

ブロック変換に基づく画像非可逆圧縮 におけるウェーブレット変換

A Block-based Lossy Image Compression Using The Wavelet Transform

舞鶴工業高等専門学校 芦澤恵太* (Keita Ashizawa)

National Institute of Technology, Maizuru College

名城大学大学院都市情報学研究科 原田卓弥 (Takuya Harada)

Graduate School of Urban Science, Meijo University

名城大学都市情報学部 山谷克 (Katsu Yamatani)

Faculty of Urban Science, Meijo University

§1. はじめに

高効率な画像圧縮の実現には、信号のもつ相関を利用し、より冗長性の低い信号に変換することが必要である。1990年代以降、JPEG標準方式をはじめとし、デジタル画像の圧縮には、依然としてブロック単位の離散コサイン変換(DCT)に基づく方式が広く用いられている[1]。この場合、画像を再構成する為の基底がブロック全体を覆う為、ブロック内に局在する特徴(例えばエッジ)の近傍で生じる歪による画質劣化を避けることが出来ない。また、変換係数の値をゼロ近傍に集中させるという意味では、必ずしもDCTが最善の選択とは限らないことに着目し、周波数変換に用いる直交変換の選択肢をブロックごとに多様化することによって、より高効率な圧縮性能を実現する方法が模索されている[2, 3, 4]。そこで我々は、基底が局在する離散ウェーブレット変換(DWT)に着目した。これまでに、矩形波を基底とし最も単純なDWTであるハール変換(HT)による基礎検討を重ねている。

* 〒625-8511 京都府舞鶴市字白屋 234 番地, E-mail: ashizawa@maizuru-ct.ac.jp

近年、我々が目にする画像自体が変化してきている。4K 規格のディスプレイやビデオカメラの市販の拡大に見られるように、高精細画像は医療・博物館・広告といった高品質な画像が要求される分野のみならず、家庭にも普及してきている。また、8K UHD が ITU-R 勧告で定められるなど、高精細画像の普及は、世界的に今後もさらに加速すると予想される。

本稿では、ブロック単位で周波数領域への変換を行うという制約下においてウェーブレット変換を利用するアプローチとして我々が提案している、組み合わせ式直交変換 (COT) について報告する [5, 6]。ブロック単位での変換を基本とするのは、研究の発展として高精細画像における並列処理を意識しているだけでなく、動画像処理における予測符号化への組み込みを今後検討していくためである。

§2. で我々が考える高精細画像について整理を行い、§3. で組み合わせ式直交変換の概略について述べた後、2つの新たなアイデアについて報告する：

- ・ 1次元 (垂直もしくは水平) 処理後の直流成分に限定し DCT と HT を組み合わせる。
- ・ DCT と HT とアルパート変換 (AT) を組み合わせる。

§4. で数値実験により画像圧縮に適用した際の性能予測について述べ、§5. でまとめる。

§2. 高精細画像と多画素画像

現在は、表示装置の開発に対し、伝送技術や撮影機器も含めたコンテンツの提供能力が遅れている為、従来技術で撮像された画像を並べた多画素画像を目にすることが多い。表示装置の画素数の増加に伴い表示対象領域を広げることは、画像の高精細化には直結しない。言い換えるならば、多画素数とは、他階調などと並び、高精細な画像を構成する為の1つの指標でしかない。本研究で意味するところの高精細画像の概念図を図1に示す。JPEG 標準方式が規格化された当時は、画質の評価尺度である PSNR 値が 35[dB] 程度の



図1 従来画像の $N \times N$ 領域の表現に高精細画像ではより多く $2N \times 2N$ の標本値が必要

圧縮画像であれば、視覚的に劣化を感じないといわれていた。しかしながら、本研究を進める上では、高精細画像圧縮では従来よりも高い画質 (35[dB] 以上) が求められることに着目し、PSNR 値の上限を 42[dB] まで考えることとした。

