

映像符号化における DWT の直交性と歪みの相関

安達 充幸[†] 比企 豊[†] 村松 正吾[†] 菊池 久和[†]

[†]新潟大学工学部電気電子工学科

〒950-2181 新潟市西区五十嵐 2 の町 8050

E-mail: †{adatsu,mhiki}@telecom0.eng.niigata-u.ac.jp, ††{shogo,kikuchi}@eng.niigata-u.ac.jp

あらまし 本研究では、映像符号化における離散ウェーブレット変換 (DWT) の直交性と歪みの相関の検討を行う。JPEG2000 で採用されている DWT は映像応用において霞状の視覚的劣化を生じる。この原因について変換の直交性に着目する。具体的には先に筆者らが提案した直線位相直交非分離 DWT や直交性を満たす様々な DWT を用いてのシミュレーションと実際の動画データを用いた実験により、直交性と歪の相関の関係を調べる。
キーワード ウェーブレット, 映像符号化, 非分離フィルタバンク

Orthogonality of DWTs and Correlation of Distortion in Video Coding

Atsuyuki ADACHI[†], Minoru HIKI[†], Shogo MURAMATSU[†], and Hisakazu KIKUCHI[†]

[†] Dept. of Electrical and Electronic Eng., Faculty of Eng., Niigata University

8050, Ikarashi 2-no-cho, Nishi-ku, Niigata-shi, 950-2181

E-mail: †{adatsu,mhiki}@telecom0.eng.niigata-u.ac.jp, ††{shogo,kikuchi}@eng.niigata-u.ac.jp

Abstract In this work, we evaluate the relation between the spatial correlation of the distortion in the reconstructed picture and the orthogonality of DWTs in video coding. The DWTs used in JPEG2000 are known to yield fluctuation of hazy distortion when applied to motion pictures. Our conjecture is that the orthogonal property affects the artifacts. Simulation and coding experiments for real videos are shown for evaluating the relation between the spatial correlation of the distortion in the reconstructed picture and the orthogonal property of the transform.

Key words Wavelets, Video coding, Nonseparable filter bank

1. はじめに

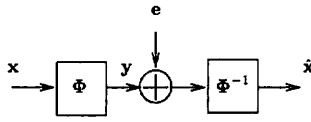
JPEG や JPEG2000 に代表される静止画像符号化国際標準方式や MPEG-1/2 や H.264/AVC に代表される動画画像符号化国際標準方式では、変換符号化方式が空間相関を利用する要素技術として採用されている [1-3]。その変換符号化の性能は、選択する変換に大きく影響を受ける。画質に与える影響の他にも演算量やメモリ使用量など実装におけるコストも重要な要素となる。

JPEG や MPEG-1/2 では、 8×8 の 2 次元 DCT を採用している。DCT は相関の強い 1 次自己回帰モデルに対して準最適変換であり、余弦関数を変換核として高速アルゴリズムが存在する。このため、静止画像、動画画像共に好まれて利用されてきた。一方、ブロック毎の独立な処理のため、低ビットレート符号化時にブロックの境界が現れるブロックノイズが問題となる。また、高周波数成分に対応する基底画像が大きいためにモスキートノイズが生じてしまう。JPEG2000 では、DWT を採用することでこの問題を克服した。

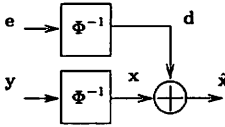
DWT は符号化利得の向上や空間スケーラビリティの実現に

寄与するばかりではなく、ブロック間で重複した変換を与え、周波数に応じて異なる大きさの基底画像が利用される。結果として、ブロックノイズやモスキートノイズが生じない。しかしながら、JPEG2000 で採用されている 5/3 DWT や 9/7 DWT では低ビットレート符号化時に霞状ノイズが生じ、Motion-JPEG2000 のような映像応用において視覚的に不快な揺らぎを与えてしまう [4]。これは、直交性を満たさないことが原因と考えられる。なお、最新の映像符号化方式である H.264/AVC では、DWT ではなく整数型 DCT 等のブロック毎の変換を採用している。

