \\ \vdots \\ \phi_1 \end{pmatrix}$$

次数の低い $(m-1)$ 次の場合の基本式は次式となる。

$$\begin{pmatrix} \phi_0 & \phi_1 & \cdots & \phi_{m-2} \\ \phi_1 & \phi_0 & \cdots & \phi_{m-3} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{m-2} & \phi_{m-3} & \cdots & \phi_0 \end{pmatrix} \begin{pmatrix} a_1^{(m-1)} \\ a_2^{(m-1)} \\ \vdots \\ a_{m-1}^{(m-1)} \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{m-1} \end{pmatrix}$$

従って，整理すると次の漸化式が求まる。

$$\begin{pmatrix} \alpha_1^{(m)} \\ \alpha_2^{(m)} \\ \vdots \\ \alpha_{m-1}^{(m)} \end{pmatrix} = \begin{pmatrix} \alpha_1^{(m-1)} \\ \alpha_2^{(m-1)} \\ \vdots \\ \alpha_{m-1}^{(m-1)} \end{pmatrix} + \alpha_m^{(m)} \begin{pmatrix} \alpha_{m-1}^{(m-1)} \\ \alpha_{m-2}^{(m-1)} \\ \vdots \\ \alpha_1^{(m-1)} \end{pmatrix} \tag{3.4}$$

これが Levinson 漸化式である。この式を用いれば，数学的帰納法を用い，$\alpha_m^{(m)}$ の推定させできれば，順次高次の予測誤差係数を推定することができる。この $\alpha_m^{(m)}$ の推定方法には，Yule-Walker 法と Burg 法の 2 種類が提案されている。

3.2 赤池モデル

[4] は，Yule-Walker 法を用いた解法を提案している。式 (3.1) を変形すると，次式が導かれる。

$$P_m = \sum_{i=0}^{m} \alpha_i^{(m)} \phi_i$$

従って，この式を式 (3.3) の第一行に挿入し，次式が導かれる。

$$\begin{pmatrix} \phi_0 & \cdots & \phi_m \\ \vdots & \ddots & \vdots \\ \phi_m & \cdots & \phi_0 \end{pmatrix} \begin{pmatrix} \alpha_0^{(m)} \\ \alpha_1^{(m)} \\ \vdots \\ \alpha_m^{(m)} \end{pmatrix} = \begin{pmatrix} P_m \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{3.5}$$

式 (3.4) を変形すると，次式が導かれる。

$$\begin{pmatrix} \alpha_0^{(m)} \\ \alpha_1^{(m)} \\ \vdots \\ \alpha_m^{(m)} \end{pmatrix} = \begin{pmatrix} \alpha_0^{(m-1)} \\ \alpha_1^{(m-1)} \\ \vdots \\ \alpha_{m-1}^{(m-1)} \\ 0 \end{pmatrix} + \alpha_m^{(m)} \begin{pmatrix} 0 \\ \alpha_{m-1}^{(m-1)} \\ \alpha_{m-2}^{(m-1)} \\ \vdots \\ \alpha_1^{(m-1)} \end{pmatrix} \tag{3.6}$$

この式 (3.6) の両辺に m 次の Toeplitz 行列をかけ，式 (3.5) を用いると次式となる。

$$\begin{pmatrix} P_m \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} P_{m-1} \\ 0 \\ \vdots \\ 0 \\ \Delta_{m-1} \end{pmatrix} + \alpha_m^{(m)} \begin{pmatrix} \Delta_{m-1} \\ 0 \\ \vdots \\ 0 \\ P_{m-1} \end{pmatrix} \tag{3.7}$$

但し，ここで次式が成り立つ。

$$\Delta_{m-1} = \sum_{i=0}^{m} \alpha_{m-i}^{(m)}$$

この関係を用い，$\alpha_m^{(m)}$ を推定することができる。

$$\begin{aligned} \alpha_m^{(m)} &= -\Delta_{m-1}/P_{m-1} \\ P_m &= P_{m-1}\left(1 - \left(\alpha_m^{(m)}\right)^2\right) \end{aligned} \tag{3.8}$$

3.3 Burg モデル

上述のモデルでは，式 (3.7) および (3.8) を用い，与えられた自己相関係数列から帰納法的に m 次の $\alpha^{(m)}$ および ¶$_m$ が推定される。しかしながら自己相関係数推定に用いられる時系列 x_n が有限長で切断されている場合，自己相関係数そのものに切断の影響が出てバイアスされている可能性がある。[2] は，切断された時系列の自己相関係数を直接推定せず，与えられた時系列から直接 $\alpha_m^{(m)}$ および予測誤差 P_m を推定する方法を提唱した。　さて，自己相関係数は前述の通り $\phi_i = \phi_{-i}$ を満たす偶関数なので，予測誤差フィルタは正の時間の向きに適用しても負の時間の向きに適用しても，同様な予測誤差を与える筈である。与えられる時系列の項数を N（すなわち時系列は $x_0, x_1, \ldots, x_{N-1}$），$m$ 次の予測誤差フィルタを正および負の時間の向きに適用した場合の予測誤差系列をそれぞれ $F_i^{(m)}$，$B_i^{(m)}$ とし，この正および負の時間向きの予測誤差二乗和をそれぞれ，$F^{(m)}$，$B^{(m)}$ とする。すなわち，次式が成り立つ。

$$\begin{aligned} F^{(m)} &= \textstyle\sum_{i=0}^{N-m-1} \left(F_i^{(m)}\right)^2 \\ B^{(m)} &= \textstyle\sum_{i=0}^{N-m-1} \left(B_i^{(m)}\right)^2 \end{aligned} \tag{3.9}$$

但し，予測誤差系列は，以下の式で導かれる。

$$\begin{aligned} F_i^{(m)} &= \textstyle\sum_{j=0}^{m} \alpha_j^{(m)} x_{m-j+i} \\ B_i^{(m)} &= \textstyle\sum_{j=0}^{m} \alpha_j^{(m)} x_{j+i} \end{aligned} \tag{3.10}$$

この式に，式 (3.6) の漸化式関係を代入すると，次式が導かれる。

$$\begin{aligned} F_i^{(m)} &= F_i^{(m-1)} \alpha_m^{(m)} B_i^{(m-1)} \\ B_i^{(m)} &= B_i^{(m-1)} \alpha_m^{(m)} F_i^{(m-1)} \end{aligned} \tag{3.11}$$

係数 $\alpha_m^{(m)}$ の推定には，式 (3.9) を用い，最小二乗法的に次式を用いる。

$$\frac{\partial (F^{(m)} + B^{(m)})}{\partial \alpha_m^{(m)}} = 0$$

この式を評価すると，次式となる。

$$\alpha_m^{(m)} = -2 \frac{\sum_{i=0}^{N-m-1} F_i^{(m-1)} B_i^{(m-1)}}{\sum_{i=0}^{N-m-1} \{ \left(F_i^{(m-1)}\right)^2 \left(B_i^{(m-1)}\right)^2 \}} \tag{3.12}$$

従って，自己相関係数を用いることなく，予測誤差系列を推定することが可能である。

4. ウィーナ・フィルタの推定

式 (2.4) で推定された最小二乗法の正規方程式を解くためには，上述の予測誤差係数列を用いる必要がある。式 (3.4) の導出に用いた方法を使って，式 (3.4) と同様な漸化式関係を導くと次式となる。

$$\begin{pmatrix} h_1^{(m)} \\ h_2^{(m)} \\ \vdots \\ h_{m-1}^{(m)} \end{pmatrix} = \begin{pmatrix} h_1^{(m-1)} \\ h_2^{(m-1)} \\ \vdots \\ h_{m-1}^{(m-1)} \end{pmatrix} + h_m^{(m)} \begin{pmatrix} \alpha_{m-1}^{(m-1)} \\ \alpha_{m-2}^{(m-1)} \\ \vdots \\ \alpha_1^{(m-1)} \end{pmatrix} \tag{4.1}$$

即ち，ウィーナ・フィルタが，予測誤差係数列とほぼ同一の漸化式関係を満たすこと，やはり同様に $h_m^{(m)}$ の推定方法が必要であること，そして予測誤差係数列と密接に関係していることが導かれた。

このウィーナ・フィルタ導出にも，予測誤差係数列と同様に，赤池モデルに代表される Yule-Walker 法と Burg 法の 2 種類の手法が存在する。

4.1 赤池モデル

予測誤差係数列で用いた方法をそのまま適用し，次式を得る。

$$\begin{aligned} \Gamma_m &= \Gamma_{m-1} - h_m^{(m)}(1 - P_m) \\ h_m^{(m)} &= (c_m - \Gamma_{m-1})/P_{m-1} \end{aligned} \tag{4.2}$$

但し，ここで Γ_m は次式を満たす。

$$\Gamma_m = \sum_{i=1}^{m} h_i^{(m)} \phi_{m-i}$$

4.2 Burg モデル

データ切断のバイアスが含まれると考え，相互相関係数を用いないようにする。この場合，先ずフィルタの推定誤差を次式で示す。

$$e_i^{(m)} = d_i - \sum_{j=0}^{m} h_j^{(m)} x_{i-j} \tag{4.3}$$

推定誤差の二乗和を最少にする $h_m^{(m)}$ を推定する。即ち，$h_m^{(m)}$ は次式を満たす。

$$\frac{\partial \left(\sum_{i=0}^{N-m-1} e_i^{(m)} \right)^2}{\partial h_m^{(m)}} = 0$$

従って，次の式が導かれる。

$$
\begin{aligned}
e_i^{(m)} &= e_i^{(m-1)} - h_m^{(m)} B_i^{(m)} \\
h_m^{(m)} &= \textstyle\sum_{i=0}^{N-m-1} e_i^{(m)} B_i^{(m-1)} / \sum_{i=0}^{N-m-1} (B_i^{(m-1)})^2
\end{aligned}
\tag{4.4}
$$

5. まとめ

数学的に良く記載された技術であっても，実際のデータへの適用では，データの切断やデータに混入するノイズの影響を考慮した処理が必要となる。当然存在するべき自己相関関数や相互相関関数にさえ，データ切断によるバイアスがかかる可能性を否定することはできないからである。[2] は，短時間データからスペクトルをどのように推定するかという問題から，ウェーブレットと呼ばれる基本波形の推定にも影響する理論を構築した。今後も，実際のデータを処理する際には，こうしたノイズや切断誤差に十分配慮したデータ処理方法への改編が必要になることもあると考えられる。

謝辞 大阪教育大学の芦野隆一先生には，本シンポジウムへの参加のご打診から，筆の遅い著者の作業を暖かくお見守り戴く等，筆舌に尽くし難いお世話を戴いた。この場をお借りして，心から感謝の意を表したい.

参考文献

[1] Gaupillaud, P.L., An approach to inverse filtering of near surface layer effects from seismic records, Geophysics, **26** (1961), 754–760.

[2] Burg, J.P., Maximum entropy spectral analysis, Proc. 37th Meet. Soc. Exploration Geophysicists, (1967), 34–39.

[3] Claerbout, J.F., Fundamentals of Geophysical Data Processing, Blackwell Science Publ., New York, 274pp, 1976.

[4] 赤池弘次・中川東一郎, ダイナミックシステムの統計的制御と解析, サイエンス社，東京, 179pp, 1972.

三ケ田　均 (京都大学大学院工学研究科)
〒615-8540 京都府京都市西京区京都大学桂４
E-mail: mikada@tansa.kumst.kyoto-u.ac.jp

ウェーブレットによるコンテンツ保護

新井康平*
*佐賀大学大学院工学系研究科知能情報システム学部門

概要．ウェーブレットによるコンテンツ保護に係る従来方法の問題は、(1)コンテンツの保護のための埋め込み情報の秘匿性、視認可能性、(2)流通コンテンツの改ざん、加工処理等に対する耐性がその主なものとして挙げられる。本稿はこれらの問題に対する解決手段を論ずるものである。

Content protection method (Data Hiding) based on wavelets

Kohei Arai,
Graduate School of Science and Engineering, Saga University

Abstract. One of the major problems on wavelet based content protection methods is invisibility of the secret information which is embedded in the circulation contents followed by robustness of the method against information manipulations or processing to the circulation contents. This paper describes some solutions to solve the aforementioned problems.

１．まえがき

DVD の違法コピーや音楽の放送・放映ごとの課金等が社会的問題になっている。これを解決するためには、コンテンツID、デジタル署名を秘匿性高く行う必要がある。また、コピー禁止や一度のみのコピー許可等の世界標準方式が開発されようとしている。ISO15489(記録保存管理ガイドライン)に基づく法人の電子記録保存すべきデジタルコンテンツ、顧客情報、知的財産、電子カルテ等個人情報として保護しなければならないデジタルコンテンツの著作権保護(デジタル署名や改ざんの検知)のため、また、動画、静止画、音楽等のマルチメデアに時刻のタグを挿入し、それを用いることによって編集を効率よく行う、または、放送形式により種別した課金情報を抽出するため、さらに、デジタル署名を用いて許可を得たものだけがコンテンツを共有できるようにする等のためにデータハイディング技術が使用されている[1)]。

データハイディングは情報ハイディングとも呼ばれ、コンテンツに何らかの情報を隠す技術のことである。ここではデータハイディングの呼称を用いることにする。データハイディングは透かし(Digital Watermark)技術やステガノグラフィ(Steganography)技術の総称である[2)-4)]。埋め込む情報が重要で、その存在が知られないようにする技術をステガノグラフィと呼び、秘密情報が埋め込まれたコンテンツ自体が重要な場合には、電子透かしと呼ぶ。オリジナルのコンテンツを原コンテンツ、署名のような隠すべきデータを秘密データ、ま

た、秘密データが埋め込まれたコンテンツを流通コンテンツと呼ぶ。透かし情報を埋め込む技術(透かし技術)は既往研究がある[5)-8)]。コンテンツに秘密データを埋め込む方法として、コンテンツの実空間上において秘密データを埋め込む手法[9)]と周波数空間上において秘密データを埋め込む手法[10)、11)]が提案されている。埋め込み後の流通コンテンツに対して圧縮等の処理を施されてもコンテンツ品質の比較的影響を受けにくい特定の周波数帯に秘密データを埋め込むことが可能であるという観点から、前者に比べ後者は秘密データの情報を隠蔽する能力が高い。前者は、また、例えばコンテンツが画像の場合のように、原画像のエッジ部分等を操作して秘密データを埋め込む工夫が必要となる[9)]。後者は、秘密データを埋め込むべき原コンテンツの周波数帯の決定が必要となる[7)、10)]。また、原コンテンツがカラー画像の場合のデータハイディング手法も提案されている[11)、12)]。原コンテンツに埋め込む秘密データは著作権等の情報であり、画像処理や除去攻撃等への耐性、秘匿性の高さがデータハイディングに求められる。画像のような原コンテンツを直交周波数空間に変換後の係数を選択し、その周波数空間に対して秘密データを埋め込む手法はこれら要求を満たしていることから多用され、中でもウェーブレット多重解像度解析[13)、14)]により周波数成分ごとの画像に分割し、分割後の各周波数成分画像のいずれかを秘密データに置き換え、再構築(合成)して流通コンテンツ(画像)を得る方法が既に提案されている。従来手法は、秘匿性が必ずしも十分ではなく(いずれかの周波数成分画像に秘密データが入っているので探し出すことも可能である)、また、秘密データの不可視化(視認困難性)が不十分である。原コンテンツを画像とする場合、原画像の情報量の観点からカラー画像を用いたデータハイディングはその他の方法に比べ秘密データを隠蔽する能力が高い。カラー画像を用いたデータハイディングは、したがって、原コンテンツ(画像)のある成分(赤、青、緑のいずれかの成分)に対して秘密データを埋め込む手法が一般に用いられる。例えば、原コンテンツ(画像)の緑色成分に秘密データを埋め込む手法が採用されている[4)]。したがって、この場合、原画像の赤色成分および青色成分の情報は使用しないことになる。

ウェーブレット多重解像度解析に基づくデータハイディングはウェーブレット周波数成分を調べ、比較的画質に影響が少ない成分に秘密データを埋め込む方法であり、多用されている。しかし、何処の周波数成分に埋め込んでいるかを探ることによって秘密データを捜し当てることもできるため、秘匿性に課題が残されている。この課題克服のためデータハイディングの前処理として主成分変換を原コンテンツ(画像)に施す方法も提案されている[15)-17)]。原コンテンツ(画像)の固有値を知り得るのは原コンテンツ(画像)を保有している著作者のみであることからこれを復元することは著作者のみである。しかし、この方法は秘密データを埋め込んだ流通コンテンツ(画像)を用いてある程度の誤差を許容すれば固有値の近似値を推定できることから秘匿性が不充分である。そのため、さらに、主成分変換後に斜交座標変換を施す前処理も提案されている[18)]。斜交座標の斜交角度は任意に設定でき、この角度に基づき主成分変換ができる。そのため、秘匿性が向上する。また、秘密データの走査方法を変換することにより、流通画像における秘密データの視認困難性および秘匿

性の向上を図った方法も提案されている[19)-22)]。

本小文はこれらウェーブレット多重解像度解析に基づくデータハイディングの方法を概説し、データ圧縮の標準画像として多用されている画像を用いて効果を評価している。

２．多重解像度解析に基づくデータハイディング

2.1 多重解像度解析

双直交ウェーブレット分解(双直交基底関数に基づく離散ウェーブレット変換)は原データ(1次元スカラーデータ：$x=(x_1、x_2、..、x_n))$に対して正方行列Cnを適用して$y=Cn\,x$と定義できる。ここで、Cnは$CnCn^t=I$となる双直交基底関数に基づく変換行列である。このCnの要素は双直交基底関数としてDaubechiesを用いた場合は付録に示すように決定できる[1]。変換後のyは低周波分Lと高周波分Hからなる。すなわち、xは$y=(H_1、L_1)$と変換される。ここで、H、Lの添え字は変換回数、すなわち、段数(レベル数)を表す。また、このL_1にCnを適用することによりH_2、L_2に、さらに、これをn段繰り返すことにより、H_n、L_nと変換される。これを分解という。逆変換は$Cn^{-1}=Cn^t$をyに適用することにより行う。この逆変換をn段繰り返すことにより、$x=(x_1、x_2、..、x_n)$が復元されることになる。これを再構成という。このウェーブレット変換･逆変換(分解･再構成)を繰り返し、ウェーブレット周波数成分に分解し、分解後の成分を使って原データを復元することを多重解像度解析と呼ぶ。これを2次元データ、たとえば、画像に適用すると、$y=(HH_1、HL_1、LH_1、LL_1)$に変換される。ここでHH_1は縦･横次元ともに高周波成分の意味であり、以下同様にLL_1は縦･横次元ともに低周波成分を意味する。これを2次元ウェーブレット変換と呼ぶ。これを繰り返し、原画像データを各周波数成分に分解することができる。また、1次元データの場合と同様に逆変換を繰り返し、原画像データを安全に復元することができる(再構成)。3段までの2次元ウェーブレット変換をFig.1に例示する。データ圧縮方法の評価に多用される標準画像データベース(SIDBA)[2]における[Lenna]に2次元ウェーブレット変換を1段および2段適用して得られる画像をFig.2に示す。このとき、双直交基底関数としてDaubechiesを用いている。この原画像をFig.3に、また、このウェーブレット変換･逆変換に基づく2次元画像データに対する多重解像度解析をFig.4に示す。これをラプラシアンピラミッドと呼ぶ。

---

1 この係数方法は基底関数に依存しており、付録の例示は Daubechies 基底関数の場合のものである。

2 http://vision.Kuee.kyoto-u.ac.jp/IUE/IMAGE_DATABASE/STD_IMAGES/

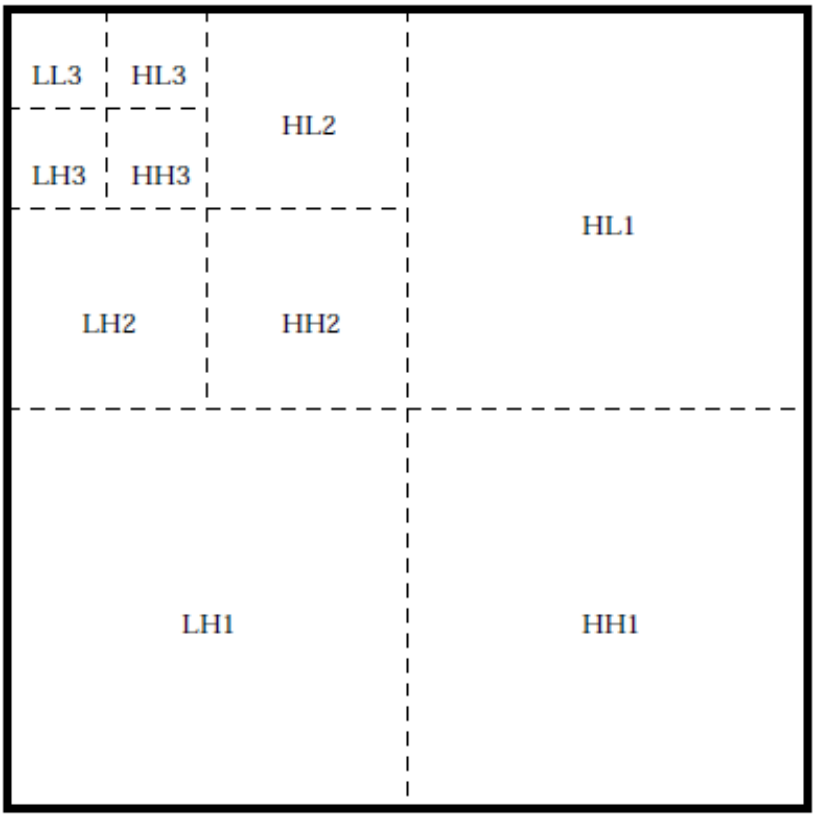
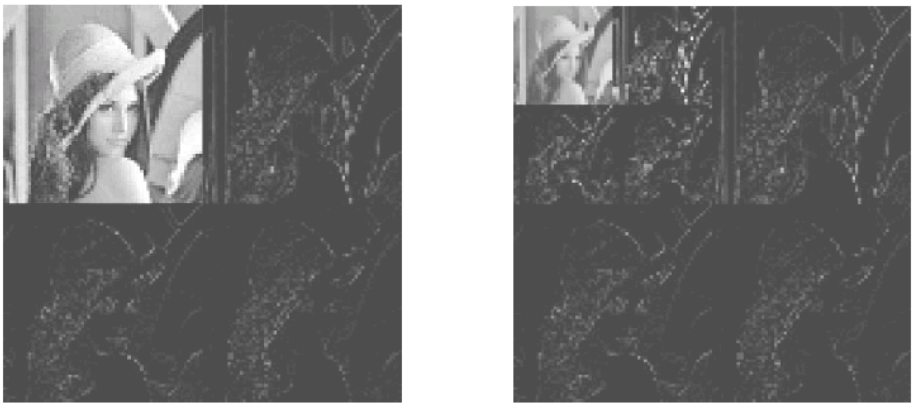


Fig.1 Discrete Wavelet Transformation: DWT (up to the third level)



(a) level 1 (b) level 2

Fig.2 Examples of images of which 1st and 2nd levels of DWT applied to Lena of original image.

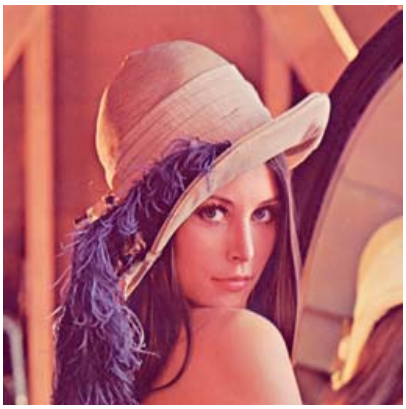


Fig.3 Original image of Lena

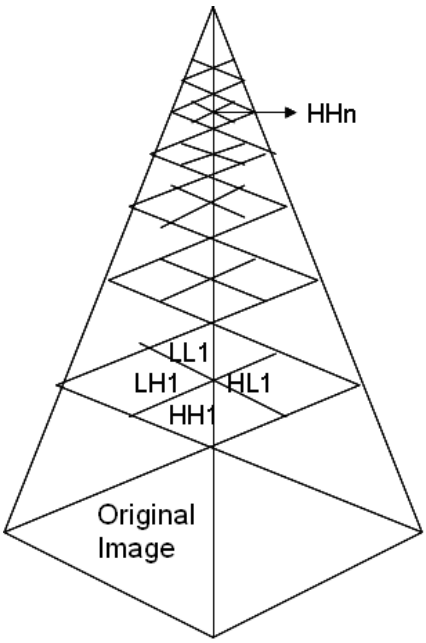


Fig.4 Laprasian Pyramid representing Wavelet Multi-Resolution Analysis (MRA)

2.2 MRA に基づくデータハイディング

原コンテンツ(画像)に埋め込む秘密データは著作権、署名等の情報(画像を含む)であり、画像処理や除去攻撃等への耐性、秘匿性の高さがデータハイディングに求められる。画像を直交周波数空間に変換後の係数を選択し、その周波数空間に対して透かしを埋め込む手法はこれら要求を満たしていることから多用され、中でもここで紹介するウェーブレット多重解像度解析(画像の周波数成分ごとの画像に分割し、分割後の各周波数成分画像のいずれかを秘密データ、または、秘密データ(画像、署名)に置き換え、再構築(合成)して流通データ(画像)を得る方法)が多用されている。Fig.5にデータハイディングの処理の流れを示す。

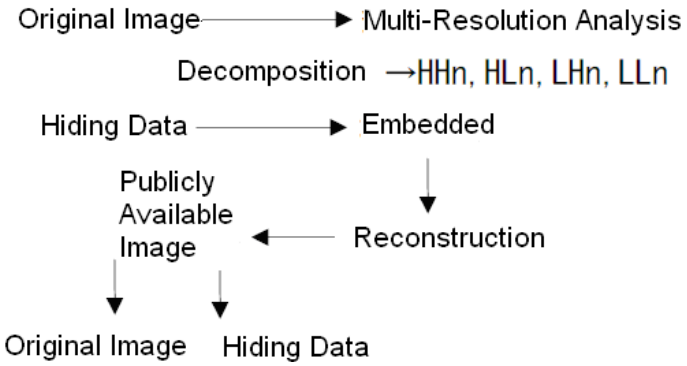


Fig.5 Process flows of data hiding based on MRA

ここで原画像を多重解像度解析によって分解後のいずれの成分画像に秘密データを埋め込むかは流通画像における秘密画像の視認困難性、秘匿性によって決定する。たとえば、Fig.6 に示すように秘密データを周波数成分が少なく、流通画像に影響が少ない HH1、HH2 等に埋め込むことが可能である。ここで秘密データには SIDBA の[Girl:Zelda]を選んだ場合を示した。

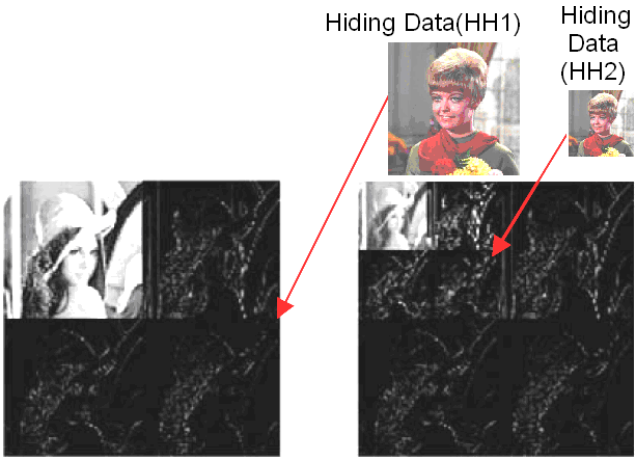


Fig.6 Examples of data hiding of which hiding data are put into the HH1 or HH2 of the image after the DWT.

縦横高周波数成分である HHn に Fig.6 の秘密データ(Zelda 画像)を原コンテンツ(Lena 画像)埋め込んだ場合の流通コンテンツ(画像)を Fig.7 に例示する。ここに示すように埋め込む場所(周波数成分および段数)によって流通画像における秘密画像の視認困難性は依存していることが分かる。明白に Fig.7(d)のように、高段数の高周波数成分に埋め込んだ場合の秘密データの視認性は他に比べ高い。また、いずれの場合も埋め込んだ場所を探れば、秘密データは復元可能であり、単純に多重解像度解析を用いていずれかの周波数成分画像を秘密データに置き換えて流通画像を生成するようなデータハイディングの場合は秘匿性、視認困難性に問題がある。

(a) Level 1

(b) Level 2

(c) Level 3

(d) Level 4

Fig.7 Examples of the distributing images (Hiding data are imbedded into HH1 (Top left)、 HH2 (Top right)、 HH3 (Bottom left) and HH4 (Bottom right).

このとき、原コンテンツ(Lena画像)に多重解像度解析を施し、そのレベルn のHH 成分HHnに秘密データ(Zelda画像)を埋め込み、その後、再構成して流通画像を生成したときの原画像と流通画像との平均二乗偏差(RMS)[3]を調べるとTable 1のようになっており、高段数の高周波数成分に埋め込んだ場合の秘密データの視認困難性は他に比べ高いことが分かる。

Table 1 Root Mean Square difference between the original and the distributing images.

Level	RMS
1	0.0137
2	0.0144
3	0.0153
4	0.0169

2.3 Daubechies 基底関数のサポート長、レベル、秘密画像の置き換えコンポーネント先

Daubechies 基底関数のサポート長、レベル、秘密画像の置き換えコンポーネント先を変え、

[3] 画素数を m、原画像および流通画像の画素値をそれぞれ x_i、、y_i とすると、 $RMS = \frac{1}{m}\sqrt{\sum_{i=1}^{m}(x_i - y_i)^2}$ と定義される。

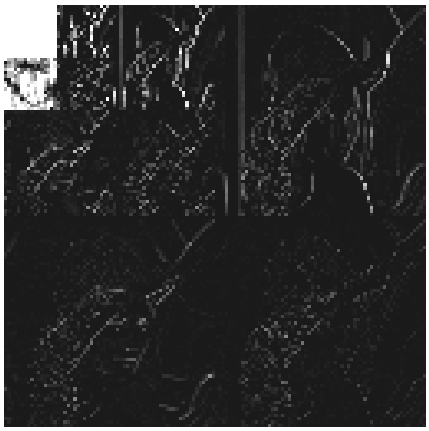
データハイディングにおけるこれらパラメータの影響の度合いを実験的に検証した。Fig.8 に Lena 画像に Mandrill 画像を埋め込む場合を例としてそれらの影響を示す。

(a) MRA image with secret image

(b) Reconstructed image (Level 1)

(c) MRA image with secret image

(d) Reconstructed image (Level 2)



(c) MRA image with secret image

(d) Reconstructed image (Level 3)

Fig.8 Resultant images of wavelet based MRA which contains the secret image of Mandrill (left bottom), LL (left top), LH (right top), and HH (right bottom) as well as reconstructed image (circulation image) which contains embedded secret image

また、Table 2 に Daubechies 基底関数のサポート長、レベル、秘密画像の置き換えコンポーネント先の影響度を原画像と再構成画像との RMS 偏差によって示す。この表から、Daubechies 基底関数のサポート長および秘密画像の置き換えコンポーネント先による RMS 偏差の差はさほどなく、レベルの影響が支配的であることが分かる。

Table 2 Root Mean Square: RMS difference between the original image and reconstructed image containing secret image (Lv_i: Level of MRA, Db_j: Daubechies base function with support length of j).