§3. ブロック単位での基底系選択

§3.1. 組み合わせ式直交変換方式

文献 [5] では、JPEG 標準等で用いられている 2 次元 DCT に加え、DCT と HT の組み合わせによる直交変換および非標準型 HT をブロック単位で適応的に用いる方式 (DHCT 方式) について報告している。図 2 は、 8×8 画素ブロックを対象とする場合の DHCT で用いる基底系を表している。図 2 の DCT-DCT は 2 次元 DCT の基底系を表しており、また DCT-HT および HT-DCT は、それぞれ、DCT-DCT において、水平方向のみを 1 次元 HT に、垂直方向のみを 1 次元 HT に置き換えて構成した基底系である。さらに NHT は非標準型 HT の基底系を表している。

DHCT 方式では、これらの基底系に基づく 4 通りの変換をそれぞれ適用し、変換係数の 1-ノルムの値が最小となる変換をブロック単位で選択する。これは、一定の情報量の下での量子化歪、および、圧縮後の情報量をできるだけ小さくするという意味では、変換係

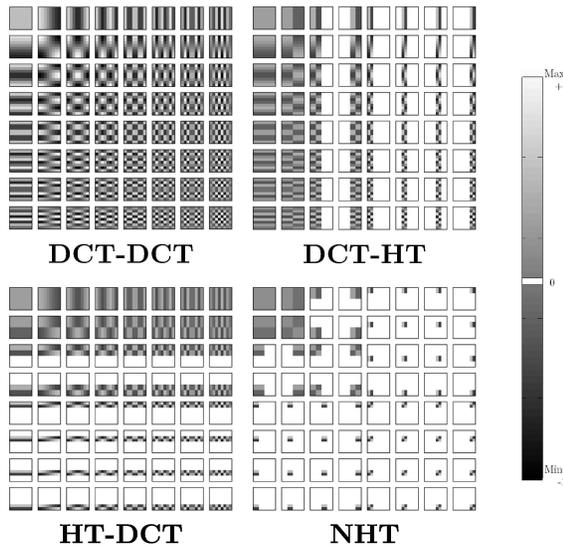


図 2 DHCT の基底系

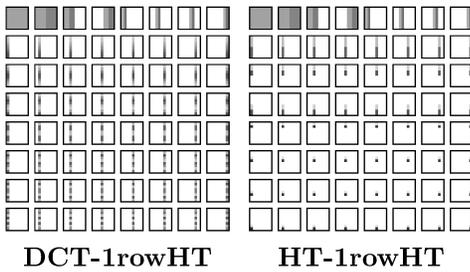
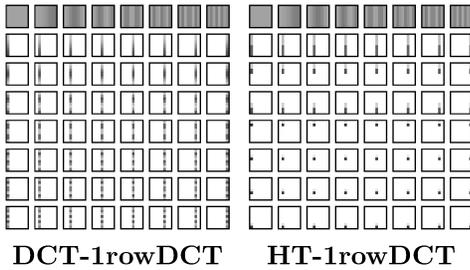


図3 水平方向の処理を限定した基底系

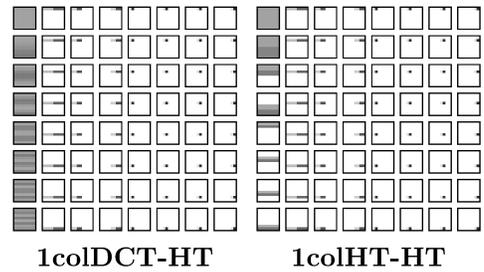
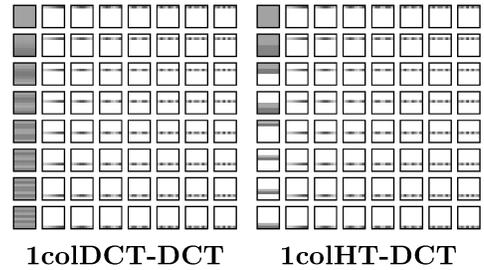


図4 垂直方向の処理を限定した基底系

数の値がゼロ近傍に集中しているほうが良いためである。以下では、異なる直交変換を垂直・水平方向に組み合わせた変換をブロック単位で適応的に用いる方式を、組み合わせ式直交変換（COT：Combinational Orthogonal Transform）と呼ぶこととする。