本研究では、直線位相直交非分離 DWT を用いて直交性と歪の相関の検討を行う。筆者らは先に提案した 2×2 分割直線位相直交非分離フィルタバンク [5, 6] に 2 次のパニッシングモーメントを満たすための制約を導入し、 2×2 分割を繰り返すツリー構成として DWT を与えた [7-9]。直線位相直交非分離 DWT は重複性を満たし、周波数に応じて基底画像の大きさが異なる。したがって、ブロックノイズやモスキートノイズを抑圧することができる。また、直交性を満たすために、5/3DWT や 9/7DWT のような霞状ノイズの低減が可能となる。双直交の DWT では、ラ



(a) 基本構成



(b) 画像再構成の等価表現

図1 変換符号化の概略

ランダムサブバンド信号の再構成画像が空間的相関の高い歪みを示すのに対し、直交変換では空間的に相関の低い歪みが与えられる。信号対雑音比には現れない性能であるが、映像符号化応用では重要と考えられるため、実験結果を視覚的及び数値的に示し、直交性と歪みの相関の関係を検討する。

本報告の構成は以下のとおりである。まず第2章では、変換符号化における直交変換の重要性を述べる。第3章では、ラティス構成に基づく直線位相直交非分離フィルタバンクについて紹介する。さらに2次のパニッシングモーメントを満たすための制約について述べ、 2×2 分割を繰り返すツリー構成としてDWTを与える。第4章では、ランダムサブバンド信号の再構成シミュレーションと実際の動画像データに対する符号化実験を行う。従来の5/3 DWT, 9/7 DWTと直交変換である非分離DWT, Daubiches DWT, Haar DWT, Symlets, 8×8 ブロック DCT などと比較し、直交性と歪みの相関の関係を検討を行う。

2. 映像符号化のための変換技術

本章では、変換符号化について概説する。変換領域における量子化誤差と再構成画像における歪みの関係を説明し、双直交変換の問題点を述べる。

2.1 量子化誤差と再構成歪みの関係

DCT, フィルタバンク, DWT は、対象信号全てを要素としたベクトルに対する行列演算として表現できる [10]。この考えに基づいて、変換符号化の概略を図1(a)に示す。図中の x と \hat{x} はそれぞれ原信号と再構成信号をベクトルとして表現している。 Φ は順変換行列を、 Φ^{-1} はその逆変換行列を表している。 y は変換係数、或はサブバンド信号を、 e は量子化によって生じる誤差を表している。本研究では量子化誤差を加法的雑音と仮定する。図1(b)は、図1(a)の行列変換の線形性から得られる等価表現である。図中の d は、量子化によって生じる再構成信号に重畳する歪みを表している。したがって、

$$\hat{x} = x + d \quad (1)$$

$$d = \Phi^{-1}e \quad (2)$$

という関係が成り立つ。

いま、量子化誤差 e の平均が零であると仮定すると、その共分

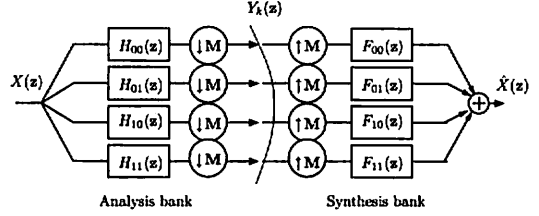


図2 2×2 分割フィルタバンクの並列構成。

散行列は

$$R_{ee} = E[ee^T] \quad (3)$$

と与えられる。ここで、 $E[\cdot]$ は期待値である。また、再構成歪み d の共分散行列は

$$R_{dd} = E[dd^T] = \Phi^{-1}R_{ee}\Phi^{-T} \quad (4)$$

と求められる。

2.2 直交性と再構成歪みの関係

R_{dd} の非対角要素が零となれば、歪み d は空間相関を持たない。しかしながら、5/3 DWT や 9/7 DWT などの双直交変換ではこの達成が困難である。このことを以下に示そう。