RMS Difference			
Lv_1	Db_2	Db_4	Db_8
LH	78.25	77.17	77.26
HH	78.27	77.26	77.41
HL	78.91	77.64	77.61
L_{V2}	Db_2	Db_4	Db_8
LH	43.57	42.75	43.68
HH	43.8	41.04	43.93
HL	45.2	43.71	44.44
Lv_3	Db_2	Db_4	Db_8
LH	21.27	21.51	20.78
HH	20.94	20.97	20.86
HL	24.84	25.05	23.77

Fig.9 に Daubechies 基底関数のサポート長、レベル、秘密画像の置き換えコンポーネント先を変えた場合の再構成画像を例示する。

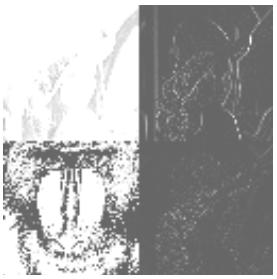
lena_lv1_db2_LH　　lena_lv1_db8_HH　　lena_lv1_db8_LH

Fig.9 Some examples of reconstructed images of MRA based data hiding (lv_i db_j denotes i level and Daubechies base function with support length of j, respectively)

2.4 流通画像を加工した場合の秘密画像の頑健性

秘密画像を高周波コンポーネント成分に置き換えた後に再構成した流通画像を、悪意を持つ利用者が改ざんした場合、埋め込んだ秘密画像は当該処理加工後の流通画像からの抽出可能性について検討した。サポート長が 8 の Daubechies 基底関数に基づくレベル 1 のパラメータセットにおける再構成画像を作成して流通画像とし、それに輝度反転、上下反転、画像拡大(2 倍)を適用し、処理後の画像から秘密画像の抽出を試みた結果を Fig.10 に示す。これらから画像拡大処理に対しては MRA に基づくデータハイディングは比較的に頑健であるが、その他の処理に対しては殆ど秘密画像の抽出は困難であることが分かった。

(a) Reconstructed: LH_Lv1_Db8

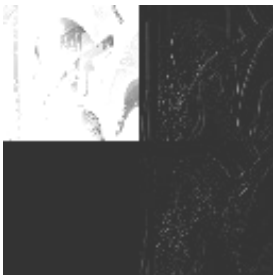


(b) Extracted: without image process

(c)Reconstructed: LH_Lv1_Db8

(d) Extracted: without image process

(e)Reconstructed: LH_Lv1_Db8

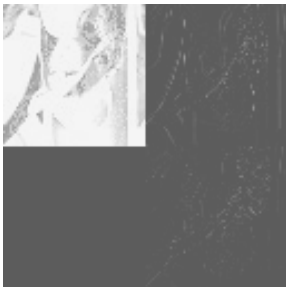


(f) Extracted image: Intensity inversion

(g)Reconstructed: LH_Lv3_Db8

(h) Extracted image: Intensity inversion

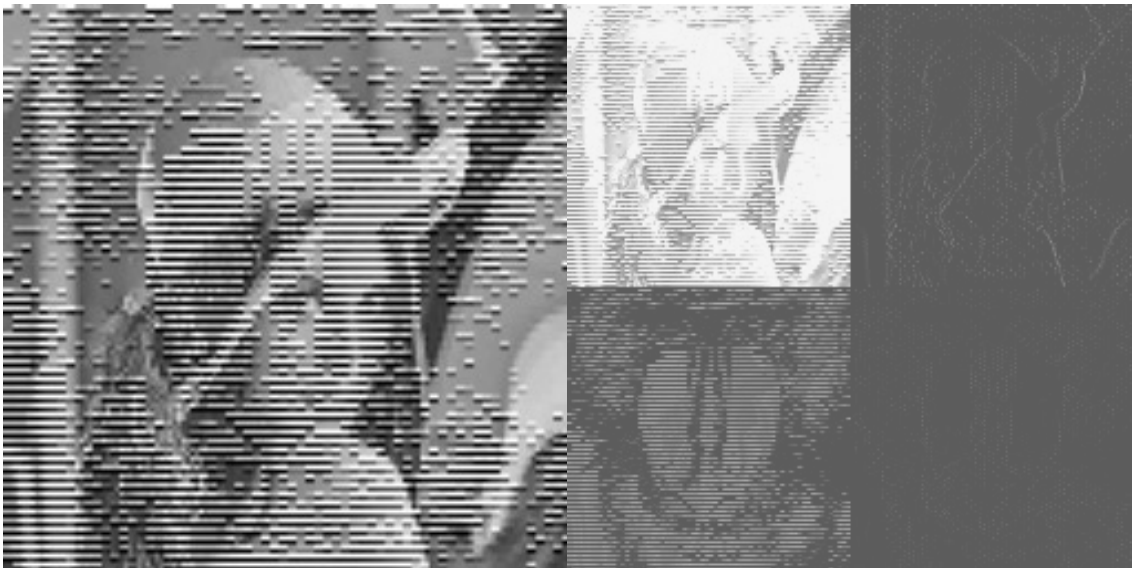
(i) Reconstructed: LH_Lv1_Db8



(j) Extracted image: up-side -down conversion

(k)Reconstructed: LH_Lv3_Db8

(l) Extracted image: up-side -down conversion



(m)Reconstructed: LH_Lv1_Db8

(n) Extracted image: Magnification

(o)Reconstructed: LH_Lv3_Db8

(p) Extracted image: Magnification

Fig.10 Robustness of the proposed data hiding against image processing, intensity inversion, up-side-down of geometric conversion, magnification with the different level of MRA of which the secret image is replaced to it.

2.5 データ圧縮に対する耐性

流通画像にデータ圧縮を適用した場合、圧縮法が既知であれば解凍後に秘密画像を抽出する可能性がある。ここでは代表的なデータ圧縮法としてブロック符号化で情報の損失がある JPEG およびウェーブレット圧縮であり、情報損失のない JPEG2000 を例にとり、秘密

画像の抽出を試みた。データ圧縮率を 100 に設定し、JPEG および JPEG2000 を Lena 画像に適用した。Fig.11 に圧縮後の画像を示す。JPEG 圧縮後の画像にはブロック符号化特有のブロックひずみが顕著である。また、モスキートノイズ、色ひずみが散見される。その点、JPEG2000 はそれらのひずみがなく、圧縮によるデフェクトが目立たない。

(a) JPEG (b) JPEG 2000

Fig.11 Enlarged portions of images of decompressed images of JPEG and JPEG 2000 with data compression ratio of 100.

これを裏付けるように PSNR(Peak Signal-to-Noise Ratio)を調べてみると圧縮率、ビットレートにも依存するが、すべての場合において JPEG2000 が JPEG の PSNR よりも優れていることが分かる。

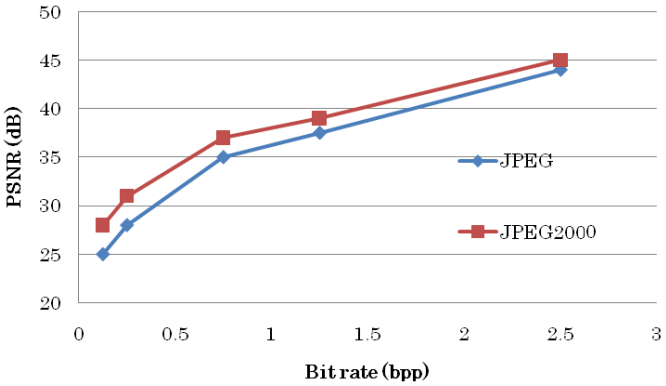


Fig.12 PSNR (Peak Signal-to-Noise Ratio) of JPEG and JPEG 2000 as a function of bit rate.

Mandrill の秘密画像の画素順序を Merssene Twister による一様乱数によって入れ替える前処理をデータハイディングに適用し、秘密画像の視認困難度を向上させた。これらに JPEG および JPEG2000 のデータ圧縮を適用した場合の画像例を Fig.13 に示す。流通画像に JPEG 圧縮を適用した画像に対して解凍を施しても情報の損失が存在するため、秘密画像は完全に復元することはできず、復元画像と原画像との RMSE を評価すると、6.38 であった。一方、JPEG2000 は情報の損失がないため、圧縮画像から完全に流通画像が解凍できるため、当該

流通画像に含まれる秘密画像が完全に抽出できる。このとき、当然ながら、RMSE=0 であった。

(a) JPEG

(b) JPEG 2000

Fig.13 Circulation images with embedded secret image of Mandrill with the data compression by JPEG and JPEG 2000 with the same permutation by the same random number which is generated by Mersenne Twister

2.5 固有値展開による秘匿性･視認困難性の向上

MRAに基づくデータハイディングには秘匿性、視認困難性が不十分である。この課題を克服するため、MRAに基づくデータハイディングの前処理として主成分変換を施す方法が提案されている。原画像の固有値･固有ベクトルを知り得るのは原画像を保有している著作者のみであることからこれを復元することは著作者のみである。そのため、著作者は当該原コンテンツの著作権を主張できる。しかし、この方法は秘密データを埋め込んだ流通画像を用いてある程度の誤差を許容すれば固有値の近似値を推定できることから秘匿性が不充分である。この課題克服のために、さらに、主成分変換後に斜交座標変換を施す前処理も提案されている。斜交座標の斜交角度は任意に設定でき、この角度に基づき主成分変換ができる。この角度情報を原コンテンツの送受当事者間のみが知り得る条件の下に当該コンテンツの秘匿性は高く守られる。また、当該角度情報を共通鍵方式等による暗号化によって原コンテンツのLSB: Least Significant Bit(量子化における最下位ビット)に挿入することにより、秘匿性はさらに高く設定することが可能になる。ここでは主成分変換および斜交座標変換を伴う多重解像度解析に基づくデータハイディングを紹介する。

Fig.14にこの方法の処理の流れを示す。まず、主成分変換によって原画像のエネルギーを集中させ、変換後の主成分画像の直交座標を斜交座標に変換してさらにエネルギー集中度を高め、これにMRAを適用して分解後のいずれかのレベル、周波数成分に秘密データを埋め込んだ後に再構成を行うことにより流通画像を得る。当該コンテンツの送受当事者間は斜交座標変換パラメータおよび原画像の固有値、固有ベクトルからなる主成分変換パラメータが既知であるので原画像および秘密データを復元できるがこれらを知り得ない第3者は復元が困難である。秘密データを挿入する成分を変更することにより、多重解像度解析

に基づくデータハイディングが秘密データの情報を保護する能力を高めることができる。

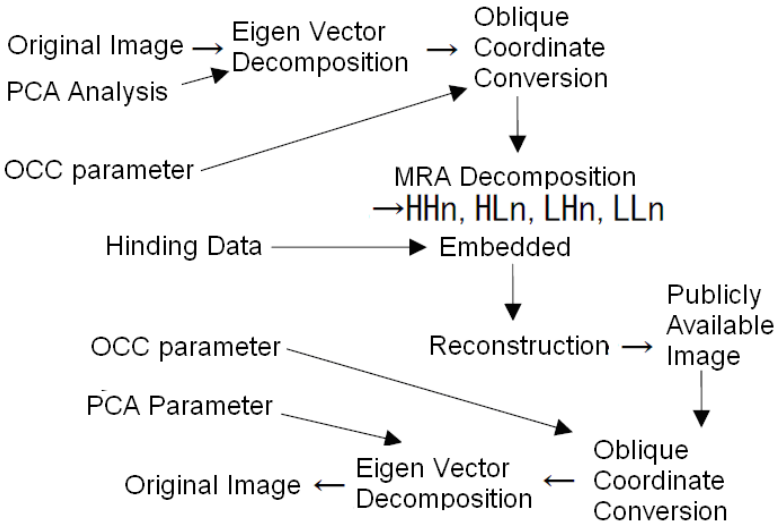


Fig.14 Process flow of the data hiding based on MRA using eigen vector decomposition (or Principal Component Analysis: PCA) and coordination conversion from Cartesian to oblique coordination systems.

二次元平面における、直交座標と斜交座標表現とは、

$$W = X + Y\cos(\theta) \quad (1)$$

$$Z = Y\sin(\theta) \quad (2)$$

という関係がある。ここでWZは斜交座標における各軸であり、XYは直交座標における各座標軸である。また、θは斜交座標変換における座標軸の角度である。この一例をFig.15に示す。

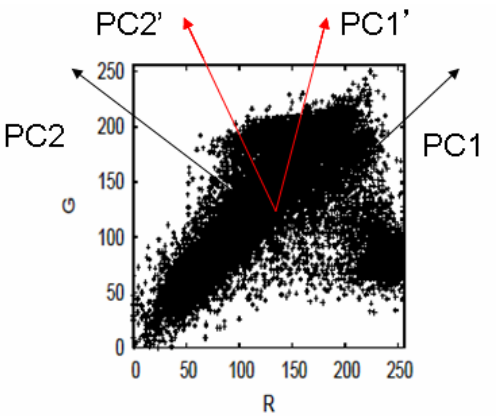


Fig.15 PCA and Oblique-coordinate conversion (OCC) as a preprocessing of the MRA based data hiding.

同図において、原画像の赤色および緑色の二次元画素分布(散布図)は第1主成分軸PC1およびそれに直交する第2主成分軸PC2から構成される二次元座標に主成分変換される。また、このPC1およびPC2からなる直交座標はPC1'およびPC2'からなる斜交角度θの斜交座標に変換する。この時、量子化ステップを変えずに斜交角度を変化させると画素の定義域が変化する。極端な例は定義域が量子化ステップを下回るような場合である。この場合、量子化雑音のみとなり、何ら情報が伝わらないことになる。これを回避するため、あらかじめ、画素定義域を拡張し(例えば、斜交角度を90 度±45 度の範囲で変化させる場合には画素値を2倍にする)、原画像および秘密データを復元する際に縮小すればよい。これは可逆な処理であり、この拡大･縮小率についてもコンテンツ保有者のみは知っているので完全に復元できる。

次に、秘密データの復号法を説明する。秘密データがハイディングされる前の多次元原画像に主成分変換を行った際の係数を用いて、流通用画像に対して第１主成分画像を構成し、その第１主成分画像に対してウェーブレット分解を行うことにより実現される。提案手法による秘密データの復号は、秘密データ情報をハイディングする前の多次元原画像に主成分変換を行った際の係数を知っている場合のみ復号可能である。すなわち、秘密データをハイディングする前の多次元原画像により、主成分変換の係数は異なる。

実験例を次に示す。原画像としてこれもデータ圧縮評価用標準画像データベース(SIDBA)の中から選んだMandrill(Fig.16)を用いた。また、秘密データとしてFig.17に示す時系列データ(グラフ)を用いた。

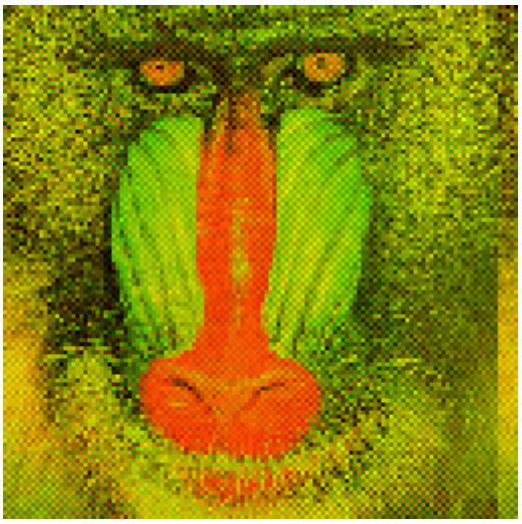


Fig.16 Original image of Mandrill of SIDBA database for data compression algorithm evaluation。

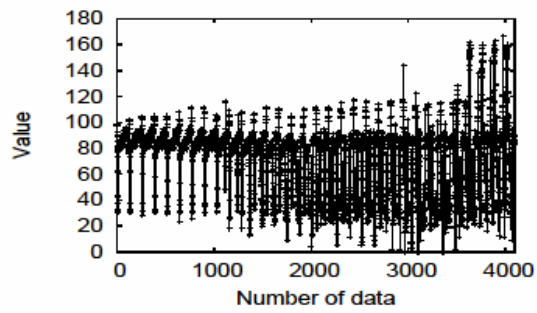


Fig.17 Hiding data of image (time series of data graph).

このカラー原画像の赤･緑･青の3 原色画像をFig.18に示す。ここで青成分は0 として便宜上、二次元多重分光画像データとしている。カラー原画像、ならびに、この主成分変換後の散布図をFig.19に示す。この時、カラー原画像の赤色、緑色の二次元散布から主成分変換を行うための平均ベクトルと変換係数行列は、すなわち、固有値および固有ベクトルはそれぞれ、

$$\lambda=(133.772、\ 129.297) \tag{3}$$

$$V=\begin{vmatrix} 0.835 & 0.550 \\ -0.550 & 0.835 \end{vmatrix} \tag{4}$$

である。

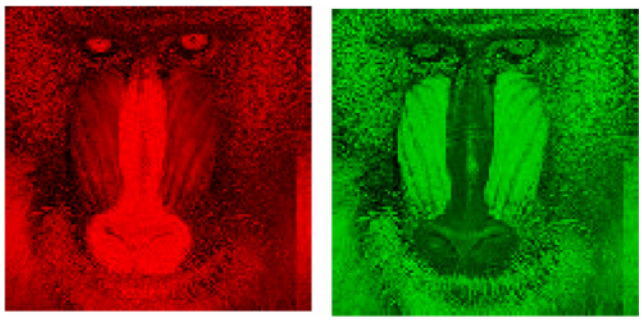


Fig.18 Red and Green colored images of the original colored image of Mandrill (Fig.13).

これら正確な固有値および固有ベクトル(主成分変換パラメータ)は原画像を保有している著作者のみが知っており、流通画像に基づき固有値展開を行っても正確な固有値･固有ベクトルを用いた場合の復元画像との差を生じることになり、正確な秘密データは得られない。しかし、秘密データが冗長性の高い画像データであるような場合、許容誤差範囲内の秘密データの復元は可能であり、十分な秘匿性とは言い切れない。より秘匿性を高めるため、斜交座標変換を前処理として追加することにした。

前出の式(1)、(2)を用いて主成分変換後の直交座標軸を任意の斜交座標軸に変換する。この時、斜交座標軸の斜交角度を斜交座標変換パラメータθとする。θ=90度は主成分変換後の直交座標軸そのものであり、例えば、θ=110 度および70 度によって斜交座標変換を施すとFig.19はFig.20になる。また、斜交座標変換後の画像をFig.21に示す。この斜交角度は前述のとおり量子化ステップの範囲まで任意に設定可能であるが、この角度に依存して画素値の定義域の拡大・縮小率を変え、それに伴い量子化ビットを加減する必要があり、処理リソースを大きくする必要に迫られるため、ここでは±20 度程度とした。第三者はFig.18から秘密データを推定することになるが斜交角度、固有値ベクトル等の情報がないため、復元は困難を極める。

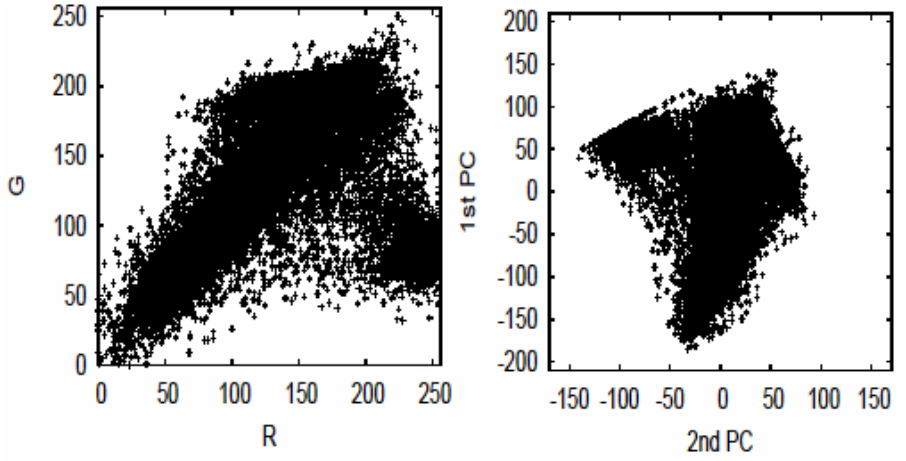
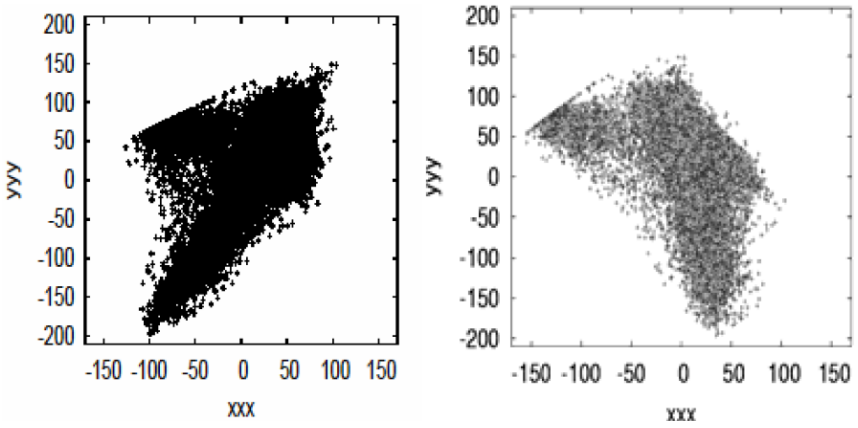


Fig.19 Scatter diagrams of the original red and green coordinate system and that of the first and second components coordinate system.



(a) 110 degree (b) 70 degree

Fig.20 Oblique coordinate converted scatter diagrams with 110 and 70 degrees.

パラメータθにより秘密データの保護性能が向上することを示す。そのため、Fig.13の第1主成分画像からウェーブレット変換を用いて秘密データを推定することを試みる。なお、第三者は、原画像の情報を知らないものとし、ウェーブレット基底および秘密データを埋め

込んだ成分(例えば、HH1成分)の情報をなんらかの方法で知り得たものとする。すなわち、当事者が保有する固有ベクトル等の情報とパラメータθを第三者は未知とする。

Fig.21 Widely available image of which hiding image (time series of data graph) is hidden in the original image (Mandrill).

斜交座標変換パラメータθを変化させて秘密データを埋め込み、それぞれの流通用データに対して第三者が秘密データを流通用データから推定することを試みた場合のRMS 偏差をFig.22に示す。

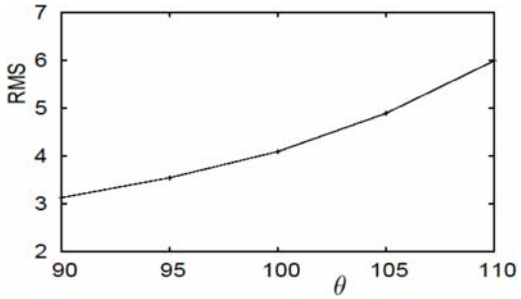


Fig.22 Root Mean Square difference between original and distributing image of the proposed watermarking with 90 to 110 degrees of OCC.

第三者が秘密データを流通用データから推定することを試みた場合のRMS 偏差は、θに依存している。すなわち、θにより秘密データの秘匿性が向上することがわかる。また、これより、原画像の情報を保護することによって秘密データの保護が可能となることが分かる。

次に秘密データの復元度を調べる。流通画像から秘密データを復元し、その復元度(秘密データと復元秘密データとの平均二乗誤差(RMS error)を評価した。その結果をFig.20 に示す。この図から、原秘密データと流通画像から復元した秘密データとの平均二乗誤差が大きくなることが分かる。すなわち、斜交角度を大きくすると秘匿性が向上することが分かる。雑音に対する耐性もデータハイディングに重要である。流通画像に画像処理、除去攻撃等による雑音が混入した場合の秘密データの復元度を評価した。混入する雑音は平均が0

で標準偏差sigma が5-20 まで5ステップで変えた正規乱数を流通画像に重畳した。その後、秘密データを復元し、原秘密データとの平均二乗誤差を評価した。その結果Fig.24に示す。重畳した雑音の分だけ平均二乗誤差は確実に増えている。また、斜交座標角度を90 度(直交座標)から110 度まで変化させることにより、約10%の平均二乗誤差の増大がみられた。したがって、雑音に対する耐性はなく、重畳する雑音の分だけ復元秘密データは劣化し、斜交座標変換の効果は10%程度であることが分かった。斜交座標角度をさらに大きくすることは可能であるが、座標変換後の再量子化による量子化誤差が増大するため、70 度から110 度程度が限界であると判断した。

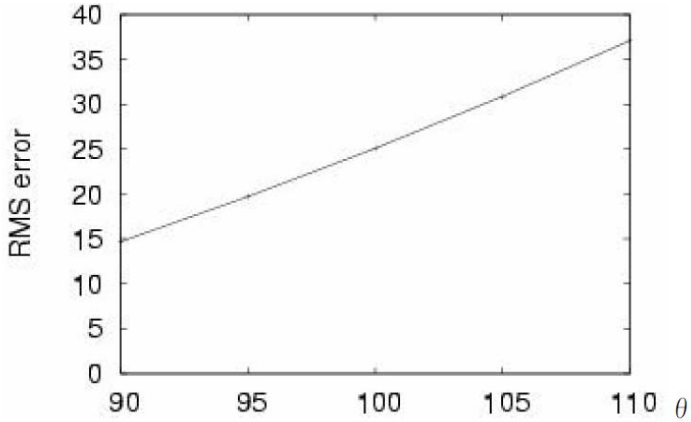


Fig.23 RMS error between original and restored hiding data derived from the distributing image.

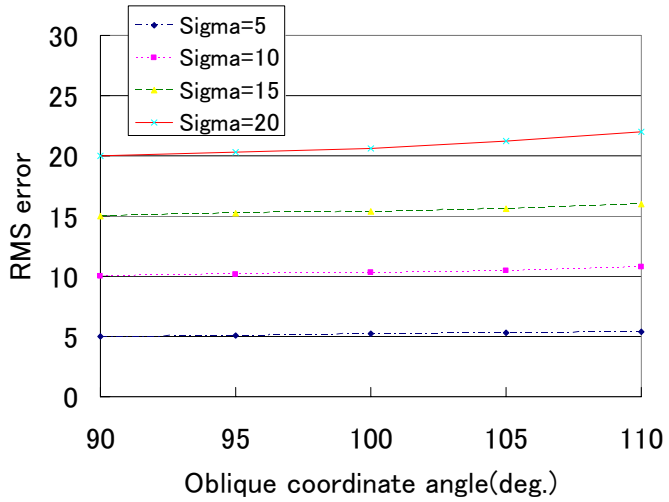


Fig.24 Hiding performance for which normal distributed noise with zero mean and standard deviation ranges from 5 to 20 (5 step) is added to the distributing image.

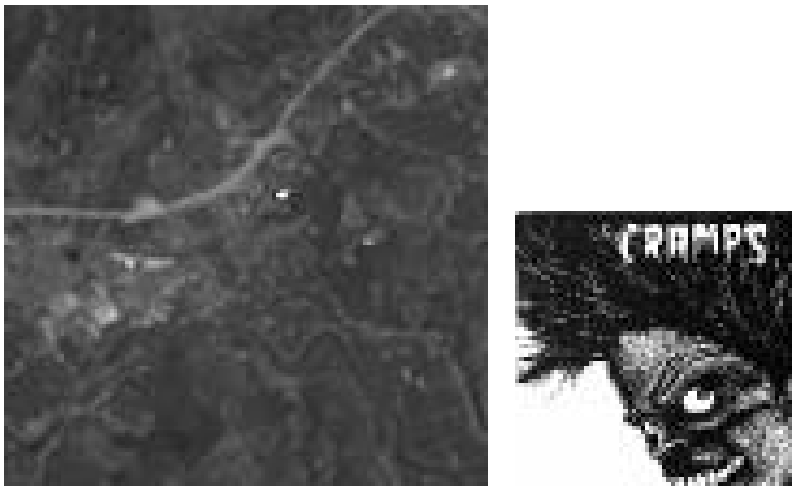
2.6 走査方式を用いる秘匿性･視認困難性の向上

秘密データを、埋め込む前に何らかの走査方式により走査し直すことにより、流通画像における秘密データの視認困難性を向上する方法も提案されている[20)]。この方法では送受当事者間において走査のための規則を共有していることを前提としている。通常の画像デー

タが線逐次走査であるのに対し、秘密画像を例えば乱数を発生させて走査順を決定するようなランダム走査に変換する。すなわち、2次元空間データを辞書式配列(1次元データ)に変換・保存するものである。

この走査方式の変換は順列変換行列によって行うことができる。このとき、乱数発生規則情報を送受当事者間で共有していれば、ランダム走査における秘密データを抽出した後に順列変換行列の逆行列を適用することにより走査方式の逆変換が可能になり線逐次走査における秘密データを再生することができる。秘密データを埋め込む際の走査方式の情報を第３者が入手することは困難であるので秘密データの秘匿性も向上する。

実験結果を次に示す。使用データはFig.25(a)に示すものを原画像(Landsat衛星に搭載されたTMセンサ[4]が佐賀市郊外の高速道路、長崎自動車道の大和インターチェンジ付近を観測したバンド3(赤の波長域)のデータ)として用い、Fig.22(b)に示す秘密データに用いた。原画像は128x128画素、秘密データは64x64画素で構成されている。量子化ビットはいずれも8である。



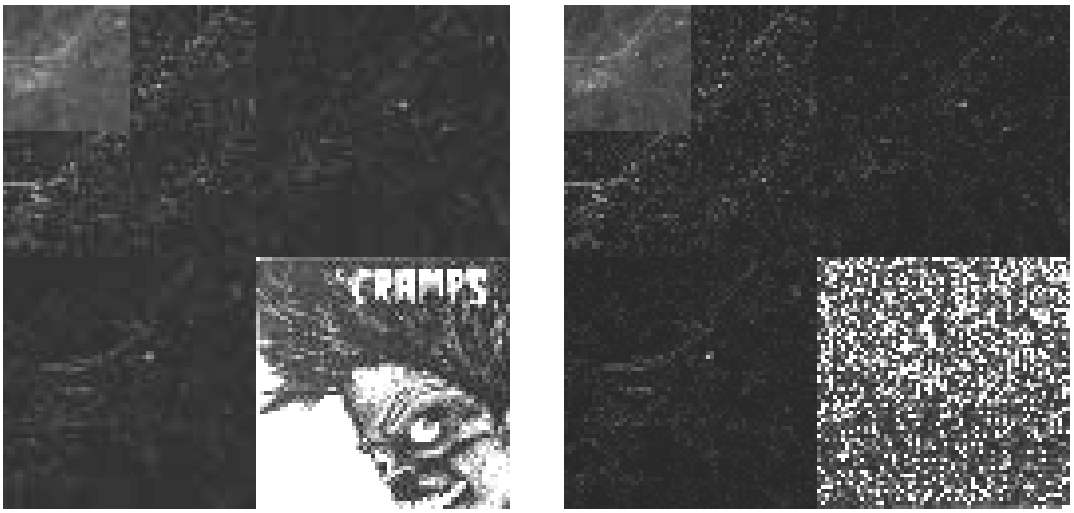
(a) Original (b) Hiding data

Fig.25 Original image of Landsat-5/TM Band 3 data of Saga city and hiding data

原画像にウェーブレット分割を1段適用し、HH1に秘密データを埋め込んだ。その際、秘密データを線逐次走査のままHH1に埋め込む方法とMersenneTwisterの乱数生成法に基づき[1-4096]の範囲の一様乱数を生成し、それに基づき走査し直してからHH1に埋め込む方法と比較した。埋め込んだ画像を用いて再構成して得られる流通画像は原画像と殆ど変わらない。しかし、当該流通画像を用いてHH1の再構成を試みてみると、線逐次走査およびランダム走査の場合、それぞれ、Fig.26(a)、(b)に示すように秘密データが復元できる。当然ながら、線逐次走査のままに秘密データを原画像に埋め込んだ場合は、HH1に秘密データそのものが

[4] Landsat/TM は青から近赤外波長域に 5 バンドおよび熱赤外波長域に 1 バンドを分光波長帯とする 30m 空間分解能のマルチスペクトルスキャナーである。

復元できてしまうが、ランダム走査の場合は乱数生成に係る支配パラメータをしらなければ秘密データを復元することはできない。



(a) Line by line scanning (b) Random scanning

Fig.26 Reconstructed hiding data from publicly available image content derived from the MRA based methods with the different scanning schemes.