§3.2. 直流成分に限定し DCT と HT を組み合わせる

我々は、より解像度の高いエッジの表現を目的とし、DCT 係数の直流成分にのみ HT を適用する場合を考慮することによる DHCT の拡張を試みた。エッジのような不連続性を含む 2 次元画素ブロックの直交変換に対し、垂直（水平）方向に直交変換を適用した後、直流成分にのみ水平（垂直）方向に直交変換を適用するという考え方は、解像度の高いエッジの表現以外にも効果的となることがわかった。垂直方向の DCT 係数を水平方向に見たときに、直流成分はなめらかに分布するが交流成分については成分間の相関が低い場合もあり、このようなときは直流成分にのみ DCT を適用するのが効果的である。また、最初の処理に DCT ではなく HT の採用することによって、さらに多様な画素ブロックに対する適応性を高めることが可能となることにも気がついた。

実際の基底系の構成方法について考える。図 2 の **DCT-HT** と **HT-DCT** の基底系に着目する。基底系 **DCT-HT** は、まず垂直方向に DCT を適用し、さらに水平方向に HT を適用することによって構成される。このとき、最初の DCT によって得られる変換係数

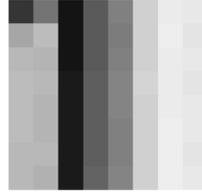


図5 エッジを含む画素ブロックの例

に対し、第1行目の直流成分のみにHTを限定的に適用するものとする、1画素単位の基底を含む新たな基底系に基づく変換係数を得ることができる。基底系HT-DCTについても、垂直と水平を入れ替えれば、同様に変換係数を得ることができる。

DCTとHTを用いて、垂直方向処理後に第一行目の直流成分に対してのみ垂直方向の直交変換を行うことで得られる基底系を図3に示す。同様に水平方向処理後に第一列目の直流成分に対してのみ水平方向の直交変換を行うことで得られる基底系を図4に示す。

例えば、図3のDCT-1rowDCTは、 8×8 画素ブロックの信号に対して、垂直方向にDCTを適用する。この処理により、1行目に 1×8 個の直流成分が得られ、2~8行目に 7×8 個の交流成分が得られる。第2の処理において、2~8行目の交流成分に対しては水平方向のHTを適用せず、1行目の8個の直流成分にのみ水平方向のHTを適用するものとする、明らかに、基底の一辺の大きさの最小値が1画素となっていることが分かる。DCT-1rowDCTは、図5に示すような特徴を持った画素ブロックを効果的に表現することを意図した。ここで図5は、標準画像Cameramanから 8×8 画素ブロックを切り出したものである。従来のDHCTでは、図2のDCT-HTに基づく変換係数の1-ノルムの値が最小となり、小数点以下を切り捨てると2966となった。それに対し、DCT-HT1に基づく変換係数の1-ノルムの値は2793となった。この結果は、水平方向に最小で2画素単位の基底を用いた場合と比較し、1画素単位の基底を用いた方が、水平方向の不連続性を効果的に表現することができることを示唆している。

§3.3. DCTとHTとアルパート変換(AT)を組み合わせる

本節では、COT方式の一つの発展として、組み合わせる直交変換を3種類とすることを検討する。我々は、マルチウェーブレットの一つであり、HTの一般化としての側面をもつアルパートウェーブレット変換(AT)を新たに組み合わせることとした。HTは局在する矩形波を基底に持つため、エッジを有するブロック信号を効果的に圧縮できた。ATの導入は、局在する線形波を基底に持つATを利用することで、部分的にグラデーショ