簡単のため、量子化誤差は係数間で無相関であり、同じ分散 σ_e^2 を持つと仮定しよう。すると、

$$R_{ee} = \sigma_e^2 I \quad (5)$$

となる。ここで、 I は単位行列を意味する。

2.2.1 直交の場合

Φ が直交ならば

$$R_{dd} = \sigma_e^2 I \quad (6)$$

となり、 R_{dd} もまた対角行列となる。すなわち、再構成画像に重畳する歪み d は画素間で無相関となる。

2.2.2 双直交の場合

一方、 Φ が双直交の場合、 $S = \Phi^{-1}\Phi^{-T}$ として、

$$R_{dd} = \sigma_e^2 S \quad (7)$$

となる。ここで、 S は対称行列である。この場合、 R_{dd} の非対角要素は非零となる。すなわち、再構成画像に重畳する歪み d は画素間で相関を有する。これは、視覚的に不快な隠状のノイズの要因となり得る。この歪み d の空間における相関は PSNR による評価には現れない。

3. 直線位相直交非分離 DWT

本章では、筆者らが先に提案した直線位相直交非分離フィルタバンクについて説明し、直線位相直交非分離 DWT を概説する。

3.1 直線位相直交非分離フィルタバンク

JPEG2000 で採用されている 5/3 DWT や 9/7 DWT は、重複性、直線位相特性を満たす。またリフティング構成で実現でき、ロスレス/ロッキー統合符号化を与える。しかしながら、双直交変換

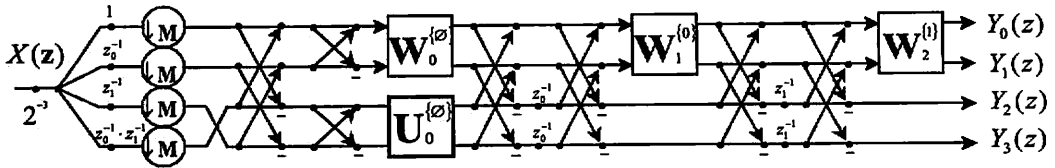


図3 直線位相直交非分離フィルタバンクの構成例(分析器)

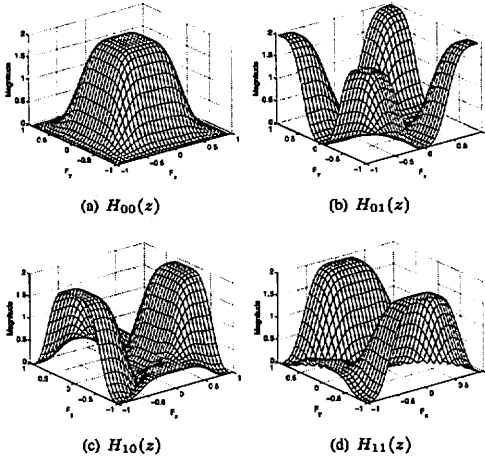


図4 2×2 分割直線位相直交非分離フィルタバンクの周波数特性 (12 × 12 タップ)

のために先の歪みの相関が問題となる。特に Motion-JPEG2000 のような映像応用において、この歪みは視覚的に不快な揺らぎの原因となる [4]。

重複性、直線位相特性、直交性を同時に満たす 2 分割フィルタバンクは存在しない。しかしながら、水平方向、垂直方向を同時に 2×2 分割する直線位相直交非分離フィルタバンクの設計は可能である [5, 6, 8]。

先に筆者らは直線位相直交非分離フィルタバンクのラティス構成とその設計法を提案した [5]。さらに、Lu Gan らは設計クラスをそのままにパラメータ数を削減する手法を示した [6]。図 2 に 2×2 分割フィルタバンクの並列構成を示す。図中の $\downarrow M$ と $\uparrow M$ はそれぞれ行列 M による間引きと零値挿入を意味しており、 2×2 分割では $M = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ と選択する。 $H_{k\ell}(z)$ および $F_{k\ell}(z)$ はそれぞれ 2 次元非分離の分析フィルタと合成フィルタを意味する。また、ラティス構成による直線位相直交非分離フィルタバンクの分析器の構成例を図 3 に示す。図 3 において $X(z)$ は入力信号、4 つの $Y_k(z)$ は 2×2 分割のサブバンド信号、 z_d は遅延器を示し、 $W_{n_1}^{[1]}$ 、 $U_{n_0}^{[0]}$ は正規直交行列を表す。この行列に制約を加えることでさまざまな性質を持つフィルタバンク設計が可能となる。実装の詳細については文献 [5, 6] を参照されたい。