３．あとがき

ウェーブレット多重解像度解析に基づくデータハイディングの前処理として主成分変換および斜交座標変換を施すことにより、秘匿性の向上を図った方法を紹介した。この手法を多重分光画像に適用しその効果を調べた。第三者が流通用データのみから秘密データを抽出することを試みた場合の秘匿性を検討した。

本論文にて紹介した手法は、原画像となる多重分光画像の特性を知る著作者、または、当事者のみが秘密データを復元できるものである。すなわち、原画像の情報を保護する必要がある場合に適用可能である。また、紹介した手法においては、固有ベクトルの存在および斜交座標変換により、秘密データの情報を保護する。すなわち、少なくとも真の原画像の情報を知らなければ、秘密データを復元できない。主成分変換の係数は、原画像毎に異なり、原画像の固有ベクトルにより構成される。さらに、秘密データの走査方式を線逐次からランダム走査に変換することにより秘密データの秘匿性、流通画像における視認困難性の向上に繋がる方法も併せて示した。このとき、ランダム走査に用いる乱数の支配方程式のパラメータを送受当事者間のみが共有することによってより秘匿性の高いデータハイディングが実現できる。

本論文では、ウェーブレットとして Daubechies基底関数を採用したが、双直交ウェーブレットであれば秘密データを復元できる。双直交ウェーブレットとして何を採用するかを隠蔽することによっても秘密データを保護することができる。

謝辞：筆者は実験にご協力戴いた(株)サインズ瀬戸　要博士に感謝します。

附録：Duabechies 基底関数に基づく多重解像度解析(係数の求め方を含む)

直交ウェーブレット[13)、14)]は、双直交ウェーブレットの１種である。すなわち、双直交ウェーブレット変換は逆変換可能である。たとえば、式(附1) のDaubechies 基底関数[13)、14)]に基づく離散ウェーブレット変換(式(附1)により直交ウェーブレットCn とベクトル表現したn 個の画素からなる画像$[\eta_1、\ \eta_2、\ldots、\eta_n]^T$ との積を求める)によりウェーブレット分解および再構成を行い、多重解像度解析とする。

$$C_n\begin{bmatrix}\eta_1\\ \eta_2\\ \vdots\\ \eta_n\end{bmatrix} \qquad \text{(附1)}$$

ここで式(附1)のCn は、例えば、サポート長(13)が2であり、画素数が8 個のベクトル表現した画像にこれを適用する場合、

$$C_8^{[2]}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\\ \eta_6\\ \eta_7\\ \eta_8\end{bmatrix}=\begin{bmatrix}p_0 & p_1 & & & & & & \\ q_0 & q_1 & & & & & & \\ & & p_0 & p_1 & & & & \\ & & q_0 & q_1 & & & & \\ & & & & p_0 & p_1 & & \\ & & & & q_0 & q_1 & & \\ & & & & & & p_0 & p_1\\ & & & & & & q_0 & q_1\end{bmatrix}\begin{bmatrix}\eta_1\\ \eta_2\\ \eta_3\\ \eta_4\\ \eta_5\\ \eta_6\\ \eta_7\\ \eta_8\end{bmatrix}$$

$$=\begin{bmatrix}p_0x_1+p_1x_2\\ q_0\eta_1+q_1\eta_2\\ p_0\eta_3+p_1\eta_4\\ q_0\eta_3+q_1\eta_4\\ p_0\eta_5+p_1\eta_6\\ q_0\eta_5+q_1\eta_6\\ p_0\eta_7+p_1\eta_8\\ q_0\eta_7+q_1\eta_8\end{bmatrix} \qquad \text{(附2)}$$

$$\left(C_n^{[2]}\right)^T C_n^{[2]}=I_n$$

$$p_0+p_1=\sqrt{2}$$

$$q_0=p_1$$

$$q_1=-p_0$$

$$0^0q_0+1^0q_1=0 \qquad \text{(附3)}$$

の条件式を満足するようなp_0、p_1、q_0、q_1を求めることにより、Cnの係数が求められる。すなわち、サポート長が2 の場合のDaubechies 基底関数に基づくウェーブレット分解(変換)、再構成(逆変換)が可能になる。画素数が大きな画像であってもサポート長が変わってもDaubechies 基底関数に基づく直交ウェーブレット変換に基づく多重解像度解析が構築でき

る[14)]。本論文における実験においてサポート長が2 のDaubechies 基底関数に基づく多重解像度解析を適用した。

参考文献

1) 小野 束、電子透かしとコンテンツ保護、オーム社、(2001).
2) 松井甲子雄、電子透かしの基礎、森北出版、(2000).
3) 川合淳郎、透かし-本物の証・紙のすき入れ（透かし）画像電子学会誌、 31、 2、 253-260、 (2002).
4) 河口英二・野田秀樹・新見道治、ディジタル・ステガノグラフィ技術について、画像電子学会誌、 31、 3、 414-420、 (2002).
5) Tanaka、 K.、 Y. Nakamura and K. Matsui: "Embedding Secret Information into a Dithered Multilevel Image" Proceedings of the 1990 IEEE Military Communications Conference、216-220、(1990).
6) Tirkel、 A.、 et al.、 "Electronic Water Mark"、 Proceedings DICTA 1993、 666-672、 (1993).
7) Cox、 I.J.、 J. Killian、 T. Leighton、 T. Shamoon: "Secure spread spectrum watermarking for multimedia"、 NEC Technical Report 95-10、 NEC Research Institute (1995)
8) Bender、 W. D. Gruhl、 and N. Morimoto : "Techniques for Data Hiding" Proceedings of the SPIE 2420、 Storage and Retrieval for Image and video Database III、 164-173、 (1995).
9) 岡本晃宏・宮崎明雄、モルフォロジカル信号処理を利用した電子透かし方式、電子情報通信学会論文誌 A、 J84-A、 8、 1037-1044、 (2001).
10) 井上尚・宮崎明雄・島津幹夫・桂卓史、ウェーブレット変換を用いた画像信号に関する電子透かし方式、映像情報メディア学会誌、 52、 12、 1832-1839、 (1998).
11) Ouellette、 R.、 H. Noda、 M. Niimi、 and E. Kawaguchi、 Topological Ordered Color Table for BPCS Steganography Using Indexed Color Images、 IPSJ Journal、 42、 1、 110-113、 (2001).
12) 瀬戸英俊・森多雄司・姜錫・青木由直、画像圧縮におけるフレネル変換を用いたカラー画像電子透かし法、信学技報、 43-48 (2000).
13) 新井康平、ウェーブレット解析の基礎理論、森北出版、(2000).
14) 新井康平、独習ウェーブレット解析、近代科学社、(2006).
15) 新井康平、瀬戸要、ウェーブレット多重解像度解析に基づくデータハイディング、可視化情報学会誌、22、 Suppl.1、 229-232、 （2002）.
16) 新井康平、瀬戸要、 主成分変換による情報の偏りを利用した多重解像度解析に基づくデータハイディング、 可視化情報学会論文誌、23、 8、 72-79、(2003).
17) 新井 康平・瀬戸 要、主成分分析に基づく情報ハイディング手法、日本リモートセンシング学会 第37 回学術講演会論文集、 (2004).
18) 新井康平、瀬戸要、座標変換に基づく情報ハイディング手法、可視化情報学会誌、25、

Suppl.1、 55-58、(2005).

19) 新井 康平・瀬戸 要、走査方式を考慮した双直交ウェーブレット多重解像度解析に基づくデータハイディング、日本リモートセンシング学会 第35 回 学術講演会論文集、(2003).

20) 新井 康平・瀬戸 要、走査方式を考慮した双直交ウェーブレット多重解像度解析に基づく鍵画像の視認困難性、秘匿性を向上したデータハイディング、可視化情報学会論文誌、29、Suppl.1、167-170、(2009).

21) Arai K., and Y. Yamada, Data hiding method which robust to run-length data compression based on lifting dyadic wavelet transformation, Proceedings of the 11th Asian Symposium on Visualization, NIIGATA, JAPAN, 2011.

22) Arai K., and Y. Yamada, Improvement of Secret Image Invisibility in Circulation Image with Dyadic Wavelet Based Data Hiding with Run-Length Coded Secret Images of Which Location of Codes are Determined with Random Number(IJACSA) International Journal of Advanced Computer Science and Applications, Vol.2, No.7, to appear, 2011.

新井康平　840-8502　佐賀市本庄 1 番地　佐賀大学工学研究科知能情報システム学専攻

1974 年日本大学大学院理工学研究科修士課程修了、1985 年 6 月工学博士、1974 年~78 年東京大学生産技術研究所、1979 年~1990 年宇宙開発事業団(現 JAXA)、1990 年から佐賀大学教授、現在に至る。1998 年からアリゾナ大学客員教授、2008 年から国際学術連合・宇宙研究委員会コミッション A 副議長、著書：独習ウェーブレット解析(近代科学社)等 29 編。

連続マルチウェーブレット変換に基づく画像分離

守本 晃 *
* 大阪教育大学　情報科学

概要． 本講演では，いくつかの画像が混合された複数の観測画像から，元画像を分離する問題を考える．一般に，自然な画像のエッジは区分的に連続な曲線であり，いくつかの元画像のエッジの交わりは点であることが期待できるので，混合画像のエッジを元画像のエッジに分離することが容易にできる．この結果から，未知の混合行列を推定し元画像を分離する．そこで，本講演では画像のエッジ抽出を行う変換として，連続マルチウェーブレット変換を提案し，混合画像の分離問題に応用する際の解決すべき問題を提起する．

Image separation based on continuous multiwavelet transform

Akira Morimoto*
*Osaka Kyoiku University

Abstract. This lecture is concerned with an image separation problem to separate original images from mixed observed images. Natural images have pixels where there is a sharp contrast in intensity. The set of such pixels is called edge. Discontinuities in natural images are generated by edges and most edges are composed of piecewise continuous curves. In other words, edges of natural images are essentially one dimensional objects. By this reason, the edges of mixed images can be easily separated into the edges of original images. This process can estimate an unknown mixing matrix and separate mixed images. A continuous multiwavelet transform is proposed for edge extraction. Several issues to be solved are raised when the continuous multiwavelet transform is applied to image separation problem.

1. はじめに

本講演は，大阪電気通信大学・工学部・基礎理工学科の萬代 武史さん，大阪教育大学・数理科学の芦野 隆一さん，大阪教育大学・情報科学・修士課程２年の片岡 秀輔さんとの共同研究をもとに，連続マルチウェーブレット変換の基本性質を概観し，混合画像の分離問題に応用する際の解決すべき問題を提起する．

2. ブラインド信号源分離

パーティ会場のような，話し声や音楽や食器の音などの様々な音声信号が複雑に重なり合った環境下でも，我々は特定の音声信号のみを選択的に聞き分けることができる．この聴覚系の能力は，実験心理学者の Cherry [8] によって**カクテルパーティ問題**と名付けら

れ，心理学から工学まで広い分野で研究されている [11] .

工学的には，複数のセンサーで捉えた観測信号から信号源の個数・混合モデルのパラメータを推定し，最終的に元信号を推定する問題であり，**ブラインド信号源分離問題** [9] と呼ばれている.

ブラインド信号源分離問題に対しては，**独立成分分析**を用いた解法が一般的であるが，**時間周波数解析**を用いた解法もある．2 個の観測信号の短時間フーリエ変換の商をブラインド信号源分離問題に利用するというアイデアは，Jourjine-Rickard-Yilmaz [12] による．彼らは，DUET（Degenerate Unmixing Estimation Technique）という分離方法を提案し，一つの信号源のみが活動している時間周波数領域が互いに素 “disjoint orthogonality” という条件下で，二つの観測信号があれば信号源が分離できることを示した．我々 [1,6,7] は，ウェーブレット変換による時間周波数情報を採用し，disjoint orthogonality の条件を緩め，一つの信号源の時間周波数情報のみが活動している時間周波数領域が広いという “時間周波数独立性” の仮定の下で，観測信号の時間周波数情報の商を用いて一つの信号源のみが活動している時間周波数領域を取り出して，モデルパラメータの推定を行ってきた.

本講演では，画像分離問題に焦点を当て，一つの元画像のみが活動している所を連続マルチウェーブレット変換を用いて抽出することを試みる.

3. 画像分離問題

簡単のために，$Q \times R$ 画素からなるグレースケール画像を扱う．画像は，実数値の $Q \times R$ 行列である．画像の混合を次のように行う.

M 個の $Q \times R$ のサイズの元画像を

$$s_m = (s_m[q,r]) \in \mathbb{R}^{Q\times R}, \qquad m = 1, \ldots, M$$

とする．元画像 s_m から J 個の観測画像

$$y_j = (y_j[q,r]) \in \mathbb{R}^{Q\times R}, \qquad j = 1, \ldots, J$$

が数理モデル

$$y_j = \sum_{m=1}^{M} a_{j,m}\, s_m, \qquad j = 1, \ldots, J \tag{3.1}$$

にしたがって，与えられる場合を考えよう．ただし，$J \geq M$ を仮定する．さらに，**混合係数** $a_{j,m} > 0$ も仮定する．$A = (a_{j,m}) \in \mathbb{R}^{J\times M}$ を**混合行列**と呼ぶ．Fig. 1 の左側が 4 個の元画像 $s_m, m = 1, \ldots, 4$ から得られた観測画像の例である.

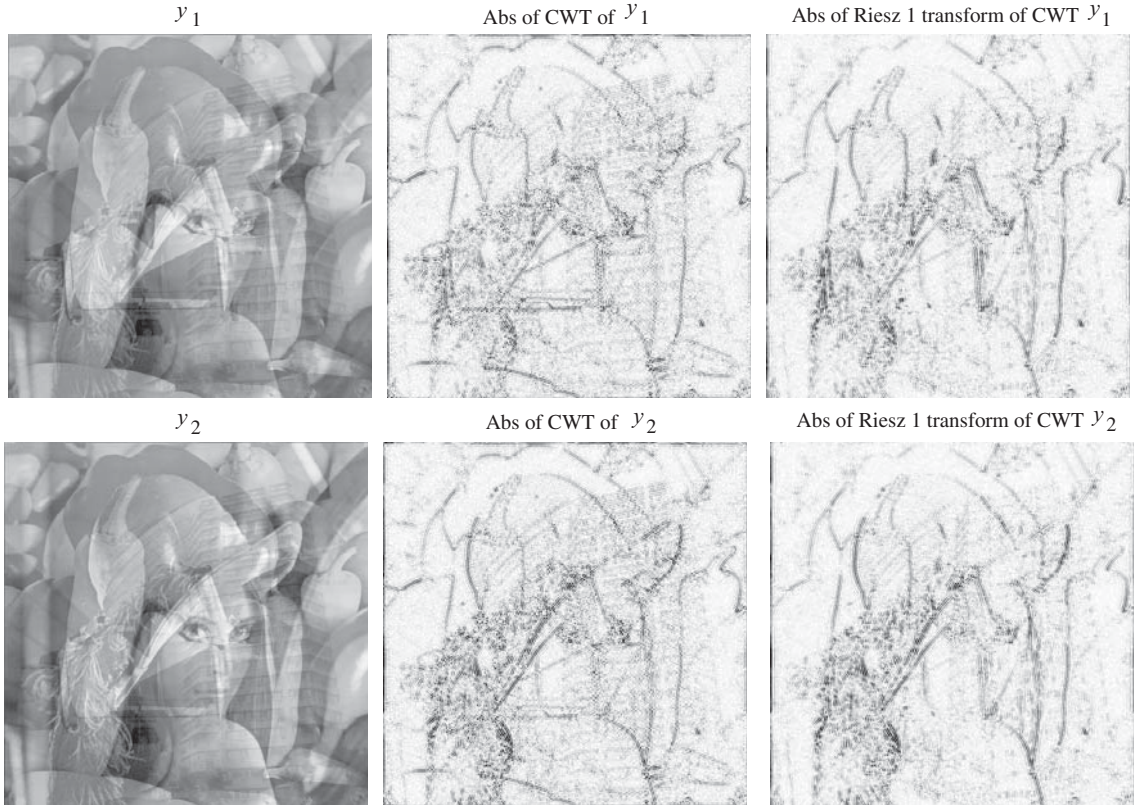


Fig. 1. 観測画像 y_1, y_2（左）とその変換画像（中）と変換画像のリース変換 $\mathcal{R}_1$ （右）.

4. 画像分離のアイデア：変換した観測画像の商

観測画像 y_j と元画像 s_m に $\mathbb{R}^{Q\times R}$ 上の線形作用素 B を作用させた変換画像を

$$Y_j^B = By_j, \qquad S_m^B = Bs_m$$

で表す．数理モデル (3.1) の線形性より，両辺に線形作用素 B を作用させると，

$$Y_j^B[q,r] = \sum_{m=1}^{M} a_{j,m} S_m^B[q,r], \quad 1 \le j \le J$$

となり，変換画像は同じモデルに従う．

ここで，変換させた観測画像の $Y_j^B[q,r]$ と $Y_1^B[q,r]$ の商

$$Q_j^B(q,r) := \frac{Y_j^B[q,r]}{Y_1^B[q,r]} = \frac{\sum_{m=1}^{M} a_{j,m} S_m^B[q,r]}{\sum_{m=1}^{M} a_{1,m} S_m^B[q,r]}$$

を取り，線形作用素 B で，k 番目の元画像のみが活動している位置の集合

$$E_k^B = \left\{(q,r) \mid S_k^B[q,r] \neq 0,\ S_m^B[q,r] = 0,\ m \neq k\right\}$$

を考えると，$(q,r) \in E_k^B$ の時，商は，

$$Q_j^B(q,r) = \frac{\sum_{m=1}^{M} a_{j,m} S_m^B[q,r]}{\sum_{m=1}^{M} a_{1,m} S_m^B[q,r]} = \frac{a_{j,k} S_k^B[q,r]}{a_{1,k} S_k^B[q,r]} = \frac{a_{j,k}}{a_{1,k}}$$

となって，推定したいモデルパラメータである混合係数の比と一致する．

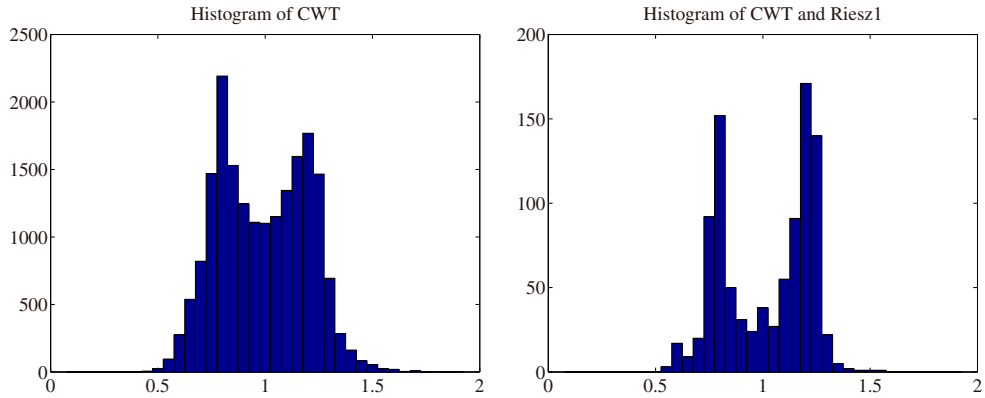


Fig. 2.　商のヒストグラム $H_2^B(t)$（左）と 2 つの商が等しいときのヒストグラム $H_2^{B,B_1}(t)$（右）.

したがって，k 番目の元画像のみが活動している位置の集合 E_k^B の要素が多ければ，商 $Q_j^B(q,r)$ のヒストグラム

$$H_j^B(t) = \sharp\left\{(q,r) \mid Q_j^B(q,r) = t\right\}$$

は，$t = a_{j,k}/a_{1,k}$（混合係数の比）でピークを持つ．ピークの個数から元画像の数 M が推定でき，ピークの位置から混合係数の比が推定できる．

ところで，観測画像にどのような線形作用素 B を作用させると，集合 E_k^B の要素が増えるのであろうか？

一般に，自然画像の輪郭線は 1 次元の曲線であるから，元画像の輪郭線が交わるのは点でしかない．したがって，混合画像からエッジを取り出すような線形作用素を作用させれば，一つの元画像のみが活動している位置が取り出せると思われる．Fig. 1 左の元画像からエッジを取り出すと中図を得る．Fig. 1 中図は，線形作用素 B として，文献 [15] で扱った連続ウェーブレット変換を用いた．Fig. 2 左が商 $Q_2^B(q,r)$ のヒストグラム $H_2^B(t)$ であり，2 つのピークがある．

このヒストグラム $H_2^B(t)$ ではきれいなピークが得られないので，同じ解像度のエッジを少し違ったように見るもう一つの線形作用素 B_1 で変換した混合画像 $Y_j^{B_1} = B_1 Y_j$ を用意する．Fig. 1 右図は，中図のリース変換 $\mathcal{R}_1$ を取った図であり，縦方向のエッジをより強調している．2 種類の線形作用素 B と B_1 での商の値が一致する点 (q,r) のヒストグラム

$$H_2^{B,B_1}(t) = \sharp\left\{(q,r) \mid Q_2^B(q,r) = Q_2^{B_1}(q,r) = t\right\}$$

を描くと Fig. 2 右図になり，左図よりだいぶ改善される．一応，元画像の個数 $M = 4$ 個のピークが見える．

一方，文献 [17] では，双直交ウェーブレット `rbio3.3` を用いた離散定常ウェーブレット変換 [16] のレベル 3 までの縦・横・斜めエッジに相当する詳細係数を取り出す 9 種類の変換を線形作用素 B とした．もう一つの線形作用素 B_1 は双直交ウェーブレット `rbio2.2` を用いた同じ解像度・方向の離散定常ウェーブレット変換である．これら 9 種類のエッジを取り出した図で商を作り，商のヒストグラムの頻度を加算したヒストグラムのピークの個数から元画像の個数を推定し，ピークの位置から混合係数の比を推定した．また，文献 [14] では，双直交ウェーブレット `rbio3.3` を用いた離散定常ウェーブレット変換とそのリース変換を，線形作用素 B と B_1 に用いた．

これらの方法を比較した結果，一つの解像度に対して Fig. 1 中図のような全方向のエッジを抽出する連続ウェーブレット変換を用いる方法 [15] より，一つの解像度に対して縦・横・斜めの 3 方向のエッジを抽出する離散定常ウェーブレット変換を用いる方法 [14, 17] の方がきれいなヒストグラムが得られ，画像分解能力に優れていることがわかった．そこで，複数種類のウェーブレット関数を用いた連続ウェーブレット変換として，連続マルチウェーブレット変換を提案する [4].

5. 連続マルチウェーブレット変換

空間 n 次元の連続ウェーブレット変換とフーリエ乗算作用素を解説してから，連続マルチウェーブレット変換を提案する．空間次元 $n = 2$ の画像分離問題に用いた周波数空間でコンパクトサポートを持つマルチウェーブレット関数を設計し，パラメータを離散化したパーセヴァルフレームとの関係を述べる．

5.1 連続ウェーブレット変換

$L^2(\mathbb{R}^n)$ における基本的な 3 つのユニタリ作用素，つまり，平行移動 $\mathcal{T}_b$, ダイレーション $\mathcal{D}_a$, モジュレーション $\mathcal{M}_\omega$ を

$$\mathcal{T}_b f(x) := f(x-b), \quad \mathcal{D}_a f(x) := a^{-n/2} f(x/a), \quad \mathcal{M}_\omega f(x) := e^{i\omega x} f(x)$$

とする．ただし，パラメータの範囲は，$a \in \mathbb{R}_+ := \{a \in \mathbb{R} \mid a > 0\}$, $b, \omega \in \mathbb{R}^n$ である．$f \in L^1(\mathbb{R}^n)$ のフーリエ変換と $g \in L^1(\mathbb{R}^n)$ の逆フーリエ変換を

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) := \int_{\mathbb{R}^n} f(x)\, e^{-i\xi x}\, dx,$$
$$\mathcal{F}^{-1}(g)(x) = \check{g}(x) := (2\pi)^{-n} \int_{\mathbb{R}^n} g(\xi)\, e^{i\xi x}\, d\xi$$

で定義する．$f, g \in L^2(\mathbb{R}^n)$ の内積とノルムを

$$\langle f, g\rangle := \int_{\mathbb{R}^n} f(x)\, \overline{g(x)}\, dx, \qquad \|f\| := \sqrt{\langle f, f\rangle}$$

とする．ここで複素数 z の複素共役を $\bar{z}$ でかく．

ウェーブレット関数 $\psi \in L^2(\mathbb{R}^n)$ に関する $f \in L^2(\mathbb{R}^n)$ の**連続ウェーブレット変換**(CWT) は

$$(W_\psi f)(b,a) := \langle f, \mathcal{T}_b \mathcal{D}_a \psi \rangle = \int_{\mathbb{R}^n} f(x)\, a^{-n/2}\, \overline{\psi\left(\frac{x-b}{a}\right)}\, dx$$
$$= (2\pi)^{-n} \left\langle \widehat{f}, \mathcal{M}_{-b} \mathcal{D}_{1/a} \widehat{\psi} \right\rangle = (2\pi)^{-n} \int_{\mathbb{R}^n} \widehat{f}(\xi)\, a^{n/2}\, \overline{\widehat{\psi}(a\xi)}\, e^{i\xi b}\, d\xi$$

で定義される．**許容条件**，すなわち，ある定数 C_ψ があって，

$$\int_{\mathbb{R}_+} \frac{|\widehat{\psi}(a\xi)|^2}{a}\, da = C_\psi, \quad \text{a.e.}\, \xi \in \mathbb{R}^n \tag{5.1}$$

という条件が満たされるとき，連続ウェーブレット変換 $W_\psi f$ から f が再構成できる．

$$f = (C_\psi)^{-1} \int_{\mathbb{R}^n \times \mathbb{R}_+} (W_\psi f)(b,a)\, T_b \mathcal{D}_a \psi\, \frac{dbda}{a^{n+1}}.$$

5.2 フーリエ乗算作用素 (FMO)

$\mu \in L^\infty(\mathbb{R}^n)$ に対するフーリエ乗算作用素 (FMO) $\mu(D)$ を

$$(\mu(D)f)^\wedge(\xi) = \mu(\xi)\widehat{f}(\xi)$$

で定義する．この記号は擬微分作用素の記号を流用したものであり，関数 $\mu(\xi)$ を作用素 $\mu(D)$ の**シンボル**または**マスク**と呼ぶ．$\mu(D)$ は $L^2(\mathbb{R}^n)$ の有界線形作用素であり，$\mu(D)$ の共役作用素は $\mu(D)^* = \overline{\mu}(D)$ である．ただし，$\overline{\mu}(\xi) := \overline{\mu(\xi)}$. $\mu(-\xi) = \overline{\mu(\xi)}$ なら，$\mu(D)$ は実数値関数を実数値関数に写す．マスク $\mu(\xi)$ が，

$$\mu(a\xi) = \mu(\xi), \qquad \forall\, a \in \mathbb{R}_+,\ \text{a.e.}\ \xi \tag{5.2}$$

を満たすとき，マスクは**次数 0 の正斉次**であるという．

例 1 (連続ウェーブレット変換) 連続ウェーブレット変換は，周波数空間での積分で

$$(W_\psi f)(b,a) = (2\pi)^{-n} \int_{\mathbb{R}^n} \widehat{f}(\xi)\, a^{n/2}\, \overline{\widehat{\psi}(a\xi)}\, e^{i\xi b}\, d\xi$$

と記述できた．ここで，a を固定して，パラメータ b に関してフーリエ変換を取ると

$$\mathcal{F}_b\, (W_\psi f) = a^{n/2}\, \overline{\widehat{\psi}(a\xi)}\, \widehat{f}(\xi)$$

となるので，マスク $\mu(\xi) = a^{n/2}\, \overline{\widehat{\psi}(a\xi)}$ と取ったフーリエ乗算作用素になる．

許容条件を満たす連続ウェーブレット変換のマスクは，$\xi \in \mathbb{R}^n$ の方向によらず積分 (5.1) が一定値 C_ψ を取る必要があるので，$\mathbb{R}^2$ のマスクとしては，Fig. 3 左図のような同心円状のパターン以外には設計しにくい．

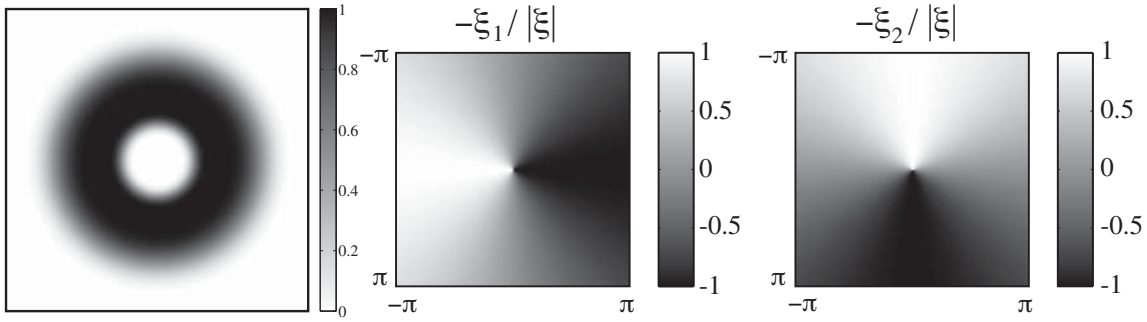


Fig. 3. $\mathbb{R}^2$ でのマスクの例．左図：許容条件を満たす連続ウェーブレット変換のマスクの例 $\mu(\xi) = a^{n/2}\,\overline{\widehat{\psi}(a\xi)}$．リース変換 $\mathcal{R}_\ell$ のマスクの虚部，中：$\Im\mu_1(\xi)$，右：$\Im\mu_2(\xi)$.