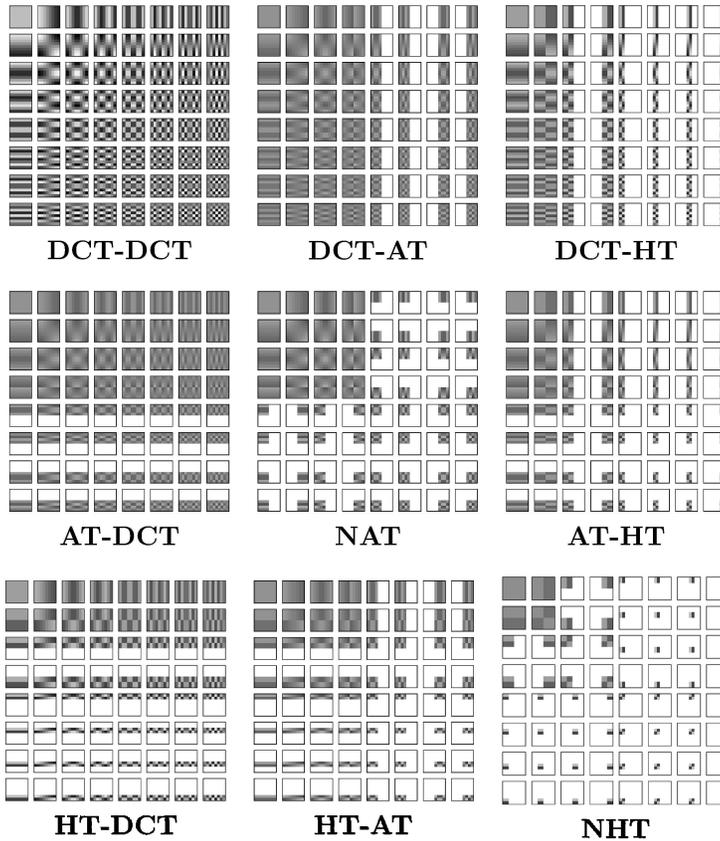


図6 3種の直交変換 (DCT, AT, HT) を組み合わせた基底系

ンしているブロック信号への対応を意図したものである。ATの数学的な議論は、文献[7, 8]に預け、本稿では画像圧縮への応用可能性を検討する。特に本研究で用いる低次のATは、Slant-Haar変換として1974年に報告されており[9]、近年では大見・吉田によって画像圧縮に適用する試みも報告されている[10]。図6に3種類の直交変換の内2種類を縦横に組み合わせることで構成可能な9種類の基底系を示す。

図6の四隅に示す基底系 (DCT-DCT, DCT-HT, HT-DCT, NHT) は、DHCTで用いる基底 (図2参照) と同一のものである。これらを除いた5種類がATを組み合わせることで新たに構成された基底系である。これまでと同様に、垂直方向と水平方向の処理において、直交変換を適用することでそれぞれの基底系が構成される。ここで、NATは非標準型ATの基底系を表している。図6より、ATを組み入れた場合、局在する基底の幅は最小でも4画素となることが見て取れる。次節では、画像圧縮において、これら9

つの基底系をブロック単位で適応的に用いることの有効性を検討する。

§4. 圧縮性能の予測

本節では、§3.2. および §3.3. で提案した COT 方式の 2 つのアイデアの画像圧縮における有効性を検討する。圧縮方式の実装においては、符号化などの詳細な検討が必要となるが、ここでは図 3, 図 4 および図 6 で提案した基底系が有意であるかを検討する為、次節に示すアルゴリズムを用いて、2 つの数値実験の結果を報告する。

§4.1. アルゴリズム

以下では、入力された元画像は $N \times N$ 画素ブロック単位に L 個のブロックに分割されているものとする。

最初に、 $\ell (= 1, 2, \dots, L)$ 番目の画素ブロック信号に対する順変換のアルゴリズムを図 7 に示す。順変換では、まず当該画素ブロックに対し m 通りの基底系に基づく変換係数が計算され、次に変換係数の 1-ノルムが最小となる変換が選択される。ここで、 m は選択候補の数とする。3.1 節で述べた DHCT では、図 2 に示す基底系が候補であるため、 $m = 4$ となる。また、3.2 節で提案した変換では、図 2 に図 3 および図 4 を加え 12 通りとなるため、 $m = 12$ となる。さらに、3.3 節で提案した AT を追加した場合は図 6 に示す 9 通りが候補となるため、 $m = 9$ となる。