3.2 2 次元のパニッシングモーメント制約

文献 [7] に 2 次元非分離 DWT のパニッシングモーメントとレギュラリティの関係が示されている。また、文献 [8] では最大のパニッシングモーメントを与える 2 次元直線位相直交フィル

タバンクの設計法が示されている。しかしながら、この方法ではフィルタの次数に応じて制約式を導出する必要がある。文献 [11] では、2 次元のパニッシングモーメントを有する 1 次元の直線位相直交 M 分割フィルタバンクの設計法が示されている。筆者らは、これと類似した手法により、2 次元直線位相直交フィルタバンクのパニッシングモーメントに関する一般的な制約式を先に提案している [9]。

文献 [6] の Lu Gan の構成に基づくことにより 2 次元のパニッシングモーメントを有するための制約式が

$$C_0^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0 \quad (8)$$

$$C_1^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0 \quad (9)$$

と導出される。ただし、

$$C_0 = \left(\prod_{n_1=1}^{N_1} W_{n_1}^{[1]} \right) \cdot \sum_{k_0=1}^{N_0} \left(\prod_{n_0=k_0}^{N_0} W_{n_0}^{[0]} \right) \cdot U_0^{[0]} \quad (10)$$

$$C_1 = \sum_{k_1=1}^{N_1} \left(\prod_{n_1=k_1}^{N_1} W_{n_1}^{[1]} \right) \cdot U_0^{[0]} \quad (11)$$

である。また、1 次元のパニッシングモーメントを有するための制約

$$W_0^{[0]T} = \left(\prod_{n_1=k_1}^{N_1} W_{n_1}^{[1]} \right) \cdot \left(\prod_{n_0=k_0}^{N_0} W_{n_0}^{[1]} \right) \quad (12)$$

を満たすものとする。記号については文献 [5] を参照されたい。なお、文献 [6] に従い $U_d^{[d]} = I (d \neq 0)$ とおく。フィルタのタップ数は $2(N_0 + 1) \times 2(N_1 + 1)$ と与えられる。

図 4 に、この制約を用いて設計された 4 つの分析フィルタの周波数特性を示す。直交性により、合成フィルタ $F_{k\ell}(z)$ は分析フィルタ $H_{k\ell}(z)$ のインパルス応答(基底画像)を垂直および水平方向に反転することで与えられる。この 2×2 分割を繰り返すことで従来の DWT に変わる直線位相直交非分離 DWT が実現できる。

3.3 基本構成

図 5(a) に分析バンク、図 5(b) にその逆変換となる合成バンクのツリー構成を示す。分析バンクでは、まず入力画像に対して 2×2 分割非分離フィルタバンク H により、4 つのサブバンド信号を出力として得る。これを 1 段階の処理とし、低域通過フィルタ $H_{00}(z)$ から得られるサブバンド信号を、次段のフィルタバンクの入力信号とする。2 段目以降も同様、階層的に処理を行うことで DWT を実現する。また、合成バンクでは、逆変換として

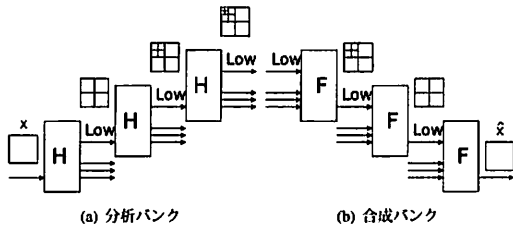


図5 2×2分割非分離DWTのツリー構成(分解レベル3)

フィルタバンク F を用いて、4つのサブバンド信号から1つの画像を合成する処理を繰り返す。本手法では、重複性を満たすことでブロックDCTとは異なり耐ブロックノイズ性を持ち、多重解像度表現が実現できる。また、従来の5/3DWT、9/7DWTは双直交変換であるのに対し、直線位相直交非分離DWTでは直交性を満たすため、第2章で述べたように霞状ノイズの原因となり得る歪みの空間的相関の低減が可能である。