例 2 (リース変換) 周波数座標 $\xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$ に対して，ℓ 番目のマスクを

$$\mu_\ell(\xi) = -i\frac{\xi_\ell}{|\xi|}, \qquad \ell = 1, \ldots, n$$

としたフーリエ乗算作用素を $\mathcal{R}_\ell$ とかきリース変換とよぶ．ただし，$|\xi| = \sqrt{\sum_{\ell=1}^n |\xi_\ell|^2}$ は，ξ の大きさである．リース変換のマスクは次数 0 の正斉次である．実数値関数 f のリース変換 $\mathcal{R}_\ell f$ は実数値関数になる．マスク $\mu_\ell(\xi)$ は虚数なので，$\mathbb{R}^2$ の場合のマスクの虚部を描くと Fig. 3 の中図と右図になる．リース変換は以下の性質を満たす．

$$\mathcal{T}_b\mathcal{R}_\ell = \mathcal{R}_\ell\mathcal{T}_b, \quad \mathcal{D}_a\mathcal{R}_\ell = \mathcal{R}_\ell\mathcal{D}_a, \quad \sum_{\ell=1}^n \mathcal{R}_\ell^2 = -I,$$

ここで，I は恒等作用素である．

5.3 連続マルチウェーブレット変換

マルチウェーブレット関数 $\Psi = (\psi^p)_{p=1}^P \in L^2(\mathbb{R}^n)^P$ に関する $f \in L^2(\mathbb{R}^n)$ の**連続マルチウェーブレット変換**(CMWT) を，

$$(W_\Psi f)(b, a) := \Big((W_{\psi^p} f)(b, a)\Big)_{p=1}^P, \qquad a \in \mathbb{R}_+,\ b \in \mathbb{R}^n$$

で定義する．ここで，Ψ や $W_\Psi f$ は P 次の縦ベクトルである．ベクトル値関数に対するフーリエ乗算作用素の作用は，

$$\mu(D)(\psi^p)_{p=1}^P := (\mu(D)\psi^p)_{p=1}^P$$

と定める．連続マルチウェーブレット変換は，次数 0 の正斉次なマスクを持つフーリエ乗算作用素と大変相性がよい．

命題 1　次数 0 の正斉次なマスク $\mu \in L^\infty(\mathbb{R}^n)$ と，マルチウェーブレット関数 $\Psi \in L^2(\mathbb{R}^n)^P$ に対して

$$\mu(D)W_\Psi f = W_\Psi \mu(D) f = W_{\overline{\mu}(D)\Psi} f$$

が成立する．ここで，$\mu(D)W_\Psi f$ において，$\mu(D)$ は，各 $a \in \mathbb{R}_+$ 毎に，b の関数 $W_\Psi f(b,a)$ に作用する．すなわち，$\mu(D)((W_\Psi f)(\cdot, a))$ である．

5.4　内積関係式と再構成公式

Gröchenig [10, Chapt.10] を参考にして，マルチウェーブレット関数 Ψ_1, Ψ_2 を考えて，内積関係式と再構成公式を与える．

$\Psi_1, \Psi_2 \in L^2(\mathbb{R}^n)^P$ とし，ペア (Ψ_1, Ψ_2) は次の条件を満たすと仮定する．

$$(A_{\Psi_1,\Psi_2}) \qquad \begin{array}{l} \xi \text{ によらない定数 } M \text{ があって,} \\ \displaystyle\int_{\mathbb{R}_+} \left|\widehat{\Psi_1}(a\xi)^* \widehat{\Psi_2}(a\xi)\right| \frac{da}{a} \le M, \quad \text{a.e. } \xi \in \mathbb{R}^n \setminus \{0\}. \end{array}$$

このとき

$$C_{\Psi_1,\Psi_2}(\xi) := \int_{\mathbb{R}_+} \widehat{\Psi_1}(a\xi)^* \widehat{\Psi_2}(a\xi) \frac{da}{a} \in L^\infty(\mathbb{R}^n) \tag{5.3}$$

なる関数が定まる．ここで，$F^* := \overline{F^{\mathrm{T}}}$ は列ベクトル $F \in \mathbb{C}^P$ の複素共役転置であり，G^*F は $F, G \in \mathbb{C}^P$ の普通の意味の内積である．$C_{\Psi_1,\Psi_2}(\xi)$ は次数 0 の正斉次関数（式 (5.2) を満たす）であることは容易に分かる．$\Psi_1 = \Psi_2 = \Psi$ のときは，上の条件と関数をそれぞれ (A_Ψ), $C_\Psi(\xi)$ とかく．もし，Ψ_1, Ψ_2 がそれぞれ (A_{Ψ_1}), (A_{Ψ_2}) を満たすなら，ペア (Ψ_1, Ψ_2) は (A_{Ψ_1,Ψ_2}) を満たす．

マルチウェーブレット関数 $\Psi \in L^2(\mathbb{R}^n)^P$ が次の**許容条件**を満たすとき，**アナライジングマルチウェーブレット**と呼ばれる．

(5.4)　　Ψ は (A_Ψ) を満たし，$C_\Psi(\xi)$ は ξ によらない非零定数である．

次の定理が連続マルチウェーブレット変換の基本定理である．

定理 2　(i) $\Psi_1, \Psi_2 \in L^2(\mathbb{R}^n)^P$ が (A_{Ψ_1,Ψ_2}) を満たすなら，$f, g \in L^2(\mathbb{R}^n)$ に対して，

$$\langle C_{\Psi_1,\Psi_2}(D) f, g\rangle = \int_{\mathbb{R}_+} \Big(\int_{\mathbb{R}^n} (W_{\Psi_2} g)(b,a)^* (W_{\Psi_1} f)(b,a)\, db \Big) \frac{da}{a^{n+1}} \tag{5.5}$$

が成立する．従って，

$$C_{\Psi_1,\Psi_2}(D) f = \int_{\mathbb{R}^n \times \mathbb{R}_+} (\mathcal{T}_b \mathcal{D}_a \Psi_2)^{\mathrm{T}} (W_{\Psi_1} f)(b,a) \frac{db\, da}{a^{n+1}}. \tag{5.6}$$

(ii) $\Psi \in L^2(\mathbb{R}^n)^P$ が (A_Ψ) を満たすなら，W_Ψ は $L^2(\mathbb{R}^n)$ から $L^2(\mathbb{R}^n \times \mathbb{R}_+, dbda/a^{n+1})^P$ （の中）への有界線形作用素である．

(iii) $\Psi_1, \Psi_2 \in L^2(\mathbb{R}^n)^P$ が (A_{Ψ_1,Ψ_2}) をみたし，C_{Ψ_1,Ψ_2} が ξ によらない非零定数とすると，$f \in L^2(\mathbb{R}^n)$ は $(W_{\Psi_1} f)(b,a)$ から

$$f = (C_{\Psi_1,\Psi_2})^{-1} \int_{\mathbb{R}^n \times \mathbb{R}_+} (\mathcal{T}_b \mathcal{D}_a \Psi_2)^{\mathsf{T}} (W_{\Psi_1} f)(b,a) \frac{db\, da}{a^{n+1}} \tag{5.7}$$

と再構成される．特に，Ψ がアナライジングマルチウェーブレットなら

$$f = (C_\Psi)^{-1} \int_{\mathbb{R}^n \times \mathbb{R}_+} (\mathcal{T}_b \mathcal{D}_a \Psi)^{\mathsf{T}} (W_\Psi f)(b,a) \frac{db\, da}{a^{n+1}}. \tag{5.8}$$

補助定理 3　アナライジングマルチウェーブレット $\Psi \in L^2(\mathbb{R}^n)^P$ に対して，
(i) マルチウェーブレット関数 $(\mathcal{R}_1\Psi; \ldots; \mathcal{R}_n\Psi)$ および $(\Psi; \mathcal{R}_1\Psi; \ldots; \mathcal{R}_n\Psi)$ も許容条件を満たす．ただし，以下のセミコロン記号を使う．縦ベクトル $F_1, \ldots, F_m$ に対して，

$$(F_1; \ldots; F_m) = (F_1^{\mathsf{T}}, \ldots, F_m^{\mathsf{T}})^{\mathsf{T}}.$$

(ii) 全ての p に対して $\operatorname{supp} \widehat{\psi^p} \cap \operatorname{supp} \widehat{\psi^p}(-\cdot) = \emptyset$ を満たせば，次の 3 種類のマルチウェーブレット関数 $\mathfrak{R}\Psi$, $\mathfrak{I}\Psi$, および $(\mathfrak{R}\Psi; \mathfrak{I}\Psi)$ は許容条件を満たす．

5.5　マルチウェーブレット関数のデザイン

ここでは，周波数空間でコンパクトサポートを持つようなアナライジングマルチウェーブレット関数をデザインしよう．許容条件を満たすウェーブレット関数を周波数空間で切り刻んでマルチウェーブレット関数を構成する．

命題 4　許容条件 (5.1) を満たすウェーブレット関数 $\psi_0 \in L^2(\mathbb{R}^n)$ と

$$\sum_{p=1}^{P} |\chi_p(\xi)|^2 = 1, \qquad \text{a.e. } \xi \in \mathbb{R}^n \setminus \{0\}$$

を満たす周波数空間の分割関数 $\chi_p(\xi)$, $p = 1, \ldots, P$ に対して，ψ^p を $\widehat{\psi^p} = \chi_p \widehat{\psi_0}$ で定義する．このとき，$\Psi := (\psi^p)_{p=1}^P$ は，定数 $C_\Psi = C_{\psi_0}$ を満たすアナライジングマルチウェーブレット関数になる．

以下では，画像分離を考えるため $n = 2$ 次元の場合に限定して，マルチウェーブレット関数をデザインする．

5.5.1 円環マルチウェーブレット関数

$\psi_\alpha(\xi)$ が周波数空間で Fig. 3 左のマスクを持つアナライジングウェーブレット関数とする．つまり，定数 $\alpha > 1$ に対して，周波数空間の 1 の分割

$$\sum_{j\in\mathbb{Z}} \left|\widehat{\psi_\alpha}(\alpha^j\xi)\right|^2 = 1, \qquad \text{a.e. } \xi \in \mathbb{R}^2 \setminus \{0\}$$

を満たすような滑らかな関数として $\widehat{\psi_\alpha}(\xi)$ を設計する．具体的には，$\alpha = 2$ の場合には，メイエの正規直交ウェーブレット関数 $\psi_{\mathrm{Meyer}}(x)$ を用いて，$\widehat{\psi_\alpha}(\xi) = \left|\widehat{\psi_{\mathrm{Meyer}}}(|\xi|)\right|$ とする．

リース変換のマスクを用いて，周波数空間の分割を行う．

$$\chi_1(\xi) = \frac{1}{\sqrt{2}}, \quad \chi_2(\xi) = \frac{-i\xi_1}{\sqrt{2}|\xi|}, \quad \chi_3(\xi) = \frac{-i\xi_2}{\sqrt{2}|\xi|}.$$

命題 4 を用いると，

$$\Psi_\alpha^{\mathrm{A}} = \left(\psi_\alpha^{\mathrm{A},p}\right)_{p=1}^3 := (\chi_1(D)\psi_\alpha, \chi_2(D)\psi_\alpha, \chi_3(D)\psi_\alpha)^T = \frac{1}{\sqrt{2}}(\psi_\alpha, \mathcal{R}_1\psi_\alpha.\mathcal{R}_2\psi_\alpha)^T$$

は，アナライジングマルチウェーブレット関数になる．また周波数空間の 1 の分解

$$\sum_{j\in\mathbb{Z}}\sum_{p=1}^{3} \left|\widehat{\psi_\alpha^{\mathrm{A},p}}(\alpha^j\xi)\right|^2 = 1, \qquad \text{a.e. } \xi \tag{5.9}$$

を満たす．

これらの関数を用いて画像分離を行ったのが文献 [15] である．マルチウェーブレット関数 Ψ_α^{A} の外形を Fig. 4 に描く．左：$\mathcal{F}(\chi_1(D)\psi_\alpha)$，中：$\chi_1(D)\psi_\alpha$，右：$\chi_2(D)\psi_\alpha$ を濃淡図で描いた．上から $\alpha = 1.25, 1.5, 2.5$ である．

5.5.2 円環分割マルチウェーブレット関数

Fig. 3 左のマスクを持つアナライジングウェーブレット関数 $\psi_\alpha(\xi)$ を周波数空間で原点からの偏角方向に P 等分すと，Fig. 5 左図で灰色に塗ったようなマスクができる．このようにして作ったアナライジングマルチウェーブレット関数を

$$\Psi_{\alpha,P}^{\mathrm{AS}} = \left(\psi_{\alpha,P}^{\mathrm{AS},p}\right)_{p=1}^P$$

とする．Fig. 6 にウェーブレット関数 $\psi_{\alpha,P}^{\mathrm{AS},p}$ のマスク（左），実部（中），虚部（右）の例をあげた．上から，$(\alpha, P, p) = (1.5, 12, 1), (2, 20, 4), (2.5, 28, 2)$ である．このマルチウェーブレット関数のマスクも周波数空間の 1 の分解

$$\sum_{j\in\mathbb{Z}}\sum_{p=1}^{P} \left|\widehat{\psi_{\alpha,P}^{\mathrm{AS},p}}(\alpha^j\xi)\right|^2 = 1, \qquad \text{a.e. } \xi \tag{5.10}$$

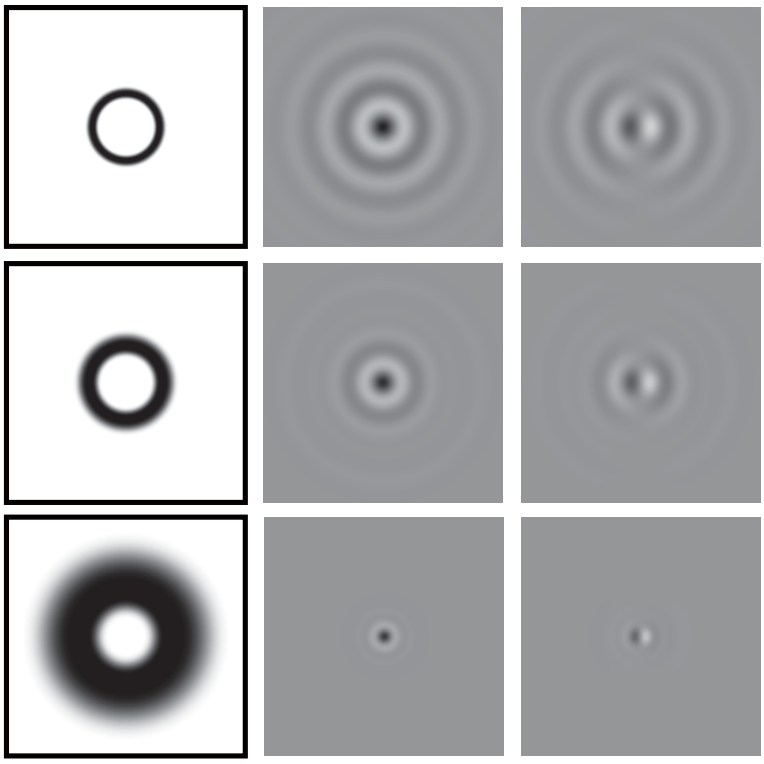


Fig. 4. $\mathcal{F}(\chi_1(D)\psi_\alpha), \chi_1(D)\psi_\alpha, \chi_2(D)\psi_\alpha$ の密度分布．$\alpha = 1.25, 1.5, 2.5$.

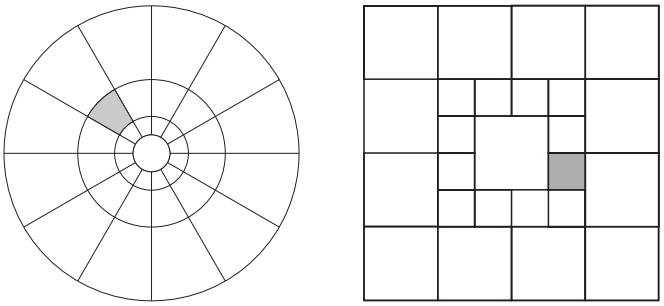


Fig. 5. 左：円環分割マルチウェーブレット関数のマスク．右：四角タイリングによるマルチウェーブレット関数のマスク．

を満たす．補助定理 3 から，$(\Re\Psi^{\mathrm{AS}}_{\alpha,P}; \Im\Psi^{\mathrm{AS}}_{\alpha,P})$ もアナライジングマルチウェーブレット関数になる．画像分離のエッジ抽出を行う線形変換 B と B_1 には，それぞれウェーブレット関数として $\Re\psi^{\mathrm{AS},p}_{\alpha,P}$ と $\Im\psi^{\mathrm{AS},p}_{\alpha,P}$ を用いた連続ウェーブレット変換を使う．

5.5.3 四角タイリングによるマルチウェーブレット関数

文献 [3] で提案したマルチウェーブレット関数を用いる．周波数平面を Fig. 5 右のように分割する．灰色の 4 角形と同じ大きさの 4 角形 12 個をマルチウェーブレット関数と

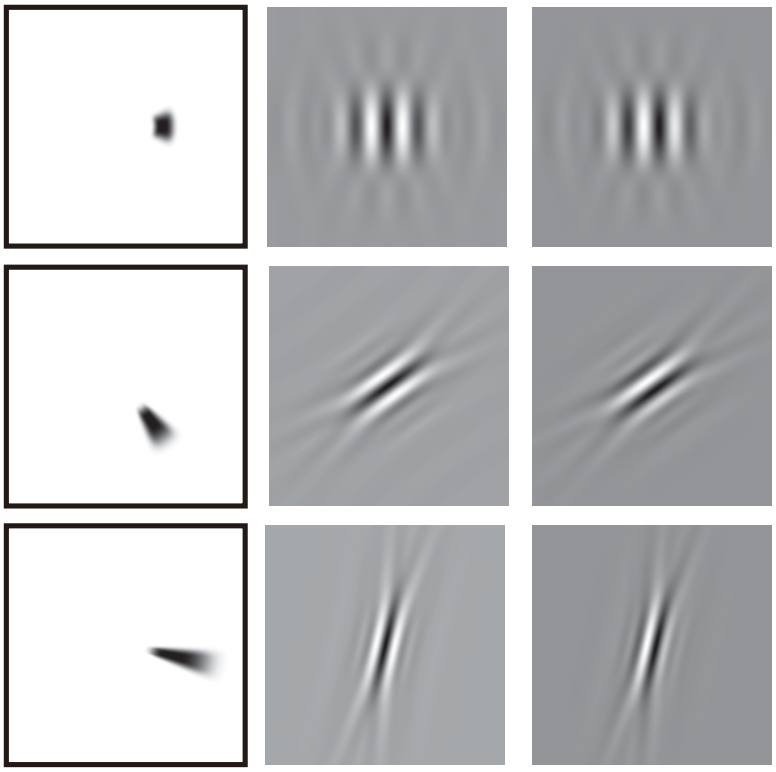


Fig. 6. 左：$\widehat{\psi^{\mathrm{AS},p}_{\alpha,P}}$, 中：$\Re\psi^{\mathrm{AS},p}_{\alpha,P}$, 右：$\Im\psi^{\mathrm{AS},p}_{\alpha,P}$. 上から $(\alpha, P, p) = (1.5, 12, 1), (2, 20, 4), (2.5, 28, 2)$.

する．

$$\Psi^{\mathrm{ST}}_{\alpha} = \left(\psi^{\mathrm{ST},p}_{\alpha}\right)_{p=1}^{12}$$

このマルチウェーブレット関数も周波数空間で 1 の分割

$$\sum_{j\in\mathbb{Z}}\sum_{p=1}^{12}\left|\widehat{\psi^{\mathrm{ST},p}_{\alpha}}(\alpha^j\xi)\right|^2 = 1, \qquad \text{a.e. } \xi \tag{5.11}$$

を満たす．Fig. 7 にこのウェーブレット関数のマスク（左），実部（中），虚部（右）を描いた．

四角タイリングによるマルチウェーブレット関数は，許容条件 (5.4) を満たしていない（満たすように変更はできる）ので，アナライジングマルチウェーブレットではないが，次小節で述べる離散化した関数系がパーセヴァルフレームになるので，そのまま用いる．画像分離のエッジ抽出を行う線形変換 B と B_1 には，それぞれウェーブレット関数 $\Re\psi^{\mathrm{ST},p}_{\alpha}$ と $\Im\psi^{\mathrm{ST},p}_{\alpha}$ を用いた連続ウェーブレット変換を使う．文献 [5] では，このマルチウェーブレット関数を用いて画像分離を行った．

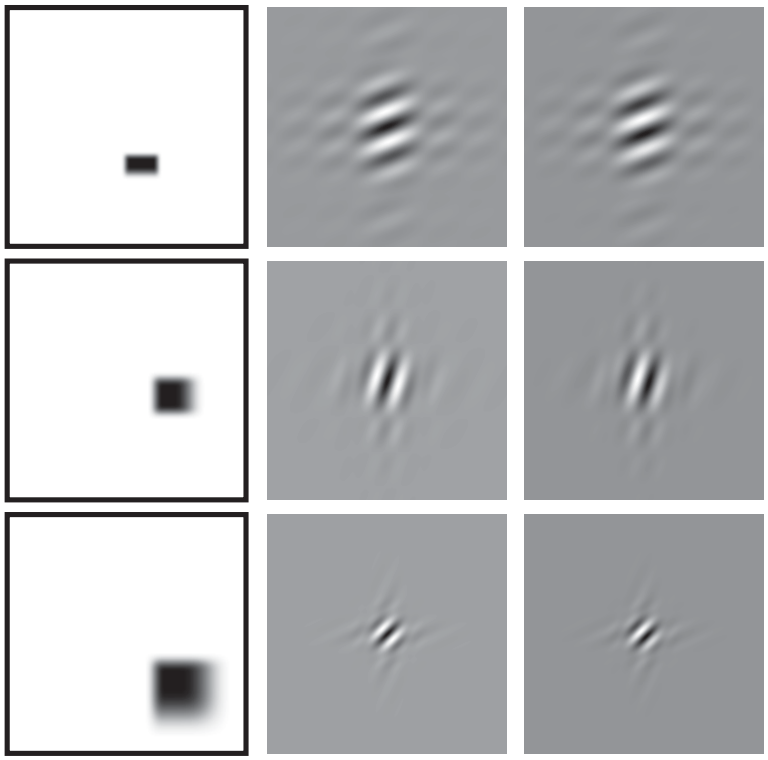


Fig. 7. 四角タイリングによるマルチウェーブレット関数左：$\widehat{\psi_\alpha^{\mathrm{ST},p}}$ のマスク中：実部 $\Re\psi_\alpha^{\mathrm{ST},p}$，右：虚部 $\Im\psi_\alpha^{\mathrm{ST},p}$．上から，$(\alpha, p) = (1.5, 1), (2, 3), (2.5, 2)$.

5.6 周波数空間の分割とパーセヴァルフレーム

前小節でデザインしたマルチウェーブレット関数は，周波数空間でコンパクトサポートを持ち，周波数空間での 1 の分解式 (5.9), (5.10), (5.11) を満たしていた．

定理 5 マルチウェーブレット関数 $\Psi = (\psi^p)_{p=1}^P \in L^2(\mathbb{R}^n)^P$ が周波数空間でコンパクトサポートを持ち，各 $\operatorname{supp}\widehat{\psi^p}$ が一辺の長さ 2π の立方体に含まれるとする．また定数 $\alpha > 1$ が取れて，周波数空間で 1 の分割

$$\sum_{j\in\mathbb{Z}}\sum_{p=1}^{P}\left|\widehat{\psi^p}(\alpha^j\xi)\right|^2 = 1, \qquad \text{a.e. } \xi \tag{5.12}$$

を満たすとする．このとき，関数系

$$\left\{\psi_{j,k}^p(x) := \alpha^{nj/2}\psi(\alpha^j x - k)\right\}_{j\in\mathbb{Z},\, k\in\mathbb{Z}^n,\, p=1,\ldots,P} \tag{5.13}$$

は，パーセヴァルフレームを構成する．つまり，任意の $f \in L^2(\mathbb{R}^n)$ に対して，

$$f = \sum_{j\in\mathbb{Z},\, k\in\mathbb{Z}^n,\, p=1,\ldots,P} \left\langle f, \psi^p_{j,k} \right\rangle \psi^p_{j,k},$$

$$\|f\|^2 = \sum_{j\in\mathbb{Z},\, k\in\mathbb{Z}^n,\, p=1,\ldots,P} \left| \left\langle f, \psi^p_{j,k} \right\rangle \right|^2$$

が成立する．

6. 位置スケール情報行列

文献 [2] で提案した方法を用いれば，混合行列を一度に求めることができる．観測画像 $y_j[q,r]$, $j = 1, \ldots, J$ にエッジ抽出用の線形作用素 B と B_1 を作用させた行列を $Y^B_j = By_j$ と $Y^{B_1}_j = B_1 y_j$ とかいた．変換行列の (q,r) 成分をならべた $2 \times J$ の行列 $G(q,r)$

$$G(q,r) = \begin{pmatrix} Y^B_1[q,r] & Y^B_2[q,r] & \ldots & Y^B_J[q,r] \\ Y^{B_1}_1[q,r] & Y^{B_1}_2[q,r] & \ldots & Y^{B_1}_J[q,r] \end{pmatrix} \in \mathbb{R}^{2\times J}$$

を**位置スケール情報行列**と呼ぶ．，k 番目の元画像のみが活動している位置の集合に含まれる点 $(q,r) \in E^B_k \cap E^{B_1}_k$ では， 位置スケール情報行列は，

$$\begin{aligned} G(q,r) &= \begin{pmatrix} a_{1,k}S^B_k[q,r] & a_{2,k}S^B_k[q,r] & \ldots & a_{J,k}S^B_k[q,r] \\ a_{1,k}S^{B_1}_k[q,r] & a_{2,k}S^{B_1}_k[q,r] & \ldots & a_{J,k}S^{B_1}_k[q,r] \end{pmatrix} \\ &= \begin{pmatrix} a_{1,k} & a_{2,k} & \ldots & a_{J,k} \end{pmatrix} \begin{pmatrix} S^B_k[q,r] \\ S^{B_1}_k[q,r] \end{pmatrix} \end{aligned}$$

と階数 1 になる．したがって，位置スケール情報行列を特異値分解 $G(q,r) = U(q,r)\Sigma(q,r)V(q,r)^{\mathrm{T}}$ し，第 1 特異値が第 2 特異値に比べてかなり大きくなる時を階数 1 として，$J \times J$ の直交行列 $V(q,r)$ の第 1 行を転置した J 次縦ベクトル $\vec{v}(q,r)$ を記録する．$\vec{v}(q,r)$ が混合行列のある列の大きさを 1 に正規化したベクトルの推定値なので，$\vec{v}(q,r)$ の分布を $\mathbb{R}^J$ の単位球面 $\mathbb{S}^{J-1}$ 上に描いて，ピークの個数から元画像の数が推定できる．またピークを取る座標から混合行列が推定できる．自己組織化地図 [13] を用いれば，多次元のクラスタリングを効率よく行える．

7. 数値実験例

4 個の元画像を混合行列 A

$$A = \begin{pmatrix} 0.3977 & 0.1940 & 0.4265 & 0.4982 \\ 0.4907 & 0.1761 & 0.2611 & 0.3875 \\ 0.3137 & 0.3519 & 0.3949 & 0.4821 \\ 0.3157 & 0.4351 & 0.4320 & 0.2544 \end{pmatrix}$$

で混合した Fig. 8 の観測画像の分離実験を行う．円環分割マルチウェーブレット関数で，

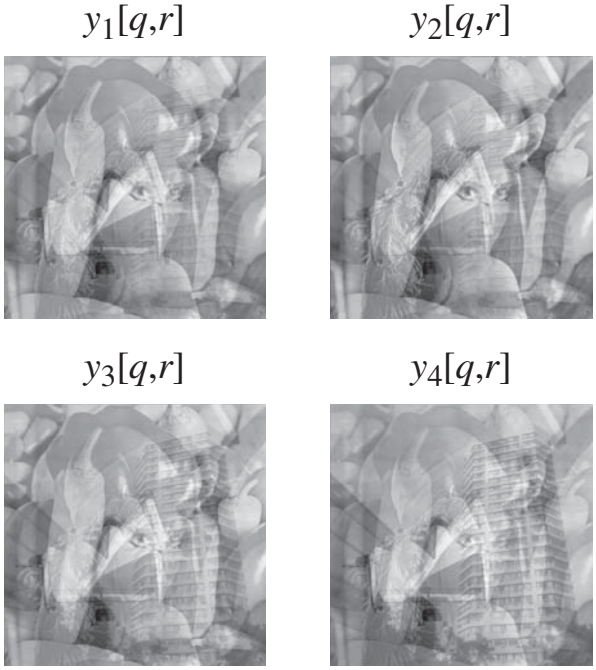


Fig. 8.　観測画像 y_j, $j = 1, \ldots, 4$.

$\alpha = 2$，円環を偏角方向に $P = 28$ 等分した関数 $\psi_{\alpha,P}^{\mathrm{AS},p}$ の実部と虚部をウェーブレット関数として用いた連続ウェーブレット変換をエッジ抽出用の線形作用素として用いた．

位置スケール情報行列は，2×4 の行列になり，階数 1 で取り出した 4 次縦ベクトル $\vec{v} = (v_1, \ldots, v_4)^{\mathrm{T}}$ を $p = 1, \ldots, 14$ に渡って記録したベクトルを用いてヒストグラムを描く．

4 次元のヒストグラムなので v_j, v_k の 2 次元に射影したヒストグラムを 6 種類描くと Fig. 9 を得る．ピークの数はどのヒストグラムでも 4 なので，元画像の数は $M = 4$ であると推定する．No. 3 のヒストグラムが見た目が一番良いとし，ピーク間の最小距離や一番低いピークの高さなどのパラメータを用いて，4 次元空間で自己組織化地図 [2, 13] を用いて混合行列を推定する．推定した混合行列 $\widetilde{A}$ と混合行列 A の各列の大きさを 1 に正規化した混合行列 $\bar{A}$ は，

$$\widetilde{A} = \begin{pmatrix} 0.5162 & 0.5975 & 0.3170 & 0.5534 \\ 0.6337 & 0.4659 & 0.2861 & 0.3464 \\ 0.4079 & 0.5773 & 0.5691 & 0.5127 \\ 0.4067 & 0.3043 & 0.7026 & 0.5573 \end{pmatrix}, \qquad \bar{A} = \begin{pmatrix} 0.5147 & 0.3140 & 0.5540 & 0.5974 \\ 0.6351 & 0.2850 & 0.3392 & 0.4646 \\ 0.4060 & 0.5695 & 0.5130 & 0.5781 \\ 0.4086 & 0.7042 & 0.5612 & 0.3050 \end{pmatrix}$$

であって，$\widetilde{A}$ の第 1, 2, 3, 4 列がそれぞれ正規化した混合行列 $\bar{A}$ の第 1, 4, 2, 3 列に十分な精度で対応していることがわかる．分離した推定画像を Fig. 10 にあげる．また分離精度を Table 1 にあげる．

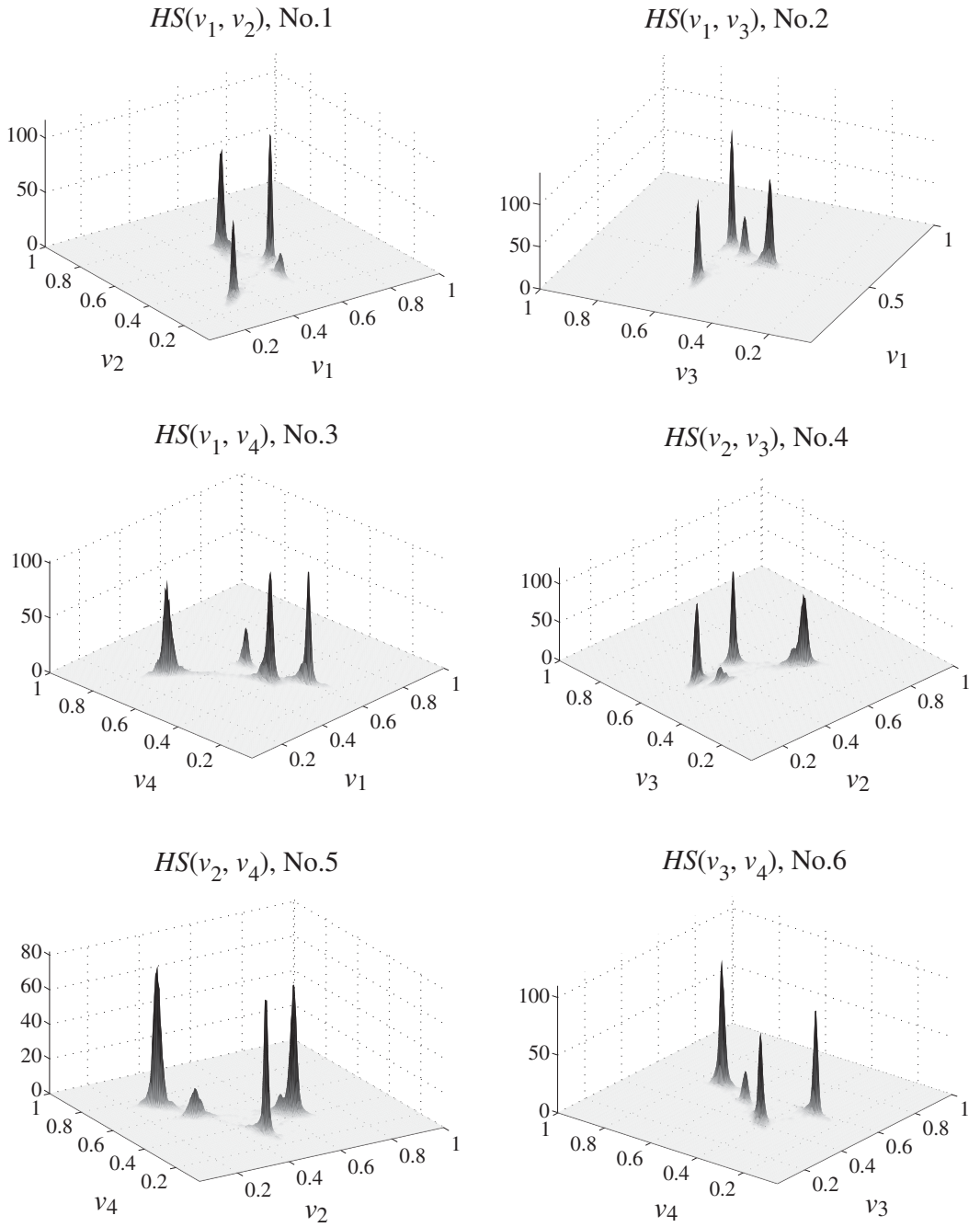


Fig. 9.　4 次元ヒストグラムの 2 次元への射影．No. 3 をベストとして選ぶ.

謝辞　本講演は，科学研究費補助金 (C)22540130, (C)23540135 および文部科学省「数学・数理科学と諸科学・産業との連携による数学イノベーションの推進」および大阪教育大学の助成を受けて行われた.

Fig. 10. 分離した画像 σ_j, $j = 1, \ldots, 4$.

Table 1. 元画像 (SI) と 推定画像 (ESI) の誤差評価.

SI	ESI	square error	SNR
s_1	σ_1	0.01296 %	38.88 [dB]
s_2	σ_3	0.00031 %	55.05 [dB]
s_3	σ_4	0.00131 %	48.82 [dB]
s_4	σ_2	0.01066 %	39.72 [dB]

参考文献

[1] R. Ashino, C. A. Berenstein, K. Fujita, A. Morimoto, M. Morimoto, D. Napoletani, and Y. Takei, *Mathematical background for a method on quotient signal decomposition*, Appli. Anal., **86** (5), 577–609, 2007.

[2] R. Ashino, K. Fujita, T. Mandai, A. Morimoto, and K. Nishihara, *Blind source separation using time-frequency information matrix given by several wavelet transforms*, Information, **10** (5), 555–568, 2007.

[3] R. Ashino, C. Heil, M. Nagase, and R. Vaillancourt, *Microfiltering with multiwavelets*, Comput. Math. Appl. **41** (1-2), 111–133, 2001.

[4] R. Ashino, S. Kataoka, T. Mandai, and A. Morimoto, *Blind image source separations by wavelet analysis*, accepted in Appli. Anal.

[5] R. Ashino, S. Kataoka, T. Mandai, and A. Morimoto, *Image separation using multi-wavelets*, Proceedings of the 2011 International Conference on Wavelet Analysis and Pattern Recognition, Guilin, 10-13 July, 2011, 245–250, 2011.