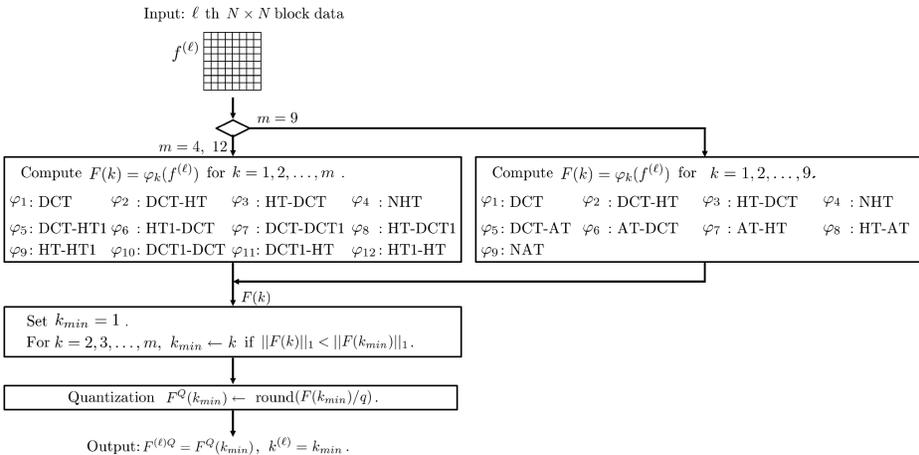


図 7 順変換

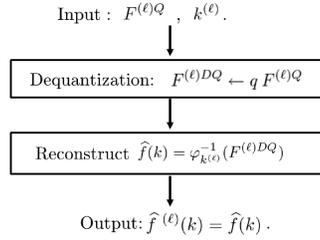


図 8 逆変換

図 7 の φ_k は、図 2、図 3、図 4 および図 6 に示した基底系に基づく、各直交変換を意味する。順変換における入力は ℓ 番目のブロックにおける画素ブロック信号 $f^{(\ell)}$ であり、出力は 1-ノルムが最も小さくなるように選択された変換を識別する変数 $k^{(\ell)}$ および量子化された変換係数 $F^{(\ell)Q}$ である。図 7 の q は量子化ステップを意味する。

ℓ 番目のブロック信号に対する逆変換のアルゴリズムを図 8 に示す。逆変換においては、量子化された変換係数 $F^{(\ell)Q}$ と選択された変換の識別変数 $k^{(\ell)}$ を入力し、逆量子化された変換係数 $F^{(\ell)DQ}$ に対して $k^{(\ell)}$ 番目の順変換に対応する逆変換 $\varphi_{k^{(\ell)}}^{-1}$ を適用する。次節以降の数値実験においては、変換係数の量子化には一様量子化を用いるものとする。すなわち、図 7 における量子化ステップ q には、全変換係数に対して一定の値を用いている。また、圧縮画像の再構成に要する情報の量を次式の 1 画素単位の平均情報エントロピー $H^{[m]}$ を用いて算出する：

$$H^{[m]} = H_{\text{DC}}^{[m]} + H_{\text{AC}}^{[m]} + H_k^{[m]}, \quad m = 1, 4, 6, 12.$$

ただし、 $H^{[m]}$ の単位は bpp とし、

$$H_{\text{DC}}^{[m]} := -\frac{(X/N) \times (Y/N)}{X \times Y} \sum_{A \in \Omega_{\text{DC}}^{[m]}} P(A) \log_2 P(A),$$

$$H_{\text{AC}}^{[m]} := -\frac{(X \times Y) - ((X/N) \times (Y/N))}{(X \times Y)} \sum_{A \in \Omega_{\text{AC}}^{[m]}} P(A) \log_2 P(A),$$

$$H_k^{[m]} := -\frac{(X/N) \times (Y/N)}{(X \times Y)} \sum_{A \in \Omega_k^{[m]}} P(A) \log_2 P(A)$$