4. 性能評価

ランダムノイズ再構成に関するシミュレーション結果と実際の動画データを用いた実験結果を以下に示し、直交性と歪みの相関について検討を行う。

4.1 ランダムサブバンド信号の再構成

第2章で概説したとおり、サブバンド領域での量子化は再構成画像の歪みとなって現われ、その空間相関は変換の直交性に大きく依存する。ここではPSNRに反映されない部分の比較として、ランダムサブバンド信号の再構成シミュレーションを行った。本シミュレーションでは無相関で一様な同一の分散を持つランダムノイズをサブバンド信号とし、それぞれの変換の合成器によって画像を再構成した。これらの結果を比較して、量子化雑音が画像の歪みに及ぼす影響を確認する。

本シミュレーションでは、大きさ 352×224 のランダムノイズを用いた。5段階での再構成結果を図6に示す。図6(a)が原画像、(b)(c)が5/3DWT、9/7DWT、(d)が非分離直交DWTで処理を行った結果である。また、直交非重複変換である 8×8 ブロックDCT、Haar DWTでの結果をそれぞれ(e)、(f)に、直交非直線位相特性を持つ Daubiches DWT (4×4 taps, 6×6 taps, 8×8 taps)、Symlet (8×8 taps)での結果をそれぞれ(g),(h),(i),(j)に示す。図に示すように、双直交変換では、無相関のランダムノイズが相関のある歪みとして再構成されてしまう。しかし、直交性を持つ変換はほぼ無相関のまま再構成できる。歪みの相関は低周波に偏り、静止画に比べ動画において視覚的に不快な揺らぎとなり得る。

図7にDWTの分解レベルに対する再構成画像の相関係数の変化を示す。図7より双直交変換である5/3DWT、9/7DWTではレベルが上がるにつれ相関も高い値となるのに対し、他の全ての直交変換ではほぼ無相関のまま再構成することがわかる。歪みの相関は低周波に偏り、静止画に比べ動画において視覚的に不快な揺らぎとなり得る。

4.2 量子化手法

以下の実験では、量子化の手法として、エントロピー符号化ス

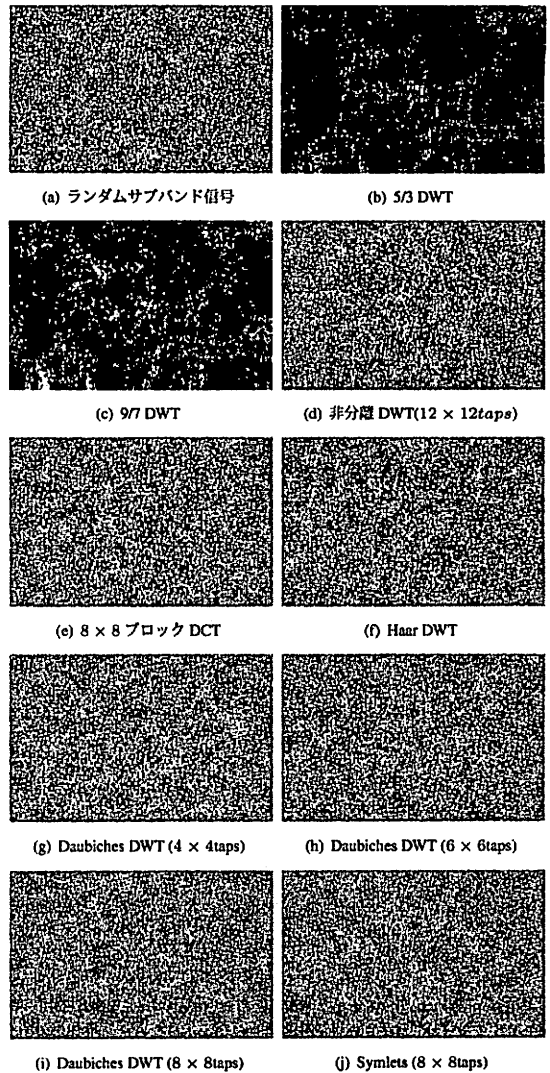


図6 ランダムサブバンド信号の再構成(5段階)

カラー量子化(ECSQ)を用いる[1]。まず、次式により各サブバンドの最適割り当てレート R_b を求める。ただし、 b はサブバンドのインデックスである。

$$R_b = R + \frac{1}{2} \log_2 \frac{G_b W_b \epsilon_b^2 \sigma_b^2}{\prod_{i=0}^{B-1} (G_i W_i \epsilon_i^2 \sigma_i^2)^{n_i}} \quad (13)$$