[6] R. Ashino, T. Mandai, A. Morimoto, and F. Sasaki, *Blind source separation of spatio-temporal mixed signals using time-frequency analysis*, Appli. Anal., **88** (3), 425–456, 2009.

[7] R. Ashino, T. Mandai, and A. Morimoto, *Blind source separation of spatio-temporal mixed signals using phase information of analytic wavelet transform*, Int. J. Wavelets Multiresolut. Inf. Process., **8** (4), 575–594, 2010.

[8] C. Cherry, *Some experiments on the recognition of speech, with one and with two ears*, Journal of Acoustical Society of America, **25** (5), 975–979, 1953.

[9] A. Cichocki and S. Amari, *Adaptive blind signal and image processing: Learning algorithms and applications*, John Wiley & Sons, Chichester, West Sussex, England, 2002.

[10] K. Gröchenig, *Foundations of Time-frequency Analysis*, Birkhäuser Boston, Inc., Boston, MA, 2001.

[11] S. Haykin and Z. Chen, *The cocktail party problem*, Neural Computation, **17** (9), 1875–1902, 2005.

[12] A. Jourjine, S. Rickard, and O. Yilmaz, *Blind separation of disjoint orthogonal signals: Demixing N sources from* 2 *mixtures*, in 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing Proceedings, **5** (2000), 2985–2988.

[13] T. Kohonen, *Self-Organizing Maps*, Second Edition, Springer-Verlag, Berlin, 1997.

[14] A. Morimoto, R. Ashino, S. Kataoka, T. Mandai, *Image separation using monogenic signal of stationary wavelet transform*, Proceedings of the 2011 International Conference on Wavelet Analysis and Pattern Recognition, Guilin, 10-13 July, 2011, 239–244, 2011.

[15] A. Morimoto, R. Ashino, and T. Mandai, *Image separation using the monogenic wavelet transform*, Proceedings of the 10th International Symposium on Communications and Information Technologies 2010 (ISCIT 2010), 707–712, 2010.

[16] G. Nason and B. Silverman, *Stationary wavelet transform and some statistical applications*, Lecture Notes in Statistics, **103**, 281–299, Springer-Verlag, New York, NY, 1995.

[17] 守本晃・神山浩之・井上大樹・大道淳史・西村一志・芦野隆一・萬代武史, **ウェーブレット解析を用いた画像分離**, 日本応用数理学会論文誌, **19** (3), 257–278, 2009.

守本 晃 (大阪教育大学情報科学)
〒582-8582 大阪府柏原市旭ヶ丘 4-698-1
E-mail: morimoto@cc.osaka-kyoiku.ac.jp

音声からの声道長推定における聴覚的ウェーブレット変換について

入野 俊夫 *
* 和歌山大学　システム工学部

概要. 音声を一声聞くだけで, 大人か子供かすぐわかる. 同時に話者の寸法に無関係に言語情報も獲得できる. このことから, 人間は, 寸法（スケール）と声道形状（音韻性）を分離抽出する機構を持っているものと考えている. この聴覚理論として, 安定化ウェーブレット-メリン変換を提案してきた. この理論を受けて, 最近, ２８人の話者間の総当たりで声道長比を求め, 推定誤差を求める実験を行った. 数種類の「聴覚的スペクトル」上で１組のスペクトルの片方をスケール変形を行いもう一方とのスペクトル距離が最も小さくなる場合のスケール係数を声道長比とした. さらに回帰分析を実施し, 安定性の指標となる推定誤差を計算した. この時, 推定精度が最も良かったのは, 実際に心理実験から求めた非線形性がある聴覚フィルタを用いた場合であった. 問題が, 線形のウェーブレット変換では扱えるスケール変形だけではないことがわかる. これらの背景と結果を紹介し, 聴覚的非線形性も含めた理論的枠組みをぜひ議論していただきたい.

Auditory wavelet transform for vocal tract length estimation from speech sounds

Toshio Irino*

*Faculty of Systems Engineering, Wakayama University

Abstract.

We hear vowels pronounced by adults and children as approximately the same although the vocal tract length (VTL) varies considerably from group to group. At the same time, we can identify the speaker group. This suggests that the auditory system can extract and separate information about the size of the vocal-tract from information about its shape. We had proposed a computational theory, named Stabilized Wavelet-Meliin Transform (SWMT), to explain the observation. Recently, we performed a VTL estimation experiments using the knowledge of the theory. We found that the nonlinear auditory filter bank , which was estimated by psychoacoustical measurement, was better than any other linear filterbanks including wavelet-like one. This implies the problem of the VTL estimation in real speech sounds is not solely the issue of the scale estimation which can be dealt with the wavelet transform. In this talk, we introduce the background and results for the discussion of the theoretical framework including the auditory nonlinearity.

1. はじめに

音声（有声音）は, 音響管である声道を声帯音源によって駆動することによって生成される. これは,「ソースフィルタモデル」と呼ばれる. 母音の違い（たとえば/a/ と/i/）は, 声道の形状の違いで表現され, スペクトル上では共振周波数の組み合わせが異なる. 一方, 大人でも子供でも, 同じ母音/a/は, /a/として発声できる. ところが, 頭の寸法が異なり声道長も異なるため, スペクトル分布は異なる. ただ, ここで共振周波数が声道長の逆数に比例したスケール関係におおよそなっている. したがって, 声道長 (Vocal Tract Length, VTL) を正規化することにより, ある程度スペクトル分布を揃えることができる考えられている.

この声道長正規化は, 不特定話者の自動音声認識で有効な手法とされ, Wakita [1] 以来様々な方法が提案されている. また, 最近, 声道長正規化による２話者間の音声モーフィング（特に男女間）の音質が改善されることが報告されている [2, 3]. これらの基本となっている手法では, 短時間フーリエ変換の直線周波数軸をメル周波数軸等の疑似対数軸に周波数ワーピング関数等により変換し, その上でスペクトルシフトやシフト不変変換を行う. このうちのどの要素を改善すれば良いかや, 推定法（たとえば学習法）をどのようにするかが議論の対象となってきた. ただし, 近年では, スペクトル表現自体の妥当性の議論はされていないようである.

本資料では, 聴覚系における声道長正規化の理論 [4] やその背景を示す. ここで, 聴覚系とウェーブレット変換との関係も紹介する. さらに, この聴覚的な知見を聴覚フィルタバンクレベルで導入した声道長推定手法と有効性について紹介する. 結果としては, 聴覚系の非線形性が入っている場合の方が線形系よりも推定精度が良かった. 最後の節でこの非線形性について述べ, 理論構築のベースを提供したい.

2. 聴覚系における寸法-形状知覚と理論

2.1 聴知覚特性

図 1 に示すように, 大人と子供が発声した同じ言葉を聞いたとき, これらの音声はスペクトル分布としては異なっていても同じ言葉として知覚することができる. また同時に話者が大人か子供かを認識することが可能である. このことから, 人間の初期聴覚系において, 聞いた音から発音体の寸法（=声道長）情報と形状（=声道の断面積関数）情報に分離し, 抽出する機能があるという仮説を立て, 理論を提案している [4–6] . これを受けて聴知覚実験も行われ, 寸法の弁別閾はおおよそ 5% 程度であることがわかっている. さらに, 通常発声範囲をはるかに超えた基本周波数-声道長の組み合わせの合成音やささやき声においても, 音韻や単語の正解率が高いことがわかっている [8–10] .

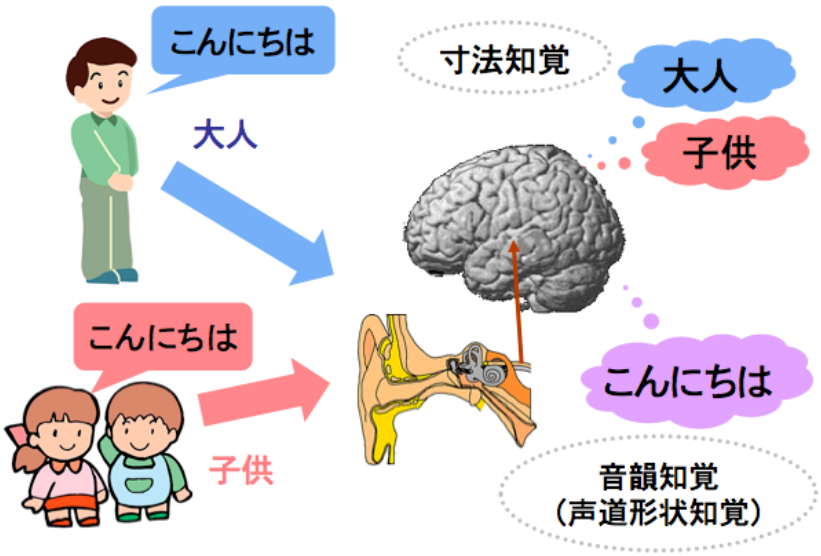


Fig. 1. Size and Shape perception from sound source.

2.2 聴覚計算理論

上記の知覚特性を説明するために, 初期聴覚系で寸法情報と形状情報の分離抽出を行っているという計算理論を提案している [4–6]. 図 2 に, このアルゴリズムである安定化ウェーブレット-メリン変換のブロック図を示す. 各々のステージについて, 背景をある程度含め説明する.

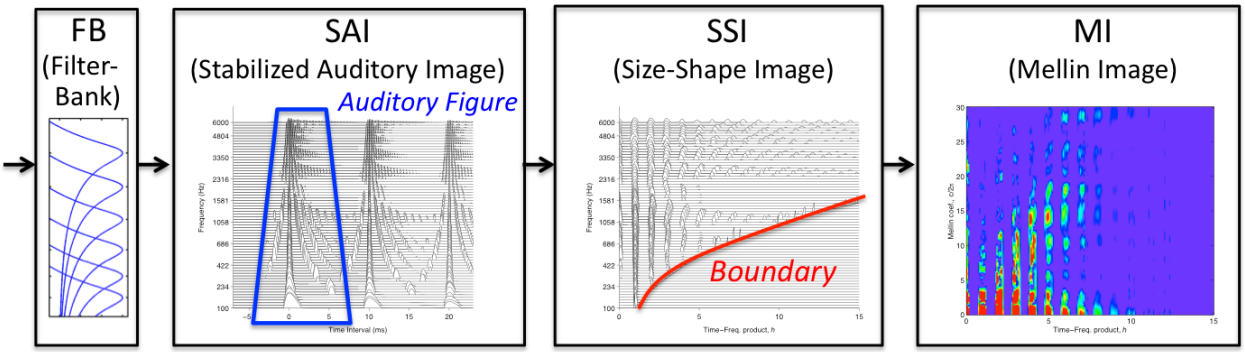


Fig. 2. Computational theory of the size-shape perception: Stabilized Wavelet-Mellin Transform

2.2.1 聴覚末梢系のフィルタバンクモデル

入力された音は, 聴覚フィルタバンク (FB) で, 時間軸と $\mathrm{ERB_N}$ 軸 [7](疑似対数周波数軸) を持つスペクトログラム的な分析が行われる. また, 実際に聴神経の活動まで模擬する

場合は半波整流を行い, 神経活動パターン (NAP) と呼ぶ表現にする. この聴覚フィルタの特性は, 心理物理実験的に推定できる [11–14]. 推定されたフィルタ特性は非線形を持ち, 入力音圧に依存して周波数特性が変化し, 利得も変化する（圧縮特性を持つ）ことが知られている. これらの非線形性に関しては 4 節で述べるが, 線形の第一次近似としてはウェーブレット変換に似ていると古くから指摘されている [17].

この聴覚末梢系の周波数分析に関しては研究の歴史は長く, 古典的な機械振動解析から, 単純ではあるが見通しの良いフィルタバンクまで, 数多くのモデルが提案されている [18]. フィルタバンクの周波数特性の一例を, 図 3 上図に示す.

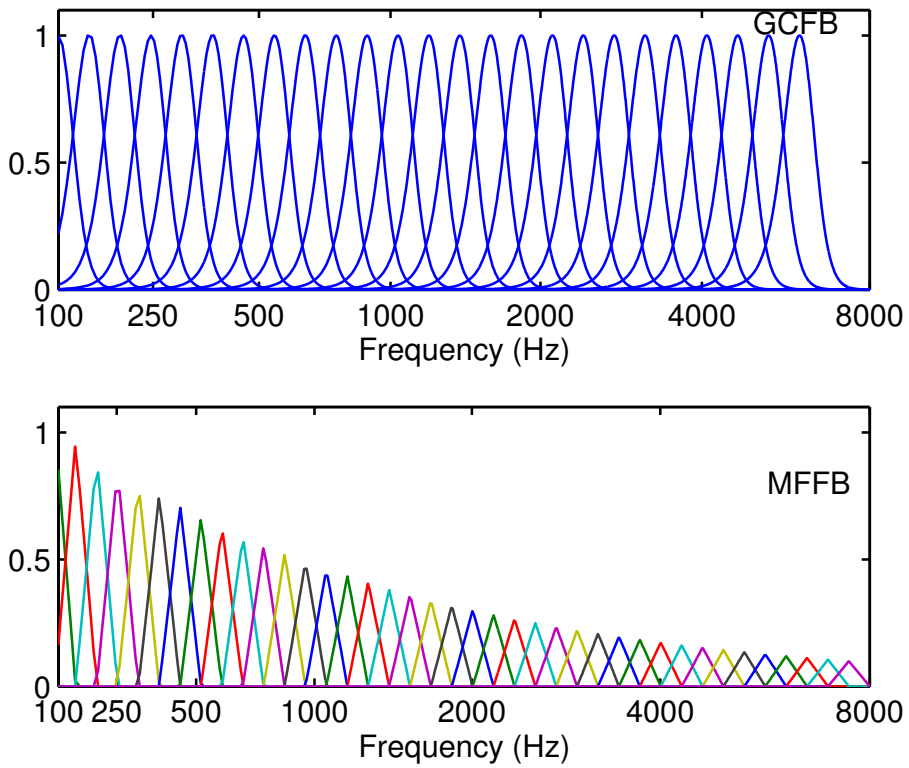


Fig. 3. Characteristics of gammachirp filterbank (upper panel). The number of the filter is restricted for the plot. Characteristics of mel-frequency filterbank (lower panel).

音響管の寸法が変化すると, インパルス応答が時間的に伸縮される, スケール変形となる. この音のスケール変形に対して, フィルタ系による歪みを与えないという意味では, 線形のウェーブレット変換が最も良い. これは, どのフィルタも同じインパルス応答 (kernel 関数) でスケールのみが違うため, 外界の音がスケール変形しても必ず同じ形のフィルタで処理されるからである. このウェーブレット変換では, 周波数と帯域幅が比例する定 Q 特性が成立していることが必要条件となる. 聴覚心理実験の結果から得られた, フィルタの中心周波数と帯域幅の関係を図 4 に示す. 縦軸は, 聴覚心理物理でよく用いられる等価矩形帯域幅 (Equivalent Rectangular Bandwidth, ERB) で, 健聴者の ERB_N (Hz) はフィルタ

の中心周波数 f_c (Hz) に対し次式で与えられる [11].

$$ERB_N = 24.7 \cdot (4.37 * f_c/1000 + 1).$$

この図を見ると, おおよそ 500Hz 以上において周波数と帯域幅が比例し, 定 Q 特性を満足していることがわかる. すなわち, その領域ではウェーブレット変換を用いてフィルタ系を構成できることになる.

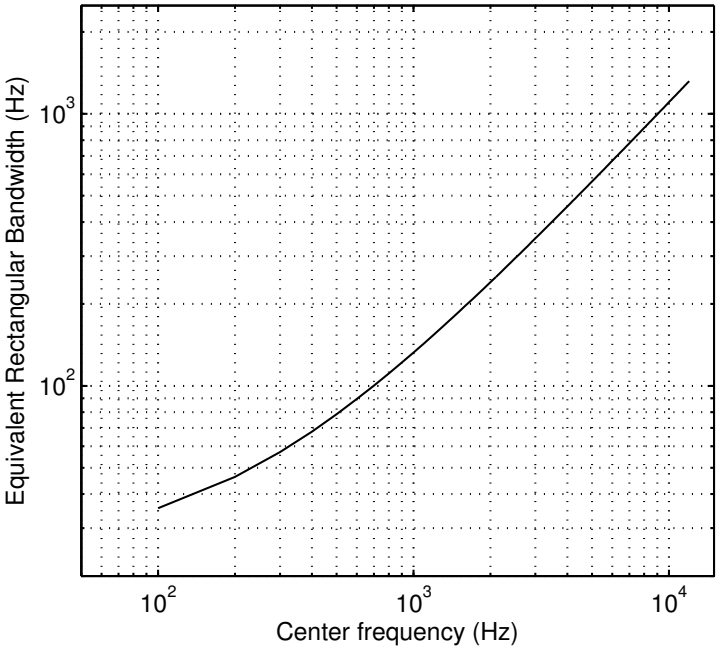


Fig. 4. The relationship between the center frequency and the bandwidth estimated by psychoacoustic experiments. This is used for the gammatone filter. The bandwidth for the gammachirp filter is about 1.5 times.

フィルタバンクを構成する各チェンネルのフィルタ（kernel 関数) としては, ガンマトーン (gammatone) の系統が最も有力である. このガンマトーン[*1]は, 生理実験で求められたネコの基底膜振動のインパルス応答を近似するための実験式として元々提案されたものである [15]. その後, 様々な変遷を経て, 現在まで最も良く使われるフィルタ系となっている. この中には, Lyon が提案した one-zero gammatone や Meddis らの DRNL, Irino and Patterson のガンマチャープ (gammachirp) などがある (経緯や文献は [16, 18–20] 参照).

このガンマチャープ[*2]. は, 以下で述べる初期聴覚系の内部表現（スケール表現）の考察に踏み込み, Mellin 変換 (3.2.4 項参照) と時間 (間隔) 軸で張る空間の最小不確定性を持つ

[*1] ガンマトーン (gamma-tone) は, 包絡線がガンマ関数 (gamma) で, 搬送波が正弦波のトーン (tone) であることからの造語である.

[*2] ガンマチャープ (gamma-chirp) は, 包絡線がガンマ関数 (gamma) で, 搬送波が周波数変化のあるチャープ波 (chirp) であることから命名された [20]. このガンマチャープ関数は Gabor 関数同様, 初期位相を適切に選ばない限り周波数 0 で値が 0 にならず admissible 条件を満たさないため, 厳密な意味でのウェーブレットカーネルとはならない. 聴覚系自体に合成系は無いので, 条件を緩めた「半ウェーブレット」的な扱いが

関数として関数解析的に求められたものである [20] . Appendix A にその導出を示す. ガンマチャープの特殊解であるガンマトーンも含めた聴覚フィルタは, 寸法情報処理に最適な系を構成していると解釈することができる.

2.2.2 初期聴覚系における時間積分

音量の小さな短音の数を増やしていくと, 聞こえる音の大きさ（ラウドネス）が徐々に大きくなることが知られている. これは, 聴覚系に時間積分の機能があることを示している [9]. この説明モデルとして, 時間窓をかける形の積分（スムージング）を考えることができる. しかし, 人間は時々刻々変化する微妙な音色も, 同時に聞き分けることもできる. そこで, この時間的な詳細構造 (temporal fine structure) を保持する機構が別途必要となるが, あまり良いモデルは提案されていない. これらを説明するために, 音の大きさ知覚や微細構造知覚といった現象ごとに別個のモデルを作ることは, オッカムの剃刀の原理からも本質から遠ざかる可能性が大きい.

そこで時間積分の特性を持ちつつ時間的詳細特性の保持するために考えだされたのが, ストローブ時間積分 (strobed temporal integration) である [5, 7, 18]. これは, 振動体をストロボスコープ*3を用いて撮影した場合や, オシロスコープの同期モードで波形を見る状況と類似のものと思えば良い. 聴覚モデルにおいては, 各々の周波数チャンネルごとに, ある時点の神経活動パターンを, 時間間隔と周波数の軸を持つ２次元のイメージバッファにピーク時点を同期させながら積分する. たとえば, 音声であれば基本周期ごとに類似な神経活動パターンが繰り返される. これをピッチパルスに同期して積分する. これで得られる表現を, 安定化聴覚イメージ (SAI) と呼ぶ（図 2 の 2 ブロック目). この 2 次元イメージは, 入力が音声ならば定常母音では定常的で, 音節の移り変わりで変化する動画的な表現となる. この SAI の上では, 基本周波数ごとに同じ活性度パターンが繰り返される. この 1 周期分が, 話者の声道の共振特性を示す聴覚図 (Auditory Figure, AF) である.

この聴覚図 (AF) を用いれば, 話者の発声している音韻や話者の声道長を安定に推定できるはずである. 一方, 音声の留まらず, この SAI の表現上での様々な場所で特徴ベクトルを取り, Web 上の音検索に使う試みも最近提案されている [21].

2.2.3 スケール共変性表現

この聴覚図 (AF) を, 縦軸のチャンネルごとに中心周波数分だけ時間間隔軸を引き延ばして正規化すると, フィルタ側のインパルス応答が全チャンネルでそろった形で表現できるようになる. この処理は, ウェーブレット変換で伸縮比を正規化して, 同一のインパルス応答 (kernel 関数) を求め, 全チャンネルにそろえて並べていることに相当する. 得られ

あれば良いのかもしれない. もっとも, 音声処理に関しては低い周波数（50 Hz 以下) は無視できるので, ガンマチャープでも実質的に分析合成系を構成できる.

*3 ストロボ/光源を一定間隔で一瞬発光させる装置. 振動体の振動周期に同期させると静止画撮影も可能である.

た表現を寸法形状イメージ (SSI) と呼ぶ (図 2 の 3 ブロック目). この表現上では, 声道長伸縮（スケール変化）の効果は, 伸縮の無い同一パターンの上下移動として単純化されて表現されることになる. これがスケール共変表現である.

ここで, 低い周波数側では, SAI において聴覚図 (AF) が基本周期ごとに重なることに注意が必要である. SSI を取るときにこの重なり部分が切り捨てられるため, 図 2 の 3 ブロック目に示した, Boundary(境界線) の下側に活性度が無い空白部分ができる. 左端のストローブしたピッチパルス時点から離れるに従い, 有効なパターンの下限周波数が高くなる. この空白部分は, 本来音源がインパルスであれば表すことのできた声道特性が, 基本周期 (=$1/F_0$) の影響により表現できない所である. これは, 声道の音響管を短い周期の声帯振動によって駆動する音声生成過程の避けがたい特徴である. 声道長を安定に推定するためには, 音響管の共振特性と駆動源の励振特性をスペクトル情報から上手に切り分ける必要がある.

2.2.4 スケール不変特徴

最終段は, SSI の縦方向にフーリエ変換をし絶対値を取って寸法を正規化した, メリンイメージ (MI) である (図 2 の最終ブロック). 対数軸に対するフーリエ変換はメリン変換に相当する. この時, 寸法情報は位相項として得られる. この処理は, 大脳の一次聴覚野で表現されている周波数軸に順序よく並んだトノトピー表現空間から, 周波数成分を取り除き, さらに内部の処理に進む段階となる. この意味で, このメリンイメージは, Shamma の提案する大脳皮質の受容野 (Receptive Field, RF) [22, 23] の一部を表現していると位置づけられる. 逆に言えば, RF の中には Mellin 変換として定式化できるものがあるものと考えられる.

3. 音声からの声道長推定

フィルタバンク（あるいはウェーブレット変換）によって, 話者の寸法（声道長）やスケールを推定する課題を考える. この声道長推定を安定に行うためには, 前節の初期聴覚系の知見, 特に 2.2.3 項の聴覚図 (AF) の下限周波数の考え方を取り入れることが重要となるであろう. ここでは, 様々なフィルタバンクを比較するためにも, 初段の聴覚フィルタバンクの出力だけから声道長を推定する問題を考える.

3.1 フィルタバンクと周波数領域の選択

「聴覚的」と称するフィルタバンクは数多く提案されている. この中から, 最も良いものを選ぶ必要がある. さらに, 前節で述べたように低い周波数領域は基本周波数や音声の駆動音源波形の形状の影響を受ける. 特に聴覚図 (AF) で表現できる下限周波数には注意が必要である. また, 逆に高域は個人性の影響が大きく, 例えば 4〜5kHz 付近に梨状窩によ

る零が存在する場合もある. この２つの領域に挟まれた間に, 声道長情報が最も良く表われる領域があるはずである. そこで周波数帯域の選択によって, 推定誤差がどのように変化するかを調べ, 誤差最小となる条件を設定する必要がある.

3.2 推定実験

実験の詳細は別報告 [3, 24–26] に譲るが, 以下で概要を述べる.

3.2.1 2 話者間の声道長比の推定

2 人の話者 i, j を設定する. 一般に声道長が異なるためスペクトル分布が異なる. そこで, 片一方のスペクトル S_j をスケール伸縮の r 倍をし, もう片一方の話者のスペクトル S_i と最もマッチングする所を探すことを考える. そこで, 2 つのスペクトルの距離が最小となるスケール伸縮比率 $r_{i,j}$ を, その 2 人の話者 i, j の組み合わせにおける声道長比の推定値とする.

3.2.2 全声道長比の推定

男女計 28 名の話者間の声道長の比を総当たり ($_{28}P_{27} = 756$ 通り) で推定する. そこでは以下の 11 種類のフィルタバンクを, 計 56 種類の周波数帯域について, 3 文章を用いて行った. スケール伸縮比率 $r_{i,j}$ を求めるアルゴリズムは, 最小化したいスペクトル距離を $D_{spec}(i, j, r)$ として, 以下のように表される.

for $N_{filterbank} = 1 \rightarrow 11$ (for all filtebanks) **do**
  for $N_{sentence} = 1 \rightarrow 3$ (for all sentences) **do**
    for $N_{Fregion} = 1 \rightarrow 56$ (for all combinations of frequency region) **do**
      for $i = 1 \rightarrow 28$ (for all speakers) **do**
        for $j = 1 \rightarrow 28, i \neq j$ (for all speakers except for the same) **do**
          $r_{i,j}(N_{filterbank}, N_{sentence}, N_{Fregion}) = \arg\min_{r}(D_{spec}(i, j, r))$
        end for
      end for
    end for
  end for
end for

声道長比推定を全 140 万回 (=$11 \times 3 \times 28 \times 27 \times 56$) 行う比較的大規模な実験である. この各々の要素について以下で述べる.

3.2.3 回帰分析

各フィルタバンク・文・周波数領域について, 声道長比 $r_{i,j}$ が求まった時点で, Appendix B に示した手法で回帰分析を行う [3, 24–26]. 回帰分析の結果求まった声道長比 $\hat{r}_{i,j}$ と, 元の

$r_{i,j}$ の差の rms 値を推定誤差とした. これは, 1 人の話者が 1 つの声道長の真値を持っているとし, 選んだ 2 話者の比を取った値に対して, どの程度ずれるかを測っていることになる. どの話者の組み合わせや, どの発話内容であったとしても, ばらつきが小さければ安定な推定と言うことができる.

理想的には声道長の真値がわかれば良いが, 実際には, たとえ MRI 装置を用いた声道断面測定を行っても明確に特定できない. また, 実際の声道長と音声スペクトルとの関係は, 第 1 次近似としてはスケール関係（比例関係) が成立するが, まだ詳細には解明されていない. さらに, ここではスペクトルマッチングだけが目的のため, 単純なスケール関係を考えている (Appendix B 参照).

3.2.4 比較対象のフィルタバンク

「聴覚的」フィルタバンクは様々提案されているが, フィルタバンクの種類によりスペクトル表現が異なるため, 性能が異なるはずである. ここでは, ガンマチャープフィルタバンク (GCFB), 広く用いられているガンマトーンフィルタバンク (GTFB), 音声認識で最も用いられているメル周波数フィルタバンク (MFFB) を比較対象として, 以下の 11 条件を設定した. STRAIGHT 以外は, 25ms の hamming 窓でパワーを平均化したスペクトログラムを用いた.

- $\mathrm{GCFB_{dyn}}$： 動的圧縮型ガンマチャープフィルタバンク [14]（非線形の時変フィルタ）. 周波数範囲 [100, 6000] で 100ch とした.
- $\mathrm{GCFB_{lin}}$：固定係数の線形の圧縮型ガンマチャープフィルタバンク.
- $\mathrm{GTFB_{100}}$：現在標準的に良く用いられる, ガンマトーンフィルタバンク [10] .
- $\mathrm{GTFB_{050}}$：帯域幅を 0.5 倍した GTFB.
- $\mathrm{GTFB_{025}}$：帯域幅を 0.25 倍した GTFB.
- $\mathrm{MFFB_{STR24}}$：メル周波数フィルタバンク . 周波数特性は図 3 下図に示した [27]. ここでは, HTK 形式に準拠し, [0,8000] で 24ch とした. STRAIGHT スペクトルを基にして構成.
- $\mathrm{MFFB_{STR40}}$：同上. 40ch.
- $\mathrm{MFFB_{STR120}}$：同上. 120ch.
- $\mathrm{MFFB_{STFT24}}$：短時間フーリエ変換 (STFT) を用いた標準的な MFFB. それ以外は $\mathrm{MFFB_{STR24}}$ と同条件.
- $\mathrm{MFFB_{STFT40}}$：同上. 40ch.
- $\mathrm{MFFB_{STFT120}}$：同上. 120ch.

$\mathrm{MFFB_*}$ は, 短時間フーリエ変換や STRAIGHT で時間-周波数表現にした上での重み関数である. その意味ではインパルス応答は定義されていない. 図 3 下図に示すように, 帯域ごとの重なり部分を加算するとちょうどピークに対して 1 の割合となるように定義されているので, 形式上コンプリートフィルタバンクの形である. これに対し, $\mathrm{GCFB_*}$(図 3 上図)

や GTFB_* はフィルタどうしの重なりが大きくオーバーコンプリートフィルタバンクの形式になっている. また, $\mathrm{GCFB}_{\mathrm{lin}}$ は, GTFB_{100} の約 1.5 倍の帯域幅を持つフィルタから構成されているため, オーバーコンプリートネスもさらに高い.

3.2.5 推定のための周波数領域と評価用音声

聴覚図 (AF) の知見から, 声道長の推定に用いる周波数領域を制限した方が良い可能性がある. そこでここでは, 様々な周波数領域を検討するため, 下限周波数 100〜800Hz で 100Hz 刻み, 上限周波数 2000〜8000Hz で 1000Hz 刻みで設定した. これらの組み合わせは 8×7(=56 通り) のメッシュ状で, 各点ごとに推定誤差を計算した.

また, 音声サンプルによって, 推定される声道長が異なる可能性もある. そこで, 推定の安定性を評価するために, 長さの異なる 3 文章（各々 10, 14, 20 音節で構成）の発話を用いた. 話者男女各 14 名が同一の文章を発話した音声サンプル間で声道長比を計算する.

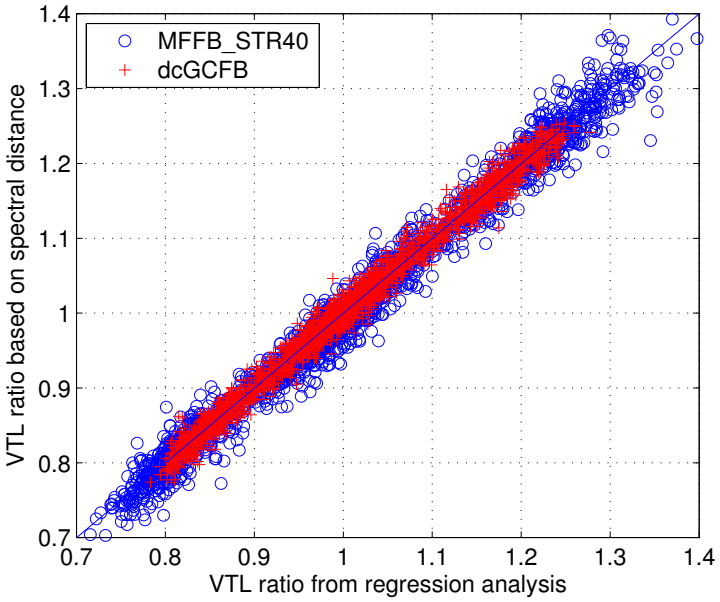


Fig. 5. Relationship between VTL ratios $\boldsymbol{r}$ and $\hat{\boldsymbol{r}}$ estimated using $\mathrm{GCFB}_{\mathrm{dyn}}$ (+) and $\mathrm{MFFB}_{\mathrm{STR40}}$ (∘) with best frequency regions.