ここで、 R は任意の平均レート、 G_b は対応する合成フィルタのゲイン(フィルタ係数の二乗和)、 $W_b = 1$ (重み係数)、 $\epsilon_b^2 = 1$ (一様分布の場合)、 σ^2 はサブバンド信号の分散である。この式で求められる R_b から次式より量子化ステップサイズ Δ_b を求める。

$$\Delta_b = \sqrt{\frac{\epsilon_b^2 \sigma_b^2}{c_b} 2^{-R_b}} \quad (14)$$

ここで、 $c_b = \frac{1}{12}$ である。最後に、この量子化ステップサイズ Δ_b を用いて、

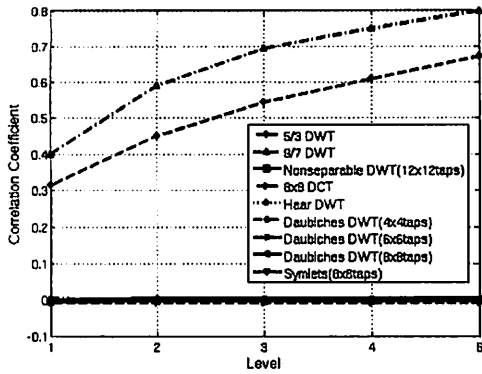


図7 レベルを変えたときの相関係数の変化(ランダムサブバンド信号再構成)

$$q = \begin{cases} \text{sign}(x) \lfloor \frac{|x|}{\Delta_b} \rfloor & (\frac{|x|}{\Delta_b} > 0) \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

のように量子化を行う。また、逆量子化は

$$\hat{x}_q = \begin{cases} 0 & (q = 0) \\ \text{sign}(q) \lfloor |q| \Delta_b & (q \neq 0) \end{cases} \quad (16)$$

のように計算する。なお今回は、負の割り当てが生じないようにビットの再配分を繰り返した [10]。

4.3 動画実験

実際の動画像に対してそれぞれの変換により処理を行い、どのような歪み与えられるかを確かめる。本実験では、ITE (映像情報メディア学会) 標準動画像 'Boy and Toys' (8-bit, gray-scale) 450 フレームを使用し、平均ビットレート 0.5[bpp] とし、従来の 5/3DWT, 9/7DWT と直線位相直交非分離 DWT, 直交 DWT などとの比較を行う。

ここでは、フレーム画像の処理例としてフレーム 350 番目の処理結果を示す。それぞれの変換により得られた画像と原画像との差分をとり、原画像に加わった歪みを図 8 に示す。図 8 より、特に 5/3 DWT の歪みがランダムサブバンド信号を再構成した図 6 の歪みと類似することが確認できる。

4.3.1 平均 PSNR 値の比較

符号化性能を評価するために、平均ビットレート R [bpp] を変化させ、それぞれ手法における PSNR を比較した。図 9 にその結果を示す。本実験では 'Boy and Toys' 450 フレームを利用し、5 段階 DWT を行った。図 9 より、提案法では 5/3 DWT よりは高い数値を示し、低ビットレート時には 9/7 DWT と同等程度の値を得ることができた。

4.3.2 レベル毎の平均相関係数の比較

各フレーム画像毎、変換により得られた再構成フレーム画像と原フレーム画像との差分をとり、原フレーム画像にのった歪をとり出す。その歪の相関係数を測定し 450 フレームでの平均値を計算する。これを DWT の分解レベルを変えて行い、結果を図 10 に示す。図 10 (a) には、5/3 DWT, 9/7 DWT, 非分離

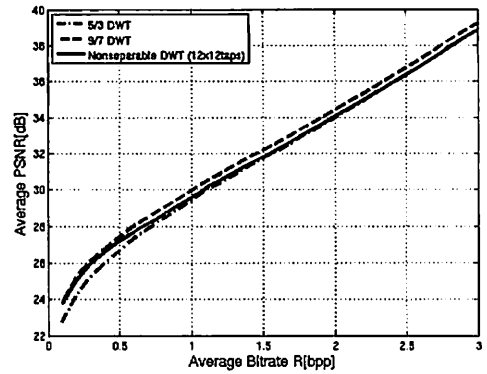
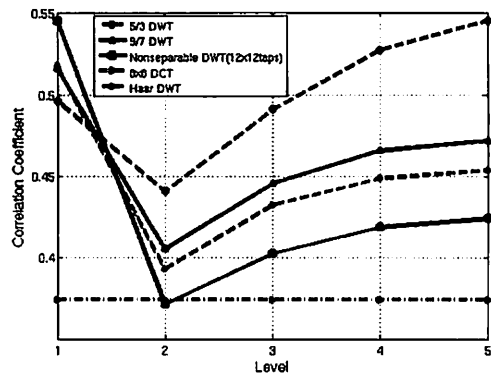
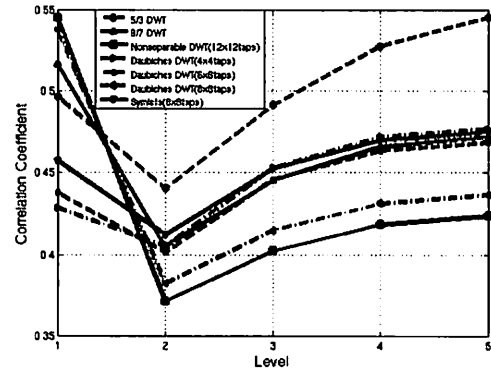


図9 平均 PSNR による比較



(a) 5/3 DWT, 9/7 DWT, 非分離 DWT と直線位相直交変換との比較



(b) 5/3 DWT, 9/7 DWT, 非分離 DWT と非直線位相直交変換との比較

図10 レベルを変えたときの相関係数の変化(平均ビットレート 0.5bpp)

DWT(12 × 12taps) と、直交性、直線位相特性を持つ 8 × 8 ブロック DCT, Haar DWT を示す。また、図 10(b) には、5/3 DWT, 9/7 DWT, 非分離 DWT と、直交性、非直線位相特性を持つ Daubiches DWT (4 × 4taps, 6 × 6taps, 8 × 8taps), Symlet (8 × 8taps) を示す。