3.3 実験結果

図 5 に, $\mathrm{GCFB}_{\mathrm{dyn}}$ (dcGCFB) (+) と $\mathrm{MFFB}_{\mathrm{STR40}}$ (∘) で推定した声道長比を示す. 横軸は, 回帰分析の結果の声道長比 ($\hat{r}$), 縦軸は, 元のスペクトル距離から求めた声道長比 (r) である. また, それぞれ最も良く推定された周波数領域を選んでいる. この図から, $\mathrm{GCFB}_{\mathrm{dyn}}$ の方が推定値のばらつきが小さいことがわかる. その分安定に推定できていると考えられる.

図 6 に, フィルタバンクの種類ごとに最良周波数帯域を選択した場合の誤差を棒グラフ

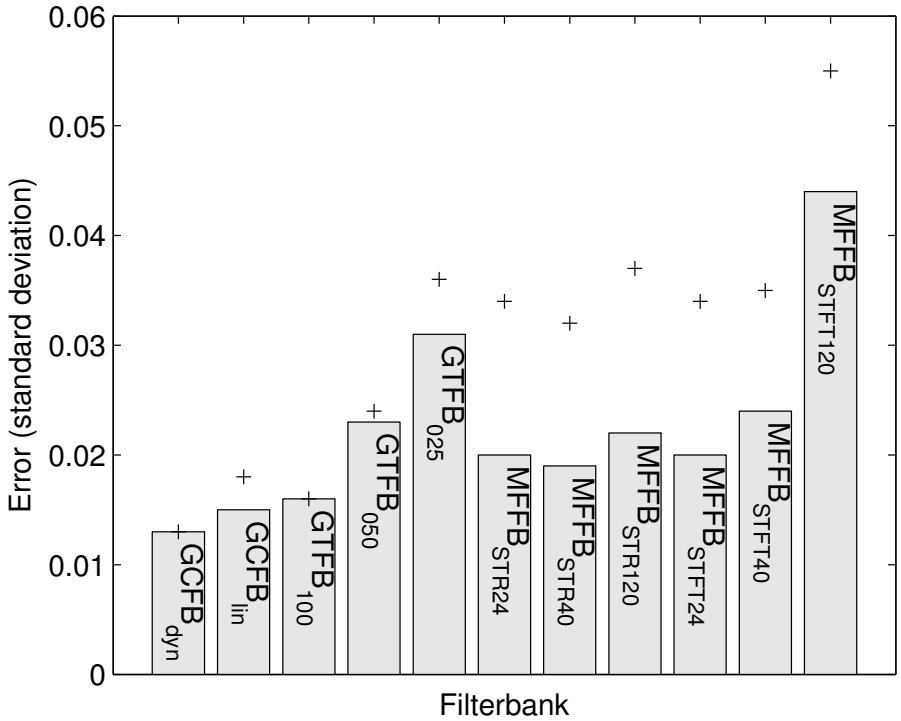


Fig. 6. Standard deviation for the filterbanks. Bar shows the minimum error when the frequency range is properly selected. + shows the error when the frequency region is [500,5000].

で示す. 周波数領域は, フィルタバンクごとに異なる. この図から以下のことがわかる.

- $\mathrm{GCFB_{dyn}}$ の場合最小誤差で, 線形の $\mathrm{GCFB_{lin}}$ よりも良い.
- $\mathrm{GTFB_{100}}$ は, $\mathrm{GCFB_{lin}}$ と同程度である.
- $\mathrm{GTFB_{*}}$ の帯域幅が狭まるにつれ, 誤差は大きくなる.
- $\mathrm{MFFB_{STR24}}$〜$\mathrm{MFFB_{STFT40}}$ は同程度の誤差で, $\mathrm{GTFB_{100}}$ と $\mathrm{GTFB_{050}}$ の中間的な値となる.
- $\mathrm{MFFB_{STFT120}}$ は, F_0 非依存の STRAIGHT スペクトルを基にした $\mathrm{MFFB_{STR120}}$ より格段に誤差が大きい.

表 1 に最小誤差を与える周波数領域を示す.

- どの場合でも周波数領域の下限周波数は 500Hz 以上で, それ以下とはならない.
- $\mathrm{GCFB_{dyn}}$, $\mathrm{GTFB_{*}}$ は, 上限周波数が 5000Hz で比較的広い領域となっている.
- $\mathrm{MFFB_{*}}$ は, 周波数領域の上限周波数が, 2000Hz〜3000Hz で比較的狭い.

一方, 音声のホルマント情報は, 2000Hz 以上にも存在する（たとえば, 母音/i/や/e/の第 2 ホルマント）. この情報を用いる方が, どのような音環境でも声道長を安定に推定できる

Table 1. Frequency region for minimum error

FB	Freq.Region	Error	FB	Freq.Region	Error
$GCFB_{dyn}$	[700,5000]	0.013	$GCFB_{lin}$	[500,3000]	0.015
$GTFB_{100}$	[600,5000]	0.017	$GTFB_{050}$	[800,5000]	0.028
$GTFB_{025}$	[800,5000]	0.033			
$MFFB_{STR24}$	[500,2000]	0.020	$MFFB_{STFT24}$	[600,2000]	0.020
$MFFB_{STR40}$	[600,2000]	0.020	$MFFB_{STFT40}$	[800,3000]	0.026
$MFFB_{STR120}$	[600,2000]	0.023	$MFFB_{STFT120}$	[800,3000]	0.045

と, 一般的には考えられる.

図 6 の＋マークに, 周波数領域を [500,5000] とした場合の誤差を示す.

- $GCFB_*$ $GTFB_*$ では, 最小を与える領域とほぼ同じで, 誤差も最小値に近い.
- $MFFB_*$ では, 誤差が大きく上昇する. すなわち, 2000Hz 以上の領域の情報はむしろ阻害要因となっていることがわかる.

3.4 考察

上記のような結果になった理由について考察する. 図 7 に, 実験に用いた 1 文の男性発話の中の母音/o/にラベル付けされている所の平均スペクトル (約 50ms 分) を, フィルタバンクごとに算出したものを示す. 図 7 上図から以下のことがわかる.

- $GCFB_{dyn}$ のスペクトルのダイナミックレンジが最も小さくなっている. これは, 音圧の低い所を強調し, 音圧の高い所を下げる圧縮特性 [12–14] があるためである. これには, ホルマント情報を強調する働きがある.
- $GCFB_{lin}$ は $GTFB_{100}$ と同程度のスペクトル形状であるが, ややなだらかである. これは, $GCFB_{lin}$ の帯域幅は約 1.5 倍程度広いことに起因する.
- $GTFB_{050}$ でスペクトルのリップルで示される resolved harmonics [*4] が大きくなる. 帯域幅の狭い $GTFB_{025}$ ではさらに大きくなる. これは, 基本周波数 F_0 の高調波成分が表れているためで, フィルタ帯域幅が狭まるほど鋭いピークとなって高い周波数まで目立つようになる.

図 7 下図からは以下のことがわかる.

[*4] このような F_0 の高調波成分によるスペクトル上のリップルは resolved harmonics と聴覚心理分野では呼ばれている. 個々のフィルタバンクで 1 つの高調波成分を分解してピークになるためで, フィルター帯域幅が F_0 に比べて大きくなるとピークは表われなくなる.

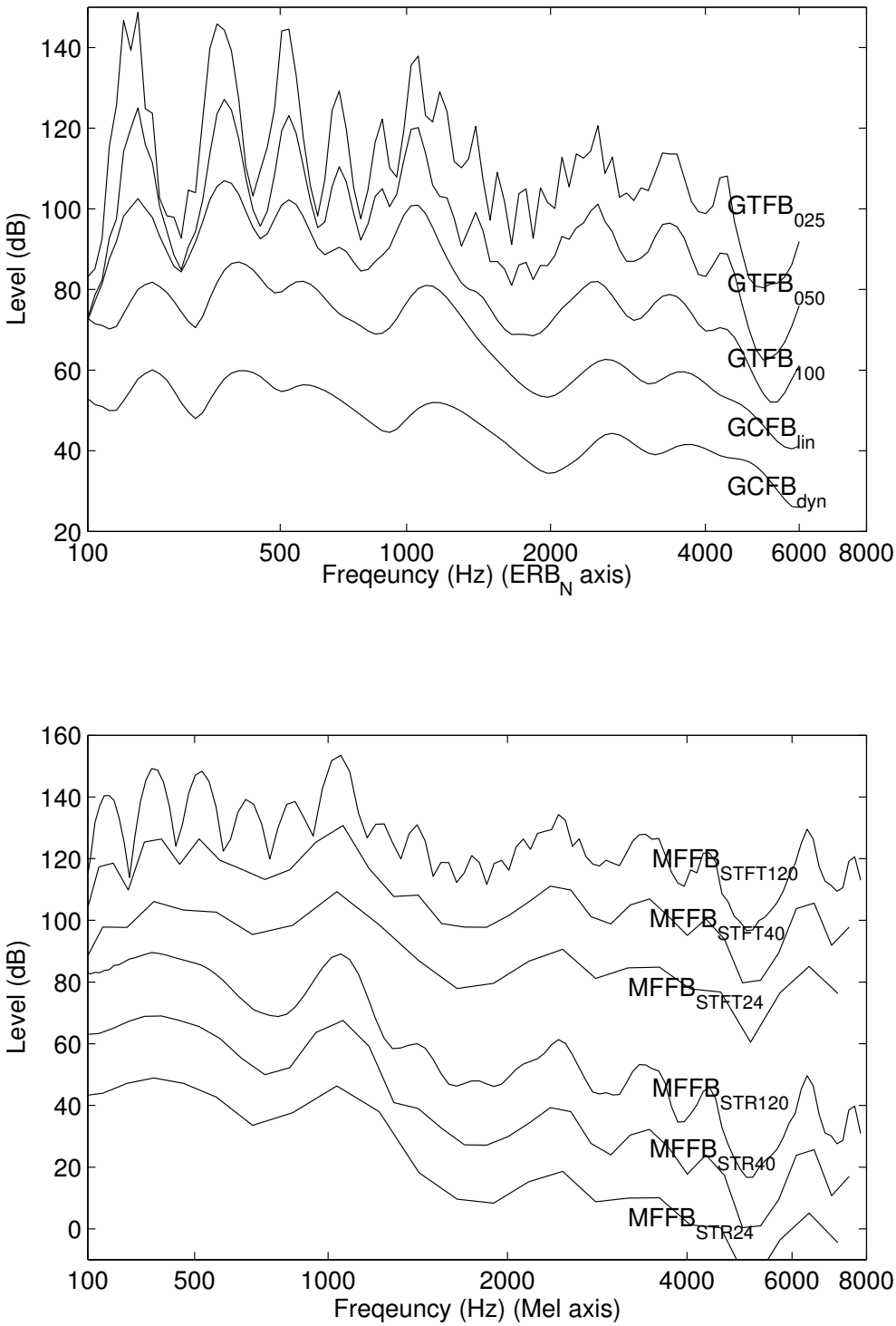


Fig. 7. Spectra of vowel /o/ derived by GCFB, GTFB, and MFFB.

- $\mathrm{MFFB}_{\mathrm{STFT120}}$ は GTFB_{025} 同様, resolved harmonics が目立っている.
- これに対し $\mathrm{MFFB}_{\mathrm{STFT24}}$, $\mathrm{MFFB}_{\mathrm{STFT40}}$ は目立たないが, フィルタ帯域幅が十分大きいためと考えられる.
- $\mathrm{MFFB}_{\mathrm{STR}*}$ で resolved harmonics が見られないのは, F_0 非依存の STRAIGHT スペクトルを基に計算しているためである.
- MFFB_* と GCFB_* や GTFB_* の比較から, オーバーコンプリートネスよりも, フィルタの帯域幅で性能が決まるようである.

4. 線形のウェーブレット変換を超えて

スケール変形に対して理論的に最適なはずの線形のウェーブレット変換よりも, 非線形の聴覚フィルタバンクの方が安定に声道長推定できることがわかった. このことは, この声道長推定の問題が単純なスケール変形だけでは表すことができないということを示している. ここでは, 聴覚末梢系の非線形特性を反映させているガンマチャープ聴覚フィルタバンク（$\mathrm{GCFB}_{\mathrm{dyn}}$）の非線形性について紹介し, 理論構築の議論の導入としたい.

4.1 周波数範囲

上記の結果は, 安定な声道長推定には 500Hz 以上の周波数領域を用いることが重要であるということを示している. 音声において声道長のスケール性だけが表出するのであれば, スケール変形に対し「透明」なはずのウェーブレット変換を用いれば十分で, 周波数を制約する条件は出ないはずである. しかしながら, 音声を駆動するための声帯振動があるため, その基本周波数 F_0 と高調波の影響がどうしても出てくる. また, 聴覚末梢系の特性に関しても, 図 4 に示したように, 定 Q 型フィルタとなるのは, 500Hz 以上である. これが偶然の一致か必然性があるのかは議論の余地がある.

4.2 非線形性の効用

図 6 の結果から, 線形フィルタバンク ($\mathrm{GCFB}_{\mathrm{lin}}$ ～$\mathrm{MFFB}_{\mathrm{STFT120}}$) に対して, 聴覚末梢系の非線形特性を反映させた $\mathrm{GCFB}_{\mathrm{dyn}}$ を用いた方が推定精度が良いことがわかる.

図 8 に, 心理物理実験によって求められたガンマチャープ聴覚フィルタの入力音圧に対するフィルタの振幅周波数特性の変化を示す [14]. まず, 音圧が高くなるにつれて, 中心周波数 (2000 Hz) におけるフィルタの利得が減少することがわかる. また, 中心周波数よりも離れた周波数 (例えば 1000 Hz 以下や 3000Hz 以上) では, レベル依存性がほとんど無いこともわかる. また, この特性を実現するフィルタにおいて, インパルス応答における瞬時周波数変化がほとんど無いことも生理学的にも知られており, ガンマチャープ聴覚フィルタにも反映させている.

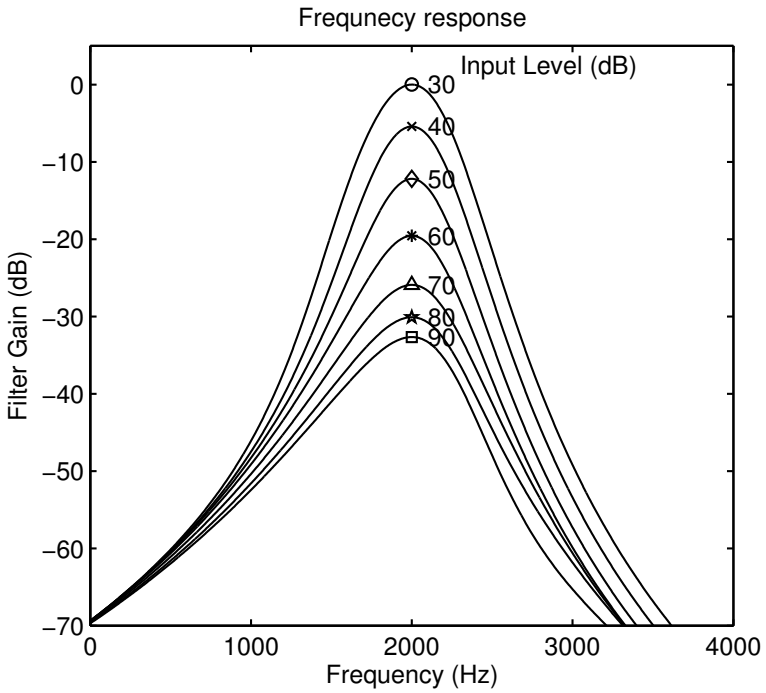


Fig. 8. Level dependent gain and filter shape when the input sound pressure level is varied between 30 and 80 dB.

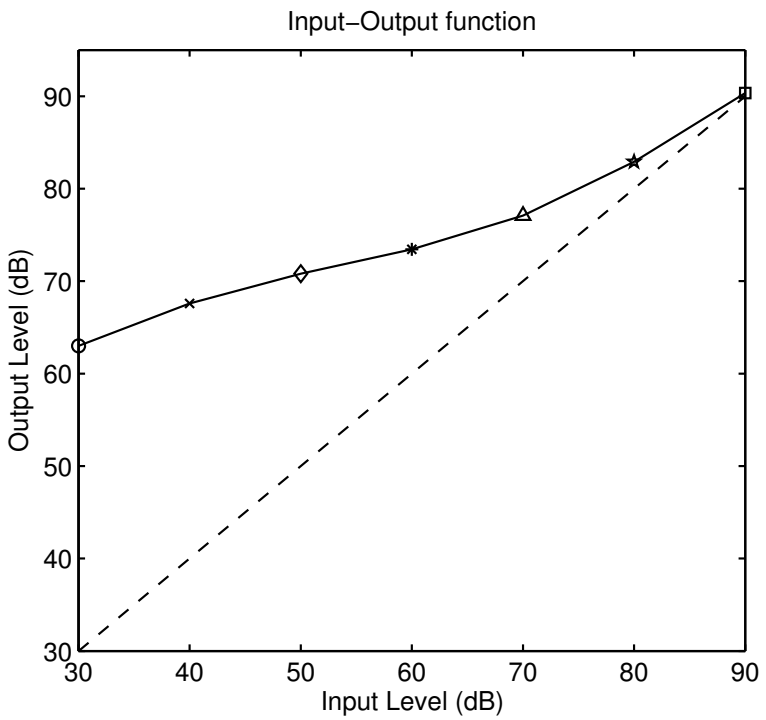


Fig. 9. Input-output function of auditory filter. The solid line shows compressive characteristics with growth rate of 0.2 $\sim$ 0.3 dB/dB.

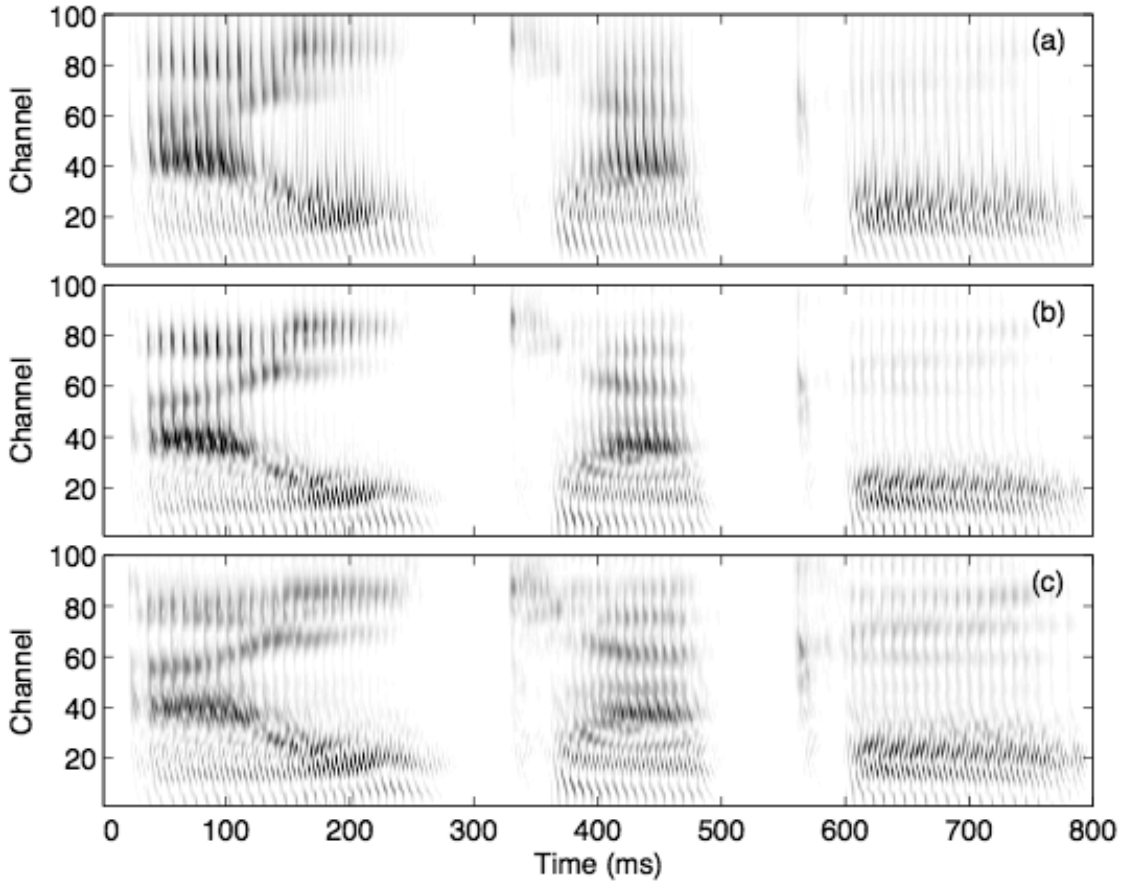


Fig. 10. Cochlear spectrograms, or cochleograms, for the Japanese word 'aikyaku,' plotted on a linear scale to reveal level differences: (a) $GCFB_{partial}$, (b)$GCFB_{lin}$, and (c) $GCFB_{dyn}$.

このフィルタを通した時の, 入力音圧と出力音圧の関係を図 9 に示す. 破線の対角線が入出力が 1:1 の線形の場合である. これに対し聴覚フィルタにおいては, 入力音圧の増加に対して出力音圧の増加の割合が少なく, おおよそ 0.2 〜0.3 dB/dB の割合である. この状況を指して, 圧縮特性と呼ならわされている. この他にも, フィルタを並べたフィルタバンクにおいては, 2 音抑圧[*5] 等の非線形特性も知られている .

ガンマチャープ聴覚フィルタバンク $GCFB_{dyn}$ で音声（発話「あいきゃく」）を分析した例を, 図 (c) に示す. 同図 (a),(b) は, 線形フィルタバンクの例である. 特に 600 〜 800 (ms) の所の 80ch 周辺におけるホルマント（声道音響管の共振特性）が強調されて表現されていることがわかる. また、40 〜 120 (ms) の 40ch 付近のホルマントは線形の場合に比べてむしろコントラストが小さくなっている. このことから, 聴覚フィルタにおける非

[*5] 中心周波数に正弦波を入れて観測した場合よりも, さらにその周辺の周波数に 2 つ目の正弦波を加えた場合の方が出力が減少する現象. 入力を増やしたにも関わらず出力が減少する.

線形性は, 音声の特徴を最も表す部分を平均的に強調するように働いていることがわかる. これが, 今回の声道長推定においても有効に働いたものと考えられる.

5. おわりに

初期聴覚系の計算理論の知見を声道長推定に応用した結果について述べた. 従来の声道長推定手法では, フィルタバンクの全帯域を用いて推定されることが多かったが, 適切な周波数領域を用いることが重要であることがわかった. また, 聴覚末梢系の非線形性を持った聴覚フィルタバンクの方が , ウェーブレット的な線形系よりも声道長推定を安定に行えることを示した. また, この非線形特性についても概略した. しかし, 線形のウェーブレット理論を拡張するには至っていない. ぜひ議論をしていただければ幸いである.

謝辞 本研究の一部は, 科学研究費補助金基盤 (A) 19200017 および (B)21300069 による支援を受けた. 聴知覚や理論に関しては, Roy D. Patterson 博士 (Cambridge 大 CNBH) との共同研究である. 声道長推定に関しては河原英紀教授 (和歌山大学システム工学部), 岡本恵里香氏との共同研究の成果である. ここに感謝する.

Appendix A. ガンマチャープ関数の導出

ガンマチャープ関数は Mellin 変換が張る空間の最小不確定性を持つ関数として求めることができる [20].

A.1 Mellin 変換

信号 $s(t), (t > 0)$ のメリン変換 [28] は

$$S(p) = \int_0^\infty s(t)t^{p-1}dt,$$

ここで p は複素変数である. 重要な特徴として,

$$\text{if} \;\; s(t) \Rightarrow S(p), \quad \text{then} \;\; s(at) \Rightarrow a^{-p}S(p),$$

が成立する. ここで矢印は変換を示し, a は実数の伸縮 (スケール) 係数である. すなわち, スケール変形に対し正規化した $S(p)$ の絶対値分布は変化せず, スケール不変表現となる.

A.2 演算子法と Mellin 変換

信号処理の時間周波数表現において, 量子力学で用いられてきた演算子法が表現の類似性から導入されている [29]. 時間演算子 $\mathcal{T} = t$, 時間領域における周波数演算子

$\mathcal{W} = -j(d/dt)$ を導入する． すると Cohen による「スケール演算子」は,

$$C = \frac{1}{2}(\mathcal{T}\mathcal{W} + \mathcal{W}\mathcal{T}) = \mathcal{T}\mathcal{W} - \frac{1}{2}j,$$

と表される. これは, 量子力学におけるアフィン変数を表現する演算子として既に知られている [30]. この演算子に対応する「スケール変換」[29] は

$$D(c) = \frac{1}{\sqrt{2\pi}} \int_0^\infty s(t) t^{-jc-1/2} dt,$$

で, メリン変換において $p = -jc + 1/2$ と設定したものである. この式の適用範囲を広げるためバイアス項の実数 c_0 と μ を導入して,

$$p = -j(c - c_0) + (\mu + 1/2).$$

と拡張する. 対応するメリン演算子は次式となる.

$$C_m = \mathcal{T}\mathcal{W} + \{c_0 + j(\mu - 1/2)\}.$$

我々の関心事は, 聴覚末梢系のフィルタバンク表現である. そこで,「周波数シフト項」ω_0 を各々のフィルタを特定するために導入する. すると演算子は以下のように変形できる.

$$C_a = \mathcal{T}(\mathcal{W} - \omega_0) + \{c_0 + j(\mu - 1/2)\}.$$

この周波数シフト項 ω_0 は, 本文の 2.2.3 節で述べた寸法形状イメージ (SSI) における周波数正規化の処理により, 完全に取り除くことができる. このため, メリン変換の枠組みからは外れない. 時間とこの演算子の交換子は以下となる.

$$[\mathcal{T}, C_a] = [\mathcal{T}, C_m] = [\mathcal{T}, C] = j\mathcal{T}.$$

交換子が 0 とならないので, 時間とこの演算子の表すメリン空間の値は独立に計測できない. この時の不確定性の関係は, 以下で表される.

$$\Delta t \cdot \Delta c_a \geq \frac{1}{2} | < [\mathcal{T}, C_a] > | = \frac{1}{2} | < j\mathcal{T} > | = \frac{< t >}{2}.$$

ここで, Δ. は 標準偏差, $< . >$ は平均を表す. 関数の時間平均値の 1/2 以上という条件となる. 次節で, この最小不確定性を満たす関数を導出する.

なお, 良く知られている時間-周波数空間における不確定性の関係は次式である [29].

$$\Delta t \cdot \Delta \omega \geq \frac{1}{2} | < [\mathcal{T}, \mathcal{W}] > | = \frac{1}{2} | < j > | = \frac{1}{2}.$$

この最小値を取るのは, もちろん Gabor 関数 [31] である.

A.3 最小不確定性を満たす関数

演算子が定義できると, 最小不確定性を持つ関数は固有値問題を解くことによって求められる. C_a や C_m は, $\mu = 0$ の場合以外 Hermitian ではない. しかし, 平均値を引いた $(C_a - < C_a >)$ は Hermitian となるため実固有値が求まる. この演算子と時間とで張る空間における最小不確定性を持つ関数は, 以下の固有値問題の解として得られる.

$$(C_a - < C_a >)\, s(t) = \lambda(\mathcal{T} - < t >)\, s(t).$$

ここで

$$\lambda = \frac{< [\mathcal{T}, C_a] >}{2(\Delta\mathcal{T})^2} = \frac{j < t >}{2(\Delta t)^2}\ ,$$

である. 固有値問題の式を展開すると以下のようになる.

$$t(-j\frac{d}{dt})\, s(t) - (\omega_0 + j\alpha_1)t\ s(t) + (-c_1 + j\alpha_2)\ s(t) = 0.$$

ここで, $\alpha_1 =< t > /2(\Delta t)^2$, $\alpha_2 = \mu - 1/2 - \mathrm{Im} < c_a > + < t >^2 /2(\Delta t)^2$, $c_1 = \mathrm{Re} < c_a > - c_0$ で, Re., Im. はそれぞれ実部, 虚部を示す. この解は, 以下のように求まる.

$$\begin{aligned} s(t) &= a\, t^{\alpha_2 + jc_1}\ \exp(-\alpha_1 t + j\omega_0 t), \\ &= a\, t^{\alpha_2}\ \exp(-\alpha_1 t)\ \exp(j\omega_0 t + jc_1 \ln t). \end{aligned}$$

ここで a は定数で, ln は自然対数である.

この包絡線 $t^{\alpha_2}\exp(-\alpha_1 t)$ はガンマ分布関数 $\gamma(t)$ である. 搬送波は $\exp(j\omega_0 t + jc_1 \ln t)$ で表される. 搬送波の偏角を時間微分すると, 瞬時周波数 f_i が得られる.

$$f_i = \frac{1}{2\pi}(\omega_0 + \frac{c_1}{t}).$$

これは, 時間的に瞬時周波数が変化することを示しており, 音として再生するとチャープ音である. そこで, この関数を「ガンマチャープ (gammachirp)」と命名した. ここで, $c_1 = 0$ とすると, 搬送波は一定周波数の正弦波となり,「ガンマトーン (gammatone)」関数となる. すなわち, ガンマチャープは, 元々実験式として与えられたガンマトーンを特殊解として持つ, 自然な形の拡張となっていることがわかる.

Appendix B. 声道長推定手法

ここでは, 文献 [24–26] における声道長推定法について簡単に紹介する.

B.1 スペクトル距離に基づく声道長比の推定

同じ文章を発話した話者 A, 話者 B の音声はフィルタバンクによって分析され, 平滑化されたスペクトログラム $P_A(\tilde{f},t)$ と $P_B(\tilde{f},t)$ が求められる. ここで $\tilde{f}$ はワープ周波数で, フィルタバンクにより ERB 周波数 f_{ERB} あるいは mel 周波数 f_{mel} のいずれかを表す. また,t は分析時刻（分析窓の中心時刻) を表す. 二つの音声の音素の出現位置は異なっているため, まず,B のスペクトログラムの時間軸を A と合うように変形する. 変形したスペクトログラムは $P_{Bn}(\tilde{f},t)$ で表される. A と B の声道長を一致させるために, $P_{Bn}(\tilde{f},t)$ を, 元の周波数軸上で線形に r 倍伸縮させる [*6]. 周波数は r 倍されて rf となり, これをワープ周波数に変換すると $r\tilde{f}$ となる. したがって, 変形されたスペクトルは $P_{Bn}(r\tilde{f},t)$ で表される. 分析時刻 t のとき, dB 上でのスペクトル距離は, 実効値 (rms) として以下の式で表される.

$$D_{dB}(t,r) = \sqrt{\frac{D_P}{\tilde{f}_H - \tilde{f}_L}} \quad ,$$

$$D_P = \int_{\tilde{f}_L}^{\tilde{f}_H} \left(10\log_{10}\frac{P_A(\tilde{f},t)}{\bar{P}_A(t)} - 10\log_{10}\frac{P_{Bn}(r\tilde{f},t)}{\bar{P}_{Bn}(t)}\right)^2 d\tilde{f}.$$

$\tilde{f}_L$ と $\tilde{f}_H$ は周波数帯域の下限周波数と上限周波数, $\bar{P}_A(t)$ と $\bar{P}_{Bn}(t)$ は周波数の平均値である. 最適な声道長比の推定値 r は, 文章全体の距離 D_{dB}^{total} を最小化する値である.

$$r = \arg\min_r (D_{dB}^{total}(r)).$$

ここで, $D_{dB}^{total}(r)$ は, フレームごとのスペクトル距離 $D_{dB}(t,r)$ から以下のように計算される.

$$D_{dB}^{total}(r) = \sqrt{\frac{1}{\|V\|}\int_{t\in V} D_{dB}^2(t,r)dt,},$$

V は文章の有声音の部分である.

B.2 声道長比の推定手法

図 11 に, その計算方法の概略を示す. m 番目と n 番目の話者の VTL を l_m と l_n, とすると, 声道長比は $r_{m,n} = l_m/l_n$ で表される. 対数をとることで, 差分で表すことができる.

$$\log(r_{m,n}) = \log(l_m) - \log(l_n).$$

[*6] 単純な音響管近似で伸縮が行われていると仮定している. 実際の関係はもう少し複雑で 1 次関数以上が必要な可能性もある. しかし, この分野で十分な検討はまだ行われていない. そこでは, データのばらつきとフィッティングの良さのトレードオフや, 説明変数の少なさ（オッカムの剃刀）も含む AIC 規準等を適用して考えるべきであろう.