直線位相をもつ 8 × 8 ブロック DCT, Haar DWT では従来の双直交 DWT より低い歪みの相関を示すのに対し、非直線位相

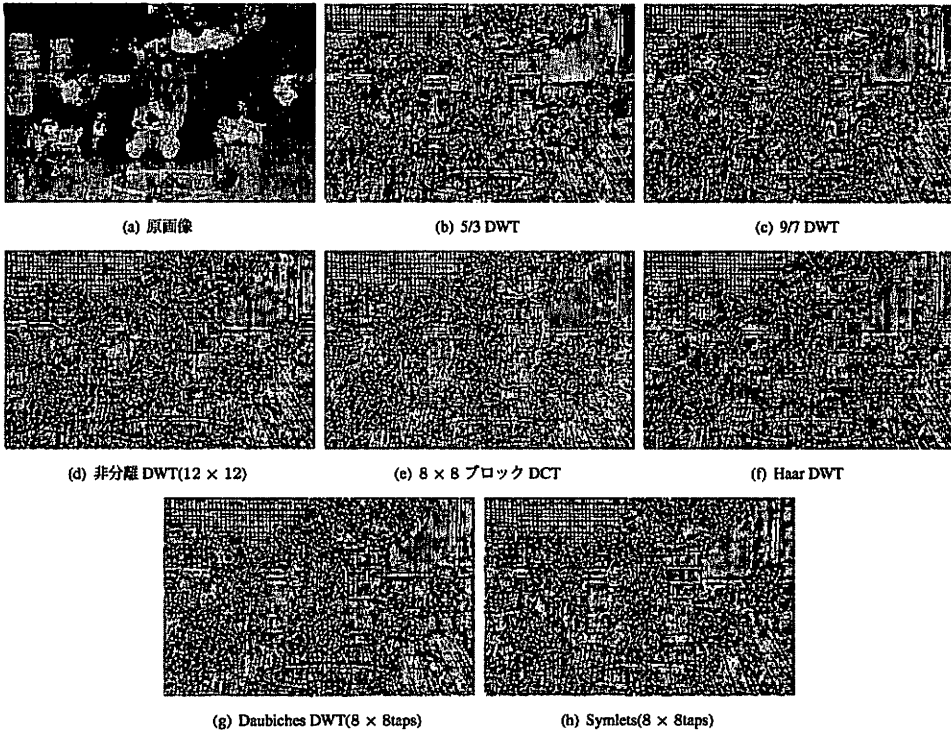


図 8 平均ビットレート 0.5[bpp] における再構成画像と原画像との差分画像

直交変換である Daubiches DWT (4×4 taps), Symlets (8×8 Staps) では、双直交 DWT と比べて弱冠高い相関を示す結果となった。今回の実験では、低ビットレートにおいて量子化の最適性が崩れたときのサブバンド内での量子化誤差の分布や量子化がサブバンドに与える影響などを考慮に入れなかったことも原因としてあげられる。また、分解レベル 1 において全体的に高い相関を示すのは、サブバンドにおける分散の偏りが少ないために量子化時に低周波部分の情報が切られてしまうためだと考える。

5. おわりに

本研究では、映像符号化のための DWT において直交性の重要性を示した。ランダムノイズの再構成シミュレーションをとおして双直交 DWT の映像応用における霞状の視覚的劣化の原因を指摘した。また、様々な直交変換を用いた実験を行い、双直交変換との比較を行った。しかしながら、実験では 9/7 DWT と比較して歪みの相関が高くなる非直線位相直交 DWT も確認された。

今後は最適化が崩れたときのサブバンド内での量子化の影響や誤差の分布の確認、また様々な特性をもつ映像での検討を行う。さらに高次なバニッシングモーメントを持つフィルタバンク設計と演算量、境界部の処理などの改善を検討する。

文 献

- [1] D.S. Taubman, and M.W. Marcellin, JPEG2000, Image Compression Fundamentals, Standards and Practice, Kluwer Academic Publishers, 2002.
- [2] M. Ghanbari, Standard Codecs: Image Compression to Advanced Video Coding, The Institution of Electrical Engineers, 2003.
- [3] I.E. Richardson, H.264 and MPEG-4 Video Compression, Wiley, 2003.
- [4] 久下哲朗, “Wavelet 画像符号化の視覚的歪みに関する考察 インタレース HDTV 高精細画像応用の課題,” 映メ学会技報, vol.25, no.79, pp.33-38, 11 2001.
- [5] S. Muramatsu, A. Yamada, and H. Kiya, “A Design Method of Multidimensional Linear-Phase Paraunitary Filter Banks with a Lattice Structure,” IEEE Trans. on Signal Proc., vol.47, no.3, pp.690-700, March 1999.
- [6] L. Gan, and K.K. Ma, “A simplified lattice factorization for linear-phase perfect reconstruction filter bank,” IEEE Signal Processing Letters, vol.8, no.7, pp.207-209, July 2001.
- [7] D. Stanhill, and Y. Y.Zeevi, “Two-Dimensional Orthogonal Wavelets with Vanishing Moments,” IEEE Trans. on Signal Proc., vol.44, no.10, pp.2579-2590, October 1996.
- [8] D. Stanhill, and Y. Y.Zeevi, “Two-Dimensional Orthogonal Filter Banks and Wavelets with Linear Phase,” IEEE Trans. on Signal Proc., vol.46, no.1, pp.183-190, January 1998.
- [9] 安達 充幸, 比企 豊, 村松 正吾, 菊池 久和, “映像符号化のための直線位相直交非分離ウェーブレット変換と歪の相関,” 第 20 回回路とシステム軽井沢ワークショップ, pp.161-166, 04 2007.
- [10] G. Strang, and T. Nguyen, Wavelets and Filter Banks, Wellesley Cambridge Pr, 1996.
- [11] S. Orintara, T. D. Tran, P.N. Heller, and T. Q. Nguen, “Lattice Structure for Regular Paraunitary Linear-Phase Filterbanks and M-Band Orthogonal Symmetric Wavelets,” IEEE Trans. on Signal Proc., vol.49, no.11, pp.2659-2672, November 2001.