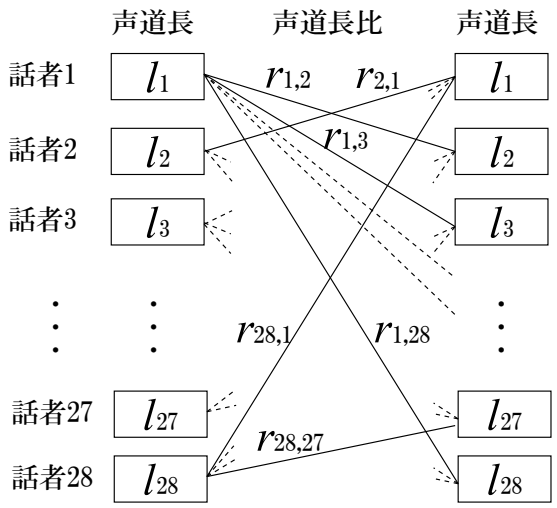


Fig. 11. *Vocal tract length of 28 speakers* $l_n\{n \mid l \leq n \leq 28\}$ *and vocal tract length ratio* $r_{m,n}\{m, n \mid m \neq n, 1 \leq m \leq 28, 1 \leq n \leq 28\}$

声道長比を求める際, フィルタバンク上で伸縮が非対称となるため, 単純な組み合わせではなく逆順も考慮した (cf. [3]). 28 名から 2 名ずつ並べる順列は 756($=_{28}P_{27}$) 通りである.

$$\begin{bmatrix} \log(r_{1,2}) \\ \log(r_{1,3}) \\ \vdots \\ \log(r_{27,28}) \\ \log(r_{2,1}) \\ \log(r_{3,1}) \\ \vdots \\ \log(r_{28,27}) \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & -1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ -1 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix} \begin{bmatrix} \log(l_1) \\ \log(l_2) \\ \vdots \\ \log(l_{28}) \end{bmatrix}$$

ここで, 最後の行は, 声道長の幾何学的平均値を正規化する制約で, 行列を正則化するために導入している.

上式の左辺を $\boldsymbol{r}_{log}(= \log(\boldsymbol{r}))$, 右辺の係数行列を H, 声道長の対数のベクトルを $\boldsymbol{l}_{log}(= \log(\boldsymbol{l}))$ と書き直す.

$$\boldsymbol{r}_{log} = H\boldsymbol{l}_{log}.$$

ここで回帰分析を行い, 声道長 $\hat{\boldsymbol{l}} = [\hat{l}_1, \hat{l}_2, ..., \hat{l}_{28}]$ の推定値を最小 2 乗法により計算する.

$$\hat{\boldsymbol{l}}_{log} = (H^T H)^{-1} H^T \boldsymbol{r}_{log},$$
$$\hat{\boldsymbol{l}} = [\hat{l}_1, \hat{l}_2, ..., \hat{l}_{28}]^T = \exp(\hat{\boldsymbol{l}}_{log}).$$

また, $\hat{\boldsymbol{l}}_{log}$ から声道長比の推定値 $\hat{\boldsymbol{r}}$ も計算できる.

$$\hat{\boldsymbol{r}} = \exp(H\hat{\boldsymbol{l}}_{log}).$$

この値と, スペクトル距離から計算した声道長比 $\boldsymbol{r}$ との間のユークリッド距離 d_{est} (rms 値) で, 推定誤差を評価できる.

$$d_{est} = \|\boldsymbol{r} - \hat{\boldsymbol{r}}\| \simeq \sigma.$$

d_{est} は, 恒等写像線 ($\hat{\boldsymbol{r}} = \boldsymbol{r}$) を中心とした標準偏差 σ とほぼ同じである. したがって σ が小さいほど, 組み合わせの条件の相違による変動が小さく, 安定に推定できていると考えることができる.

参考文献

[1] Wakita, H., "Normalization of vowels by vocal-tract length and its application to vowel identification, " IEEE Trans. Acoust., Speech, Signal Processing, ASSP-32, pp. 183–192,1977.

[2] 浅香, 西田, 赤桐, 西村, 入野, 河原, "声道長の正規化に基づく簡易モーフィング音声の品質改良について," 信学会音声研究会, SP2009-34, 109(99), pp.63–68, 2009.

[3] Okamoto, E, Irino, T., Nisimura, R, Kawahara, H., "Evaluation of voice morphing using vocal tract length normalization based on auditory filterbank," Proc. NCSP'11, pp.187–190, 2011.

[4] Irino, T. and Patterson, R.D., "Segregation information about the size and shape of vocal tract using a time-domain auditory model : The stablished wavelet-Mellin transform," Speech Communication, 36(3-4), pp.181–203, 2002.

[5] 入野,"音源の形状情報と寸法情報を分離する聴覚でのイメージング," 日本音響学会誌, 56 巻 7 号, pp. 505–508, July 2000.

[6] 入野, "初期聴覚系におけるスケール理論," 音響学会春季研究発表会講演論文集 I, pp.511–514, 2003.

[7] Patterson, R.D. . "Auditory images: How complex sounds are represented in the auditory system," J. Acoust. Soc. Japan (E), 21, pp. 183–190, ,2000 (入野抄訳 , "聴覚イメージ：複雑な音が聴 覚システムでいかに表現されるか," 日本音響学会誌, 56 巻 7 号, pp. 503-504, July 2000.)

[8] Smith, D.R., Patterson, R.D. , Turner, R., Kawahara, H. and Irino, T., "The processing and perception of size information in speech sounds," J. Acoust. Soc. Am, 117(1), pp.305–318, 2005.

[9] 青木, 入野, パターソン, 河原, "スケール変形した有声／ 無声単語の寸法弁別と音韻認識に関する検討," 日本音響学会聴覚研究会資料, H-2008-89, Vol. 38, No. 5, pp.507–512, 2008.

[10] Irino, T., Aoki, Y., Kawahara, H., and Patterson, R.D., “Size Perception for acoustically scaled sounds of naturally pronounced and whispered words,” in “Neurophysiological Bases of Auditory Perception,” Enrique A. Lopez-Poveda, Alan R. Palmer, and Ray Meddis (Eds.), pp.235–243, Springer, LaVergne, TN USA, 644p., Apr., 2010.

[11] Moore, B. C. J., “Psychology of Hearing (5th ed),” Academic Press, London, 2003. (大串訳「聴覚心理学概論（第３版)」誠信書房）

[12] 入野, “はじめての聴覚フィルタ,” 音響学会誌, 66（10）, pp. 506-512, 2010.

[13] 入野, “はじめての聴覚フィルタ - 心理物理実験デモで学ぶ聴覚フィルタ特性 -,” 秋季音講論, pp.1347 –1348, 2010.

[14] 入野, “聴覚フィルタの心理物理実験とモデル（第４章)” , in “聴覚モデル” (森, 香田編著), p.233, pp.101–128, コロナ社, 東京 , 2011.

[15] de Boer, E. and de Jongh, H.R., “On cochlear encoding: Potentialities and limitations of the reverse-correlation technique,” J. Acoust. Soc. Am., 63, pp. 115–135, 1978.

[16] Irino, T. and Patterson, R. D., “A dynamic compressive gammachirp auditory filterbank,” IEEE Trans. Audio, Speech, Lang. Process., 14(6), pp. 2222–2232, Nov. 2006.

[17] Daubechies, I., “The wavelet transform, time-frequency localization and signal analysis ,” IEEE Trans. Information Theory, Vol. 36 (5), pp. 961–1005, 1990.

[18] Patterson, R. D., Allerhand, M. and Giguére, C., “Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform, ” J. Acoust. Soc. Amer., vol. 98, pp. 1890–1894,1995.

[19] Patterson, R.D., Unoki, M., Irino, T.,“Extending the domain of center frequencies for the compressive gammachirp auditory filter,” J. Acoust. Soc. Amer., vol. 114 (3), pp. 1529–1542, 2003.

[20] Toshio Irino and Roy D. Patterson “A time-domain, level-dependent auditory filter: the gammachirp,” J. Acoust. Soc. Am., 101 (1), pp.412–419, 1997.

[21] Lyon, R.F, Ponte, J., and Chechik, G., “Sparse coding of auditory features for machine hearing, ” ICASSP2011, 2011.

[22] Versnel, H. and Shamma S.A., “Spectral-ripple representation of steady-state vowels in primary auditory cortex, ” J. Acoust. Soc. Am., 103(5), pp. 2502–2514, 1998.

[23] 津崎, 入野,“シミュレータによる内部表現と特徴量（第７章）,” in “聴覚モデル” (森, 香田編著) p.233, pp.195–229, コロナ社, 東京 , 2011.

[24] 岡本, 入野, 西村, 河原, “聴覚フィルタバンクを用いた声道長比推定, ” 電子情報通信学会 音声研究会, 電子情報通信学会技術研究報告, Vol.111, No.153, SP2011-43,

pp.11–16, 2011 年 7 月.

[25] Okamoto, E., Irino, T., Nisimura, R., Kawahara, H., “Auditory filterbank improves voice morphing,” in Proc. Interspeech 2011, Tue-Ses2-P1, Florence, Italy, 27-31 Aug., 2011.

[26] 岡本, 入野, 西村, 河原, “聴覚フィルタバンクを用いた声道長推定法の比較, ” 日本音響学会：春季研究発表会講演論文集, 3-Q-15, 2011 年 9 月.

[27] http://labrosa.ee.columbia.edu/matlab/rastamat/, HTK のメル尺度を選択

[28] Titchmarsh, E. C., “Introduction to the Theory of Fourier Integrals,” Oxford U.P., London, 2nd ed, 1948.

[29] Cohen, L. “The scale representation,” IEEE Trans. Signal Process. 41, pp. 3275–3292, 1993.

[30] Klauder, J. R., “Path integrals for affine variables,” in Functional Integration: Theory and Applications, edited by Antoine, J. P. and Tirapgui, E. , Plenum, New York, 1980.

[31] Gabor, D., “Theory of communication,” J. IEE (London), 93, pp. 429–457, 1946.

脳波によるヒトの状態推定のための ウェーブレット手法の応用

井上 勝裕 * 山口 朋成 * 前田 誠 * 藤尾 光彦 *
* 九州工業大学

概要. 本稿では，ウェーブレット変換手法や，その非線形バージョンとも考えることができるモルフォロジカルフィルタを用いた多重解像度解析手法を終夜睡眠脳波の解析や，BCI (Brain Computer Interface) における誘発電位の抽出に応用した事例を紹介する．

Application of Wavelet Analysis to estimate human state using EEG

Katsuhiro Inoue* Tomonari Yamaguchi* Makoto Maeda*
Mitsuhiko Fujio*
*Kyushu Institute of Technology

Abstract. There are wavelet method and morphological method as tool of time-series analysis. Morphological method is thought to be a non-linear version of wavelet method. In this paper, it introduces the case where these method are applied to the analysis of human sleep EEG(electroencephalogram) and feature extraction of evoked potential in BCI (Brain Computer interface).

1. はじめに

脳波 (EEG:Electroencephalogram) が脳内の状態によって変化することは, 脳波の発見者である Berger によって 1929 年に報告され, 1933 年に Adrian によって追試・確認が行われて以来, 病的疾患の診断のみならず，脳のメカニズムを解明する上で興味ある研究対象として, 多くの研究者によって研究が行われてきている.

睡眠脳波ステージは，脳波，眼球運動図 (EOG:Electro-oculogram)，筋電位図 (EMG:Electromyogram) からの情報をもとに, 判読されるが, その一晩の遷移状況である睡眠パタンは, 騒音等の睡眠環境の異常による外的要因や, 不規則な睡眠時間, 極度の疲労, 病的疾患等の内的要因により乱される. 従って, これを計測することによって, 被験者の健康状態を知ったり, 外的環境の評価が可能となることから, その解析手法について多くの研究がなされている. また, 最近では, 選択反応時の事象関連電位の変動や, 外的光刺激の周波数に応じた脳波周波数変動, 動作想像時の脳波変動を検出し, コンピュータや装置のインタフェースとして利用しようという BCI(Brain Computer Interface) あるいは

BMI(Brain Machine Interface) と呼ばれる研究もさかんに行われている. これは, 手足の不自由なハンディキャップユーザのための装置として, 研究が進められているが, 脳波マウスといった正常成人の通常使用を意識した研究も始まっている.

我々も, これまで, 脳波波形パタンを識別する手法として, 統計的パタン認識手法や波形形状認識手法, 離散ウェーブレット変換を利用した睡眠脳波ステージ自動判定システムを構築し, 有効な結果を得ており [2], BCI に関しても, 右手・左手動作想像時の脳波変動を検出し, AIBO をリアルタイムに操作できるシステムを構築してきている [9].

我々は, これまで, 脳波から視察では抽出が困難な睡眠に関するより詳細な情報を抽出することを目的に, ウェーブレット解析手法を適用する研究や, BCI システムにおいて事象関連電位や誘発電位の検出を目的に, 非線形作用素を用いるモルフォロジカルフィルタを使用した多重解像度解析手法に関する研究を進めてきており, 本稿では, これらの手法について紹介する.

2. 脳波概説

2.1 睡眠脳波

脳波は非常に複雑な波形パタンを有しているが, 正常安静時の脳波は比較的規則正しいリズムを有しており, α 波, 様々な周波数成分を含む低電位波 (以後 L 波と略す), 高振幅徐波などの背景脳波と, それらに重畳して出現する睡眠紡錘波, K 複合波, 頭頂部鋭波, 運動覚醒波, 鋸歯状波などがある. これらの波形パタンのうち, α 波, L 波, 高振幅徐波, 睡眠紡錘波, K 複合波の一例を Fig. 1 に示す. これらの脳波波形パタンの出現率や出現系列と眼球運動図からの急速眼球運動 (REMs: Rapid Eye Movements) 情報, 筋電図からの筋電位レベルをもとに, 睡眠脳波ステージは定義され, 一般に, 覚醒段階 (Stage W) と睡眠段階（Stage 1～4, Stage REM) の 6 段階に分けられている [3, 7].

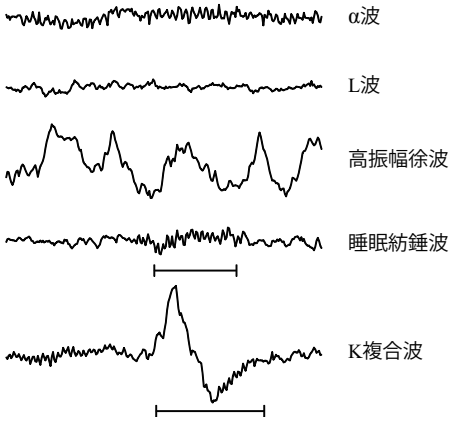


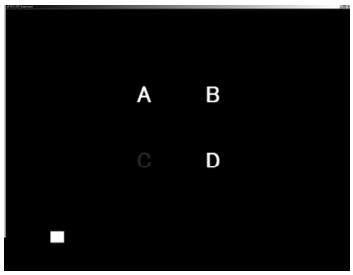
Fig. 1. Typical human sleep EEG patterns

ここでの目的は, この離散的な睡眠脳波ステージの識別にとどまらず, より詳細な睡眠脳波ステージに関する連続的な情報を得ることであり, その妥当性を検討する情報として, 覚醒中, 睡眠中を問わず生体の内部環境を調整する系として働く自律神経系リズムの情報を利用している. この自律神経リズムは, 心電位図の R-R 間隔変動を解析して得られ, LF(Low Frequency:0.04Hz~0.15Hz) 成分は交感神経活動の, HF(High Frequency:0.15Hz~0.4Hz) 成分は副交感神経活動のレベルを示すことが知られている [5].

2.2 光点滅刺激注視時の脳波変動

覚醒時においても, 種々の要因によって脳波は変動する. 認識, 記憶, 行動, 注意などの精神活動・認知活動がその要因となるが, 光点滅刺激注視時には, 視覚誘発電位（VEP:Visual Evoked Potential）が出現することが知られている. この現象を利用した SS-VEP（Steady State VEP) による BCI システムの開発が行われているが, 光てんかんの危険性を伴うこともあり, ここでは, 符号化された系列に基づく光点滅刺激を同時に与え, その応答脳波からどの光点滅刺激を注視しているかを判別する実験について説明する.

4 種類の文字がディスプレイ上に図に示すように表示され, 被験者はそのうち一文字を注視するよう指示される (Fig. 2). 各々の文字は, 1.5 秒周期の異なるパタン（Fig. 2 の下表参照）で点滅させる. 1 サイクルは 15 ビットから構成され, 光点灯時間 50msec, 光消灯時間 50msec である. 被験者とモニタ（26 インチ LCD ディスプレイ）間の距離は 1m である. 5 人の被験者がこの実験に参加し, 1 日に 2640 サイクル（120 サイクル ×22set）の実験で, 各被験者 3 日間実施した. なお, 脳波は, 17 電極 (Fig. 3) から採取し, サンプリング周波数は 512Hz とした.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	1	1	0	1	1	0	0	1	0	1	0	0	0	0	1
B	0	1	1	0	1	0	1	1	1	1	0	0	0	1	0
C	1	0	1	1	0	0	1	0	1	0	0	0	0	1	1
D	1	1	0	1	0	1	1	1	1	0	0	0	1	0	0

1: stimulus, 0: non-stimulus

Fig. 2. Stimulus patterns

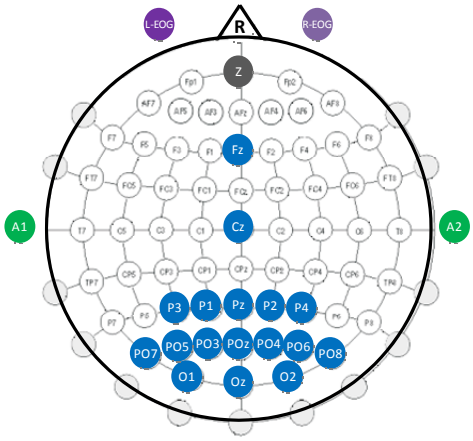


Fig. 3. Electrode positions

3. 解析手法

3.1 ウェーブレット変換

式 (3.1) で示される連続ウェーブレット変換はパラメタ a(スケーリングパラメータ) と b(シフトパラメタ) を有する積分変換である. この変換によって, 時間-周波数情報を得ることができる. ここでは, 周波数情報と関連付けの容易な式 (3.2) に示すガボール関数を使用する.

$$(W_\psi f)(b,a) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t)\psi\left(\frac{t-b}{a}\right) dt \tag{3.1}$$

$$\psi(t) = \frac{1}{2\sqrt{\pi}\sigma} \exp(-t^2/\sigma^2) \exp(jt) \tag{3.2}$$

連続ウェーブレット変換は, 式 (3.2) による (減衰係数) の値 σ を調整することによって, 周波数と時間の分解能を調整できる.

通常のウェーブレット変換は, 窓幅を周波数に応じて自動的に変化させ，時間-周波数解析を行うものであるが, 前節で説明したように, 睡眠脳波の場合，α 波や睡眠紡錘波のように比較的周波数の高い波は, 連続した波形として定義されているのに対し, 高振幅徐波や K-複合波のような低い周波数の波は単一波として定義されている. このことは, 通常のウェーブレット変換が処理窓内の波形数を同じになるよう窓幅を調整することに対し, 睡眠脳波波形処理用に他の窓幅選択法が存在することを示唆している.

そこで, ここでは, 理論的な直交性は失われるが, 次のような σ 調整式を導入した修正ウェーブレット変換法を考えることにする.

$$\sigma = \sigma_{\min} + \frac{K}{1 + \exp[-K_T(1/a - 1/a_c)]} \tag{3.3}$$

ただし, $\sigma_{\min}, K, K_T, a_c$ は定数である.

ここでは，各パラメータの値を, $\sigma_{\min} = 0.8$, $K = 10$, $K_T = 2.5$, $a_c = 0.1$ と設定しているが, この場合, 低周波領域 $(1/a \ll 1/a_c)$ や, 高周波領域 $(1/a \gg 1/a_c)$ において σ の値は以下のようになる.

$$\begin{cases} a \gg a_c & : \quad \sigma \approx \sigma_{min} + K/\exp[K_T/a_c] & = 0.8 \\ a = a_c & : \quad \sigma = \sigma_{min} + K/2 & = 5.8 \\ a \ll a_c & : \quad \sigma \approx \sigma_{min} + K & = 10.8 \end{cases}$$

この設定を用いた場合の EEG 波形と対応する修正ウェーブレット関数を Fig. 4 に示すが, 徐波帯域では単一波, 睡眠紡錘波帯域では 6 波程度の窓幅を持つよう設定していることがわかる.

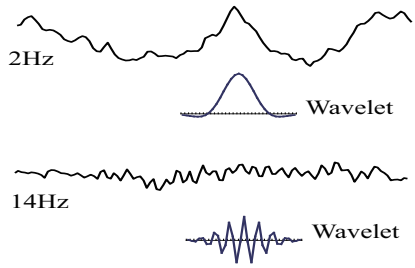


Fig. 4. EEG waves and the corresponding wavelets

1 エポック（20 秒）の脳波データに対する通常のウェーブレット変換結果を Fig. 5 に, 修正ウェーブレット変換結果を Fig. 6 に示す．両者を比較すると, 修正ウェーブレット変換を用いることによって, β 波帯域における周波数分解能が改善されていることがわかる. Fig. 7 は終夜のウェーブレット変換の結果を示している. Fig. 6 に示されている 1 エポック (20 秒) の平均化値をプロットしたものである. なお, ここでは, サンプリング周波数 500Hz で採取した EEG 信号 (C3-A2) を分析に用いている.

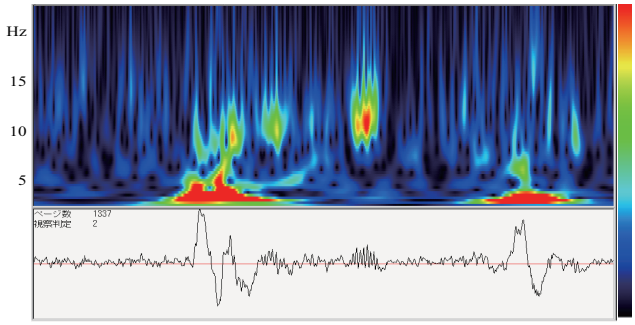


Fig. 5. Result of ordinary wavelet transform.

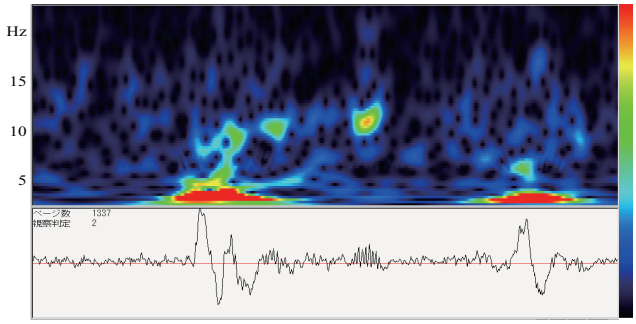


Fig. 6. Result of modified wavelet transform.

3.2 モルフォロジカルフィルタを用いた多重解像度解析

本節では, 非線形演算による信号処理を行うモルフォロジカルフィルタを利用した多重解像度解析手法について説明する.

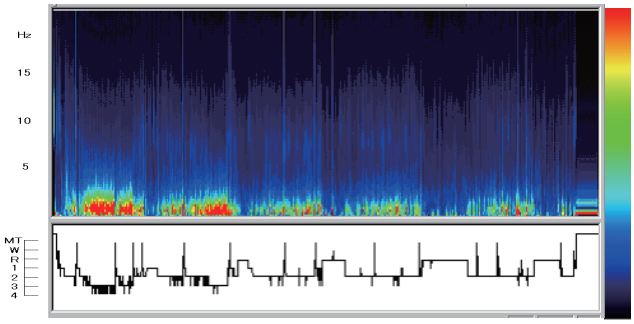


Fig. 7. Modified wavelet transform result of one night.

3.2.1 モルフォロジカルフィルタ

モルフォロジーの基本演算として, Minkowski 和 ⊕, Minkowski 差 ⊖ があり, その定義は以下の通りである.

$$[f \oplus g](t) = \max_{\substack{t+u \in F \\ u \in G}} \{f(t-u) + g(u)\} \tag{3.4}$$

$$[f \ominus g](t) = \min_{u \in G} \{f(t-u) - g(u)\} \tag{3.5}$$

ここで, $f(t)$ は処理対象である原波形, $g(t)$ はフィルタの特性を決定する構造関数である. また, F は処理対象の定義域, G は構造関数の定義域を表し, 定義域外で信号は負の無限大の値を持つものとする. これらを組み合わせた, Erosion, Dilation, Opening, Closing の 4 つをモルフォロジカルフィルタの基本演算とする. これは,

$$\text{Erosion} \quad (f \ominus g^s)(t) = \min_{u \in G} \{f(t+u) - g(u)\} \tag{3.6}$$

$$\text{Dilation} \quad (f \oplus g^s)(t) = \max_{\substack{t+u \in F \\ u \in G}} \{f(t+u) + g(u)\} \tag{3.7}$$

$$\text{Opening} \quad f_g(t) = [(f \ominus g^s) \oplus g](t) \tag{3.8}$$

$$\text{Closing} \quad f^g(t) = [(f \oplus g^s) \ominus g](t) \tag{3.9}$$

で定義される. ここで, $g^s(t) = g(-t)$ である. Erosion は構造関数の形状で負方向に膨らませる効果があり, Dilation は構造関数の形状で正方向に膨らませる効果がある. それらを続けて処理した Opening は, 構造関数を負方向から押し当てたとき, 構造関数が処理対象に進入できる部分は残し, 進入できない部分を除去する特性がある. Closing は, Opening とはグラフ的に上下逆で, 構造関数を正方向から押し当て, 負方向の突起形状を除去する特性

がある．さらに, 正負の突起形状を除去するものとして, Opening と Closing を組み合わせた Open-Closing フィルタを定義できる [16]．これは構造関数に比べて小さな凸部を削る平滑化であり, 一種の非線形な低域通過フィルタである．構造要素のスケールを大きくすると, その通過帯域はそれだけ狭くなる．また, 高域通過フィルタは, 低域通過フィルタの出力結果を元の信号から引き去ることで構成される．

(3.10) $$\text{low pass} : \psi_g(t) = \left(f_g\right)^g(t)$$

(3.11) $$\text{high pass} : \omega_g(t) = f(t) - \psi_g(t)$$

(3.12) $$\text{band pass} : p_{g_m,g_n}(t) = \psi_{g_m}(t) - \psi_{g_n}(t)$$

(3.13) $$\text{band stop} : c_{g_m,g_n}(t) = f(t) - \psi_{g_m}(t) + \psi_{g_n}(t)$$

モルフォロジカルフィルタを使用するために構造関数を決める必要があるが, 例えば, 定数 c とウィンドウの幅 $2n+1$ をパラメータとする Haar 構造関数は以下の式で定義される．

(3.14) $$g(t) = \begin{cases} c & (-n \le t \le n) \\ -\infty & (otherwise) \end{cases}$$

パルス性の成分が重畳した信号に対してモルフォロジカルフィルタを施したときの演算結果を Fig. 8 に示す．

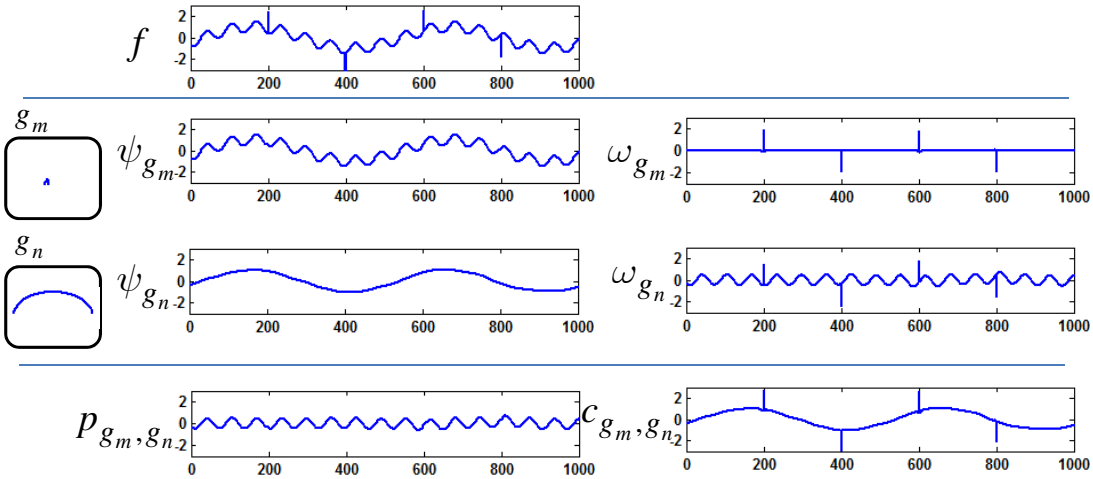


Fig. 8. Morphological filter

3.2.2 パタンスペクトル

パタンスペクトルは連続的に構造要素の異なるモルフォロジカルフィルタを適用す形で，構成される．まず，ベースとなる構造関数 B を選択し，モルフォロジカルローパスフィルタとハイパスフィルタ処理を行う．次に，同じ操作を，拡大した構造関数を用いて，低周波成分の処理を行う．最終的に、設定したレベル J まで，この処理を繰り返す．

そして、それぞれのレベルの高周波成分の加算を行うことによって得られる.

(3.15) $$P(f, B, j) = \sum_t |\omega_{jB}(t)|$$

$$(jB = B \oplus B \oplus \cdots \oplus B \quad (j-1 \text{ times } \oplus \text{operations}), \quad B \subset 2B \subset \cdots \subset JB)$$

ただし，$\omega_{jB}(t)$ は，レベル j の高周波成分である.

3.2.3 非線形多重解像度解析

Open-Closing フィルタは, 低域通過フィルタであり, これと原波形との差分をとることにより高域通過フィルタも得られる. さらに, 通過できる帯域は, 構造関数の形状に依存するため, 基準となる構造関数のスケールを変えて出力波形を出すことは, 多重解像度解析を行うことを意味する [17]. 信号の集合を V_j, W_j とする. V_j はレベル j の低周波成分, W_j はレベル j の高周波成分である. V_j の要素を x_j で, W_j の要素を y_j で表す. また, モルフォロジカルフィルタを用いた信号分離において, 式 (3.10) の低周波成分の抽出が, $\psi^{\uparrow} : V_j \to V_{j+1}$ であるとする. 入力信号 $x_0 \in V_0$ が与えられ, 次の再帰的な分解スキームを考える.

(3.16) $$x_0 \to \{x_1, y_1\} \to \{x_2, y_2, y_1\} \to \ldots \to \{x_k, y_k, y_{k-1}, \ldots, y_1\} \to \ldots$$

ここで,

(3.17) $$x_{j+1} = \psi_j^{\uparrow}(x_j) \in V_{j+1}$$

(3.18) $$y_{j+1} = x_j - x_{j+1} \in W_{j+1}$$

である. 信号分離を行う際, 前のレベルよりも大型の構造関数を用いることにより, W_j それぞれに異なる帯域の信号を抽出することができる. $k = 3$ の場合の, 信号分解の概念図を Fig. 9 に示す. 入力信号 x_0 は, 各レベルの信号を加算することにより再構成することができる.

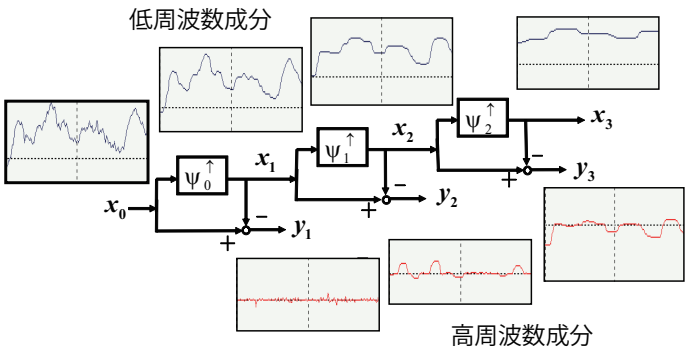


Fig. 9. Schematic diagram of signal decomposition

Fig. 10 に通常のウェーブレット（Daubechies 6) を用いた多重解像度解析結果を, Fig. 11 にモルフォロジー (構造関数 Haar) を用いた多重解像度解析結果を示す．原波形の 5〜8[sec] (動作想像時) に出現しているパルス形状成分は, 分解された高周波成分において, 線形手法であるウェーブレットでは, Level 5 程度まで影響を及ぼし, 非線形手法であるモルフォロジーでは, Level 3 程度まで影響を及ぼしていることがわかる,

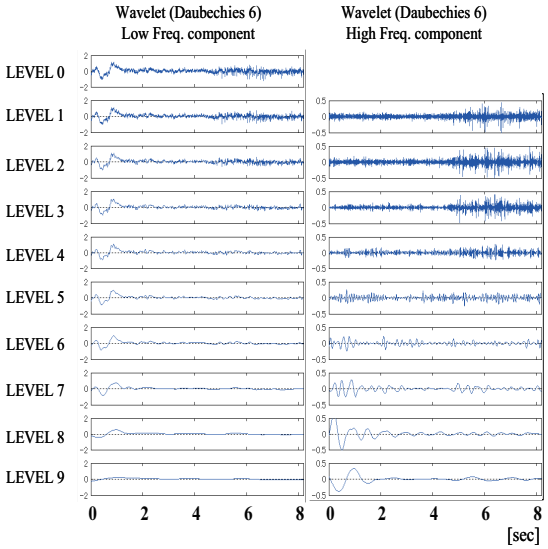


Fig. 10. Result of multi resolution analysis (Wavelet Daubechies 6)

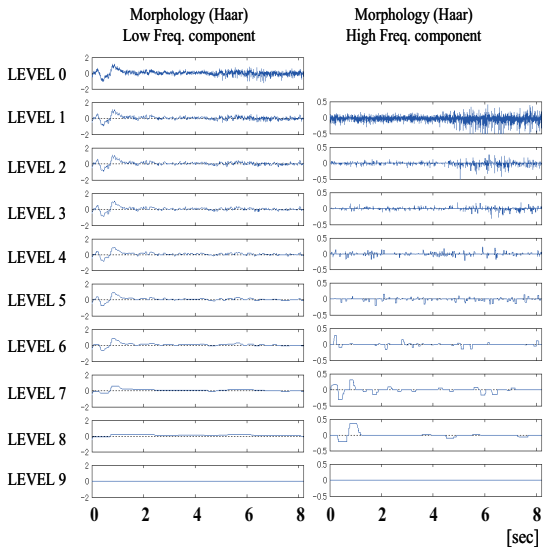


Fig. 11. Result of multi resolution analysis (Morphology Haar)

3.2.4 モルフォロジカル・ウェーブレット

Heijmans によって提案されているモルフォロジカル・ハール・ウェーブレットは，次式で定義される演算を行い，その後ダウンサンプリングを行うものである．

$$x_{j+1}(t) = \min(x_j(2t), x_j(2t+1)), \quad y_{j+1}(t) = x_j(2t+1) - x_j(2t) \tag{3.19}$$

これに対して，単に前のレベルよりウィンドウ幅を拡張して分解を行う多重解像度解析手法がある．これは，各レベルに冗長な情報が含まれることになるが，シフト不変性を有しており，通常の離散ウェーブレット変換では，SWT と呼ばれている．これと同様なものとして，モルフォロジカル・冗長ウェーブレットがある．近似信号は，最大・最小演算の組み合わせで生成され，詳細信号は，モルフォロジカル・ハイパスフィルタを式 (3.11) 通したものと同様なものとなる．

$$\begin{aligned} x_{j+1}(t) &= \epsilon_j(\delta_j(\delta_j(\epsilon_j(x_j))))(t), \quad y_{j+1}(t) = x_j(t) - x_{j+1}(t) \\ \delta_j(f) &= \underbrace{\delta(\delta(\cdots\delta(f))\cdots)}_{j}, \quad \epsilon_j(f) = \underbrace{\epsilon(\epsilon(\cdots\epsilon(f))\cdots)}_{j} \\ \delta(f)(t) &= \max(f(t), f(t+1)), \quad \epsilon(f)(t) = \min(f(t), f(t+1)) \end{aligned} \tag{3.20}$$

レベル j の演算は，ハール構造関数（幅：$j+1$）を用いたモルフォロジカル・ローパスフィルタを意味している．これを利用することによって，短時間データに対するスペクトルに相当するパタンスペクトルを得ることができる．

$$P(x_0, B, j, t_0) = \sum_{t=t_0}^{t_0+\Delta t} \left| y_j(t) \right| \tag{3.21}$$

短時間の音声データに適用した例を Fig. 12 に示すが，他の手法に比べ，ピーク周波数が検出しやすいなどの特徴を持っているものと思われる..

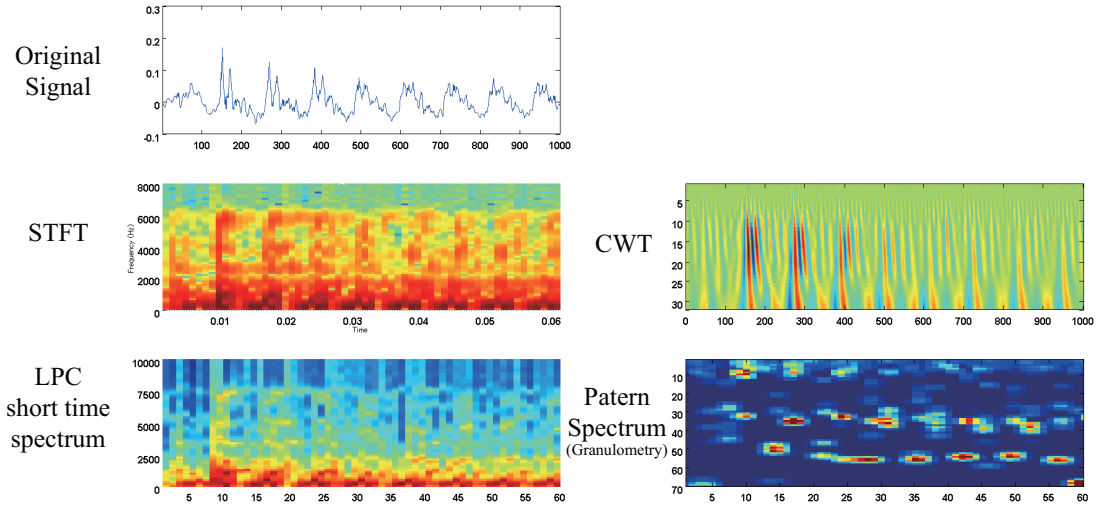


Fig. 12. Time-frequency analysis (short data)
(analyzing step: 16 [pt] (0.8 [ms]), processing length: 40 [pt] (2 [ms]))

4. 脳波解析

4.1 睡眠解析

ここでは, 脳波波形が持つ高周波成分の変動状況を抽出する方法について述べる. Fig. 13 に, そのアルゴリズムの概要を示すが, 得られた時間-周波数解析結果に, 画像処理を施し, その輪郭線を抽出することにより, 特性曲線を得ようとするものである.

この処理によって抽出した特性曲線の一例を Fig. 14 に示す. 図中, 左列上段がウェーブレット変換結果, 左列下段が睡眠脳波ステージ遷移状況, 右列上段が特性曲線抽出結果, 右列下段が心電位を処理して得られた HF, LF 成分の変動状況である. 右列上段の図に, HF 成分を上下反転させたグラフもあわせて示しているが, 本手法によって得られる特性曲線は, HF 成分（副交感神経活動）の変動と非常に良く似た変動を示していることがわかる.

8 名の被験者の 14 夜分（6 名× 2 夜＋ 2 名× 1 夜)）のデータについて，ここで抽出した特徴リズムと HF 成分との相関を調査した結果が，Fig. 15 と Fig. 16 である.

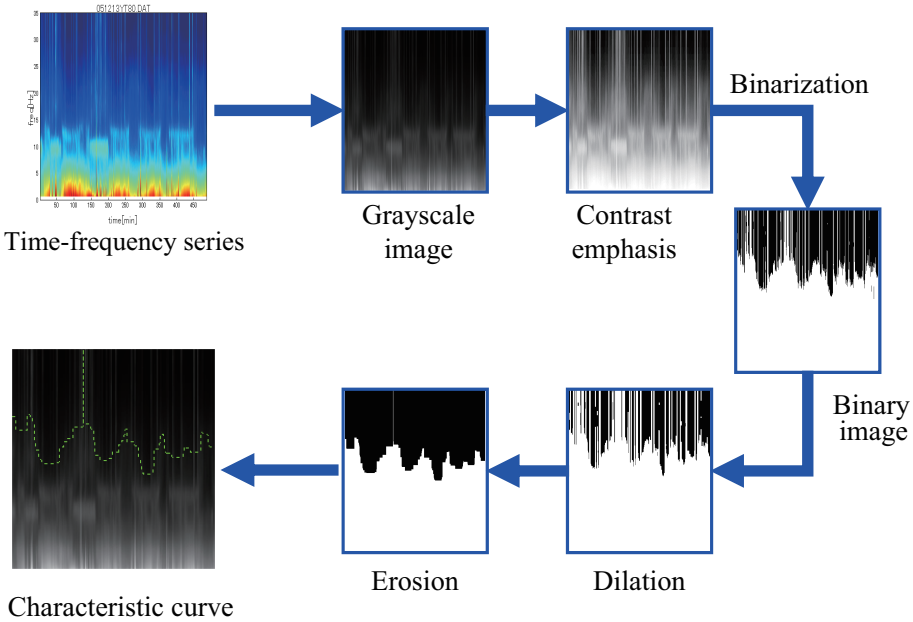


Fig. 13. Processing scheme

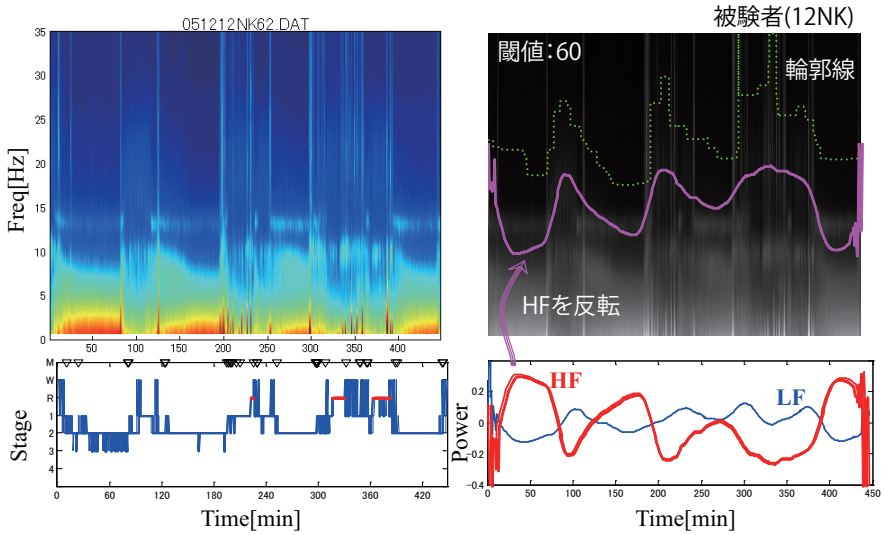


Fig. 14. Feature extraction result

本手法で抽出するリズムは，パワー閾値に依存するが，この閾値を各種設定し抽出した特徴リズムと HF 成分との相関をとり，相関が一番高くなる特徴リズムの平均周波数に関するヒストグラムが Fig. 15 である．14 例中 10 例で，21～27Hz 帯域に相関が高くなっており，このことから，抽出される特徴リズムの平均周波数が 20Hz～30Hz 帯域になるよう，パワー閾値を適応的に設定すれば良さそうであることがわかる．

また，Fig. 16 は，特徴リズムと HF 成分との時間遅れを調査した結果である．特徴リズムの位相を 1 分刻みで遅らせて相関を取り，一番相関が高くなる遅れ時間に関するヒスト

グラムをとったものであるが，この結果より，脳波の特徴リズムは自律神経リズムより位相が遅れていることが確認でき，自律神経リズムと全く等価な情報を抽出している訳ではなく，その差違にも睡眠状態に関する情報が含まれている可能性を示唆している．

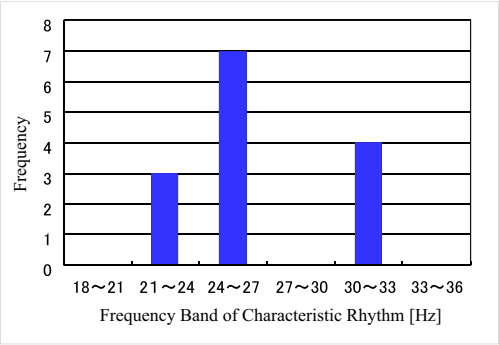


Fig. 15. Histgram of optimal band

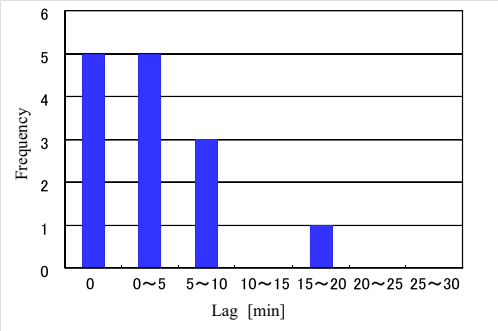


Fig. 16. Histgram of lag time

全例に対して抽出した特徴リズムと自律神経リズムとの相関をとった結果を Fig. 17 に示す．全体的にある程度の相関が得られており，本手法によって抽出した特徴リズムは，生理学的意味がある情報を有しているものと思われる．このことは，これまでの R&K 判読基準に基づく離散的な睡眠ステージ情報にとどまらず，より詳細な睡眠状態の評価へ利用できる可能性を示唆している．

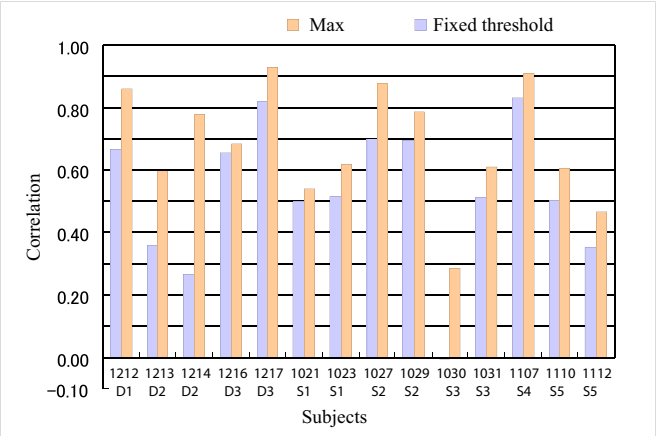


Fig. 17. Correlation between extracted rhythm and HF rhythm

[睡眠脳波のパタンスペクトル]

前節で示したパタンスペクトル解析を行った結果を Fig. 18, Fig. 19 に示す．Fig. 18 は，20 秒間の脳波データの処理を行ったものであり，Fig. 19 は，一晩分の処理結果である．

モルフォロジカル・フィルタの構造要素を連続的に変化させて抽出したものであるが，連続ウェーブレット変換と相似な情報を抽出できており，今後，詳細な解析を進めていきたいと考えている．

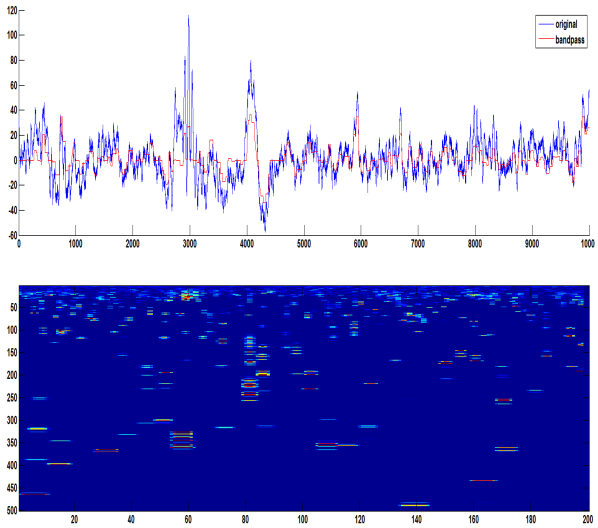


Fig. 18. Patern spectrum of sleep EEG (20sec)

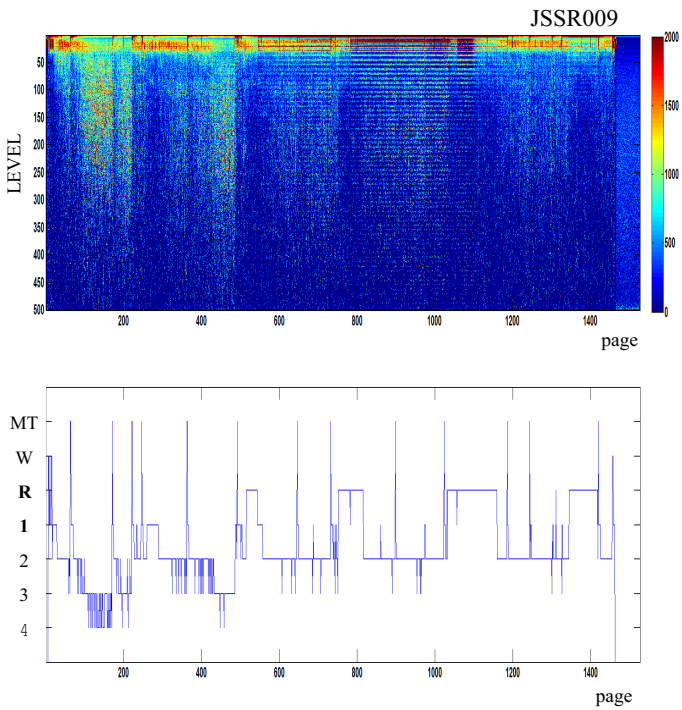


Fig. 19. Patern spectrum of sleep EEG (all night)

4.2 BCI への応用

本節では，2.2 節で説明した実験において得られた脳波に対して適用した結果を説明する．

4.2.1 特徴抽出と識別法

点滅光源が 1 個の場合と，複数の場合の平均加算をとった脳波を Fig. 20 に示すが，多光源の場合，注視点が光っていないにもかかわらず，VEP が出現していることがわかる．

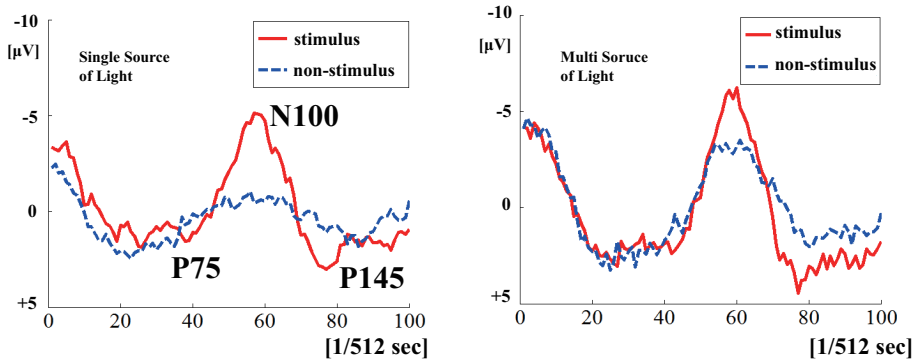


Fig. 20. Averaged signal

これらの現象は，後頭部の刺激と同側側に出現することが知られており，複数個の電極から同時に採取した多チャンネルの脳波を利用することによって，これらの現象の影響を除去できる可能性がある．そこで，ここでは，17 チャンネルの脳波に対して，独立成分分析を用いたブラインド信号分離を前処理として行い，相関情報を用いて，分離された信号群から VEP が多く含まれていると思われるチャンネルを選択後，モルフォロジカルフィルタによってその低域成分を抽出し，注視点の識別を行う．なお，ここでは，モルフォロジカル・フィルタの構造要素としては，次式で定義される超楕円を採用している（ウィンドウ幅：$2 \times n$）．

$$g(t) = k\left(1 - \left|\frac{x}{n}\right|^p\right)^{\frac{1}{p}} \qquad (-n \le t \le n) \tag{4.1}$$

その処理結果の一例を Fig. 21 に示す．分離された信号には，ある程度 VEP 成分に対応した応答が観測できるが，源信号からでは，VEP 成分の抽出が困難であることがわかる．

注視点に対応する信号は，次式で示す脳波変動とテンプレート波形との相関情報 (cor($\mathbf{x}$, $\mathbf{x}_{\mathrm{TEMP}}$)) を用いて決定する．

$$i^* = \arg\max_i \{|r(i)|\},\quad r(i) = \sum_{j=1}^{4} \sum_{k=1}^{N_b(=15)} e_j(k) \cdot \mathrm{cor}\left(\mathbf{x}(i,k), \mathbf{x}_{\mathrm{TEMP}}\right),\quad (i = 1, 2, \cdots, 5) \tag{4.2}$$

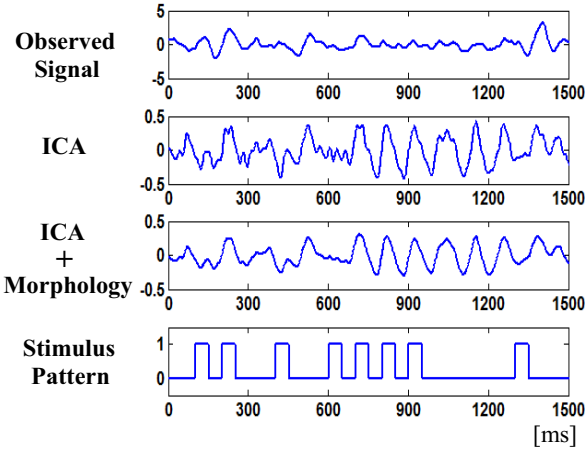


Fig. 21. Processing result

ここで，$\mathbf{x}_{\mathrm{TEMP}}$ は，単一光源刺激提示時の脳波の加算平均をとったものであり，e_j は，クラス j の刺激パタンに対応した値 (刺激提示時: 1, 無刺激時: -1) を設定した系列である．最終的には，sign($r(i)$) によって評価している．

VEP 成分を抽出するためのモルフォロジカル・バンドストップフィルタの構造要素のパラメータは，次式を用いて求める．（2 構造要素× 3 パラメータ＝ 6 パラメータ）

$$\theta^* = \arg\max_{\theta}\{u(\theta)\},\quad u(\theta) = \sum_{j=1}^{4}\sum_{k=1}^{15} e_j(k)\cdot \mathrm{cor}\left(c_{g_1,g_2}, \mathbf{x}_{\mathrm{TEMP}}\right) \tag{4.3}$$

$$\theta = [k_1, n_1, p_1, k_2, n_2, p_2]^T$$

パタン認識における特徴ベクトル $\mathbf{s} = [s_1, s_2, s_3, s_4]^T$ は 1 ビット（100msec）の脳波変動とテンプレート波形との相関情報を用いることにする．

$$s_j = \sum_{k=1}^{15} e_j(k)\cdot \mathrm{cor}\,(\mathbf{y}(k), \mathbf{x}_{\mathrm{TEMP}}) \tag{4.4}$$

判別式としては，特徴量 $\mathbf{s}$ は正規分布に従うと仮定し，ベイズ判定法を採用する．その具体的式は，次のようになる．

$$i^* = \arg\max_i \Pr(\omega_i|\mathbf{s}) = \arg\max_i \left\{\frac{1}{2}\ln|\Sigma_i| - \Pr(\omega_i) + \frac{1}{2}(\mathbf{s}-\mathbf{m}_i)^T\Sigma_i^{-1}(\mathbf{s}-\mathbf{m}_i)\right\} \tag{4.5}$$

4.2.2 認識結果

Fig. 22 に，4 次元の特徴量 $\mathbf{s} = [s_1, s_2, s_3, s_4]^T$ の中の 2 つの特徴量の分布を示す (各特徴量は，7.5 秒間（5 × 15 ビット）間の加算平均をとったもの)．この図より，処理を行うことによって特徴量の分布が，より分離できていることがわかる．

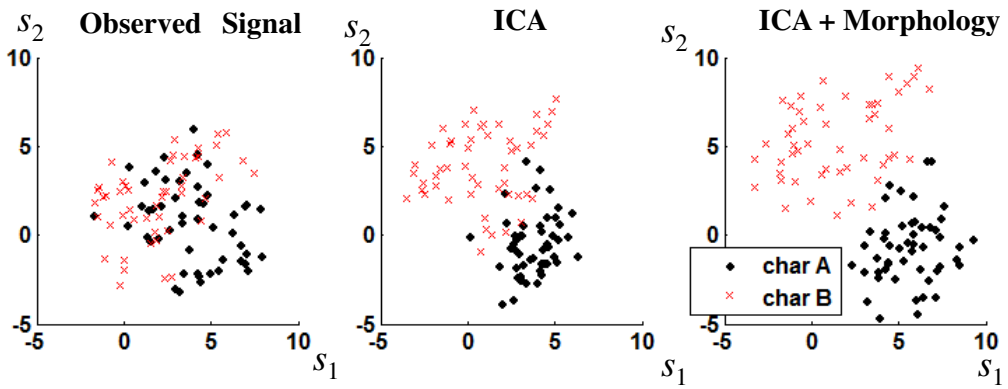


Fig. 22. Distribution of features
(Subject:S4, number of addition:5 (7.5 sec.))

Fig. 23 は精度と加算数（処理データ長）との関係を示しており，あまりに多くの加算を行うことは，処理遅延（30 回：45 秒）を意味し，現実的ではないものの，加算数を多くすることによって，精度が向上していることが確認できる．

5 名の被験者における 4 クラス間の識別結果の平均は，1 系列（1.5 秒）を使用した場合は 49% で，5 系列（7.5 秒）を使用した場合は 74% であり，被験者 S4 の場合，15 系列を使用した場合，87% の精度が得られている．また，80% の精度を得るために必要な系列数は，ICA だけの場合は 10 系列で，ICA+M-filter の場合は 9 系列であり，本質的には ICA による効果が大きいものの，モルフォロジカル・フィルタは精度の改善に寄与していることが確認できる．

ただ，本実験において，4 ビット続けて点滅せた場合，次のビット区間では，点灯しなくても VEP が現れる傾向にあることが観測されており，定常 VEP 反応を励起している可能性があることから，刺激提示パタン自身の検討が必要である．

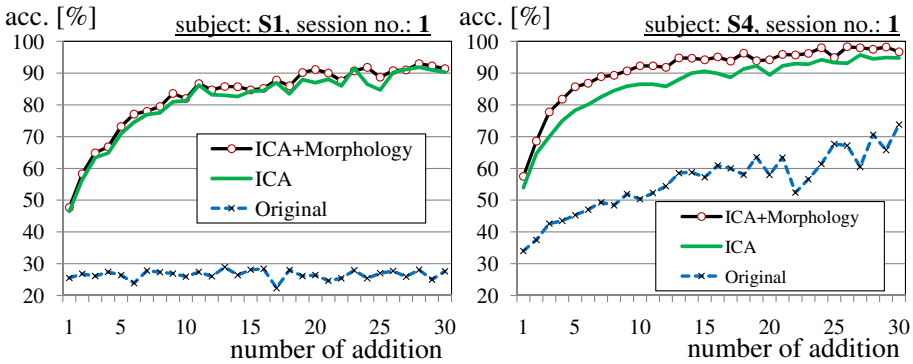


Fig. 23. Pattern recognition results with number of addition

5. 終わりに

本稿では，ウェーブレット変換手法や，その非線形バージョンとも考えることができるモルフォロジカルフィルタを用いた多重解像度解析手法を終夜睡眠脳波の解析や，BCI (Brain Computer Interface) における誘発電位の抽出に応用した事例を紹介してきた．連続ウェーブレット変換手法によって得られた結果より，睡眠のリズムに関する情報の抽出や，モルフォロジカル・ウェーブレット手法を用いた睡眠脳波解析，BCI システムにおける誘発電位の抽出が可能になるなど，その有効性を確認したが，非線形演算を要するモルフォロジカル・ウェーブレットに関する理論体系の整備はできておらず，今後検討を進めていきたいと考えている．

参考文献

[1] 大熊輝雄, 臨床脳波学 第 5 版, 医学書院, (1999)

[2] 井上勝裕, 睡眠脳波ステージ自動判定システムの構築への適用: ウェーブレット解析の産業応用: 電気学会ウェーブレット解析の産業応用に関する協同研究委員会編, 朝倉書店, (2005), 178–202.

[3] Hori T., Y. Sugita *et. al.*, Proposed Supplements and Amendments to ’A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects’, the Rechtschaffen & Kales (1968) standard, Psychiatry and Clinical Neurosciences, **55** (2001), 305–310.

[4] Inoue K., T. Tsujihata, K. Kumamaru and S. Matsuoka, Feature extraction of human sleep EEG based on a peak frequency analysis; Proc. of the 16th IFAC World Congress, Prague, July 4-8 (2005)

[5] Msee R. L., Interpretation of Normalized Spectral Heart Rate Valiability Indices In Sleep Research: A Critical Review, Sleep, **30-7** (2007).

[6] Niedermeyer E. and F. L. Da Silva, Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, Williams & Wilkins, (1999)

[7] Rechtshaffen A. and A. Kales, A manual of standardized terminology, techniques and scoring system for sleep stages of human subject, Public Health Service U.S. Government Printing Office, Washington D.C. (1968)

[8] Guger C., S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica and G. Edlinger, How many people are able to control a P300-based brain-computer interface (BCI)?, Neuroscience Letters, **462** (2009), 94–98.

[9] Inoue K., K. Kumamaru and G. Pfurtscheller, Robot Operation based on Pattern Recognition of EEG Signals: Proc. of the 3rd International Brain-Computer Interface Workshop and Training Course 2006, Graz, Austria (2006), 116–117.

[10] Inoue K., M. Fujio, T. Yamaguchi and G. Pfurtscheller: Mathematical morphological multi-resolution analysis of EEG signals during misoperation of BCI system: Proc. of the 4th International Brain-Computer Interface Workshop and Training Course 2008, Graz, Austria (2008), 74–79

[11] Martinez P., H. Bakardjian and A. Cichocki, Fully Online Multicommand Brain-Computer Interface with Visual Neurofeedback Using SSVEP Paradigm, Computer Intelligence and Neuroscience, (2007).

[12] Pfurtscheller G., C. Brunner, A. Schlögl and F. H. Lopes da Silva, Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks, NeuroImage **31-1** (2006), 153–159.

[13] Addison P. S., The Illustrated Wavelet Transform Handbook, Institute of Physics Publishing, Bristol and Philadelphia (2002)

[14] H. J. A. M. Heijmans and J. Goutsias, Nonlinear multiresolution signal decomposition schemes Part II: Morphological wavelets, IEEE Trans. on Image Processing, **9-11** (2000), 1897–1913.

[15] Hyvärinen A, J. Karhunen and E. Oja, Independent Component Analysis, Wiley-Interscience, 2001

[16] Goutsias J., *et. al.*, Nonlinear Multiresolution Sig-nal Decomposition Shemes-Part I: Morphological Pyramids, IEEE Trans. on Image Processing,**9-11** (2000).

[17] Goutsias J., *et. al.*: Nonlinear Multiresolution Signal Decomposition Shemes-Part II: Morphological Wavelets: IEEE Trans. on Image Processing,**9-11** (2000).

[18] Nobuhara H. et al., Max-plus algebra-based wavelet transforms and their FPGA implementation for image coding, Information Sciences, **180-17** (2010).

[19] Serra J., Image Analysis and Mathematical Morphology, U.K. Academic, London (1982)

[20] Yang B. and S. Li, Multi-Focus Image Fusion Using Watershed Transform and Morphological Wavelet Clarity Measure, International Journal of Innovative Computing, Information and Control, **7-5(A)** (2011), 2503–2514.

井上 勝裕 (九州工業大学大学院 情報工学研究院 システム創成情報工学研究系)
〒820-8502 福岡県飯塚市川津 680-4
E-mail: inoue@ces.kyutech.ac.jp

山口 朋成 (九州工業大学大学院 情報工学府 情報科学専攻)
〒820-8502 福岡県飯塚市川津 680-4
E-mail: t_yamagu@kiri.ces.kyutech.ac.jp

前田 誠 (九州工業大学大学院 情報工学研究院 システム創成情報工学研究系)
〒820-8502 福岡県飯塚市川津 680-4
E-mail: inoue@ces.kyutech.ac.jp

藤尾 光彦 (九州工業大学大学院 情報工学研究院 システム創成情報工学研究系)
〒820-8502 福岡県飯塚市川津 680-4
E-mail: inoue@ces.kyutech.ac.jp

文部科学省

MEXT

MINISTRY OF EDUCATION,
CULTURE, SPORTS,
SCIENCE AND TECHNOLOGY-JAPAN

OSAKA KYOIKU
UNIVERSITY