とする。上式において、 $\Omega_{\text{DC}}^{[m]}$ と $\Omega_{\text{AC}}^{[m]}$ 、および $\Omega_k^{[m]}$ は、それぞれ $F^{(\ell)Q}$ の直流成分と交流成分、および各画素ブロックにおいて選択された変換を識別するための変数 $k^{(\ell)}$ の分布集合を表す。また、 $P(A)$ はそれぞれの集合における各要素の出現確率を表している。

ここで、 X と Y は、入力画像の垂直方向と水平方向の画素数を表す。圧縮画像の画質は、次式の PSNR 値を用いて評価する：

$$\text{PSNR} = 20 \log_{10}(255/\text{RMSE}).$$

ただし、RMSE は原画像と圧縮画像の平均二乗誤差の平方根を表すものとする。

§4.2. 数値実験 1：変換を直流成分に限定し基底系を多様化

標準画像を結合して作成した実験画像（図 9）に対し、同一情報量に対する圧縮画像の PSNR 値を比較する。ここでは、JPEG 標準に準拠し、ブロックサイズは 8×8 とする。全てのブロックにおいて基底系 **DCT-DCT** に基づく変換を施した場合を基準に、前節で示した $m = 4$ と $m = 12$ を比較することで、画像内の 1 画素幅の変化に対応するように基底系を多様化する効果を検証する。図 10 に、 $m = 4, 12$ の PSNR 値から $m = 1$ の PSNR 値を引いた値を示す。なお、図 10 の横軸は、 $m = 1$ とした場合の再構成画像の画質を目盛りに情報量を表している。なお、図 10 において、下限である $m = 1$ による再構成画像の画質が $30[\text{dB}]$ となるときの情報量は $0.91[\text{bpp}]$ 、上限である $40[\text{dB}]$ のときの情報量は $2.42[\text{bpp}]$ である。

図 10 より、ここで示した圧縮帯域において、 $m = 12$ による圧縮画像の方が $m = 4$ と比較し、 $0.2[\text{dB}]$ 程度の画質の改善が見て取れる。これは、あらゆる特徴をもつ画像に対して、変換を直流成分に限定することによる基底系の多様化の効果が期待できることを示唆している。



図 9 実験画像 1 (768×768 画素)

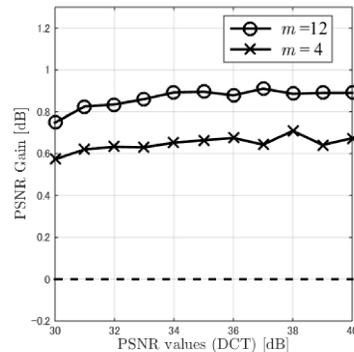


図 10 同一情報量に対する画質比較

§4.3. 数値実験 2：高精細画像への対応

図 13 に、図 11 および図 12 の SCID 高精細標準画像に対し、ブロックサイズを 64×64 とした場合の画質比較を示す。一般に、高精細画像には高画質が要求される。そこで、圧縮の帯域としては、 $m = 1$ の場合における再構成画像の PSNR 値が $37[\text{dB}]$ から $42[\text{dB}]$ という、人間の視覚では劣化が慎重に観察しなければ見て取れない範囲での検証を行う。図 10 と同様に、図 13 の縦軸は $m = 1$ との PSNR 値の差であり、横軸は $m = 1$ による再構成画像の画質を用いた情報量である。比較には、 $m = 4, 12$ に、AT を組み合わせた基底系を用いる $m = 9$ を加える。

図 13 では、すべての画像に対して、PSNR 値の差が正の値を取ることから、ブロックごとに基底系を選択することが効果的であることがわかる。しかしながら、 $m = 1$ に対する $m = 4$ と $m = 12$ の PSNR 値の改善幅が、ほぼ一致していることが見て取れる。このことから、選択率の比較などの詳細な検討が必要ではあるものの、高精細画像に対しては直流成分に限定した基底系を用意することは意味を持たない、言い換えるならば図 2 に示



Fishing goods



Silver



Woman with glass

図 11 実験画像 2 (SCID, 4096×3072 画素)

Flowers



Japanese goods



Threads

図 12 実験画像 3 (SCID, 3072×4096 画素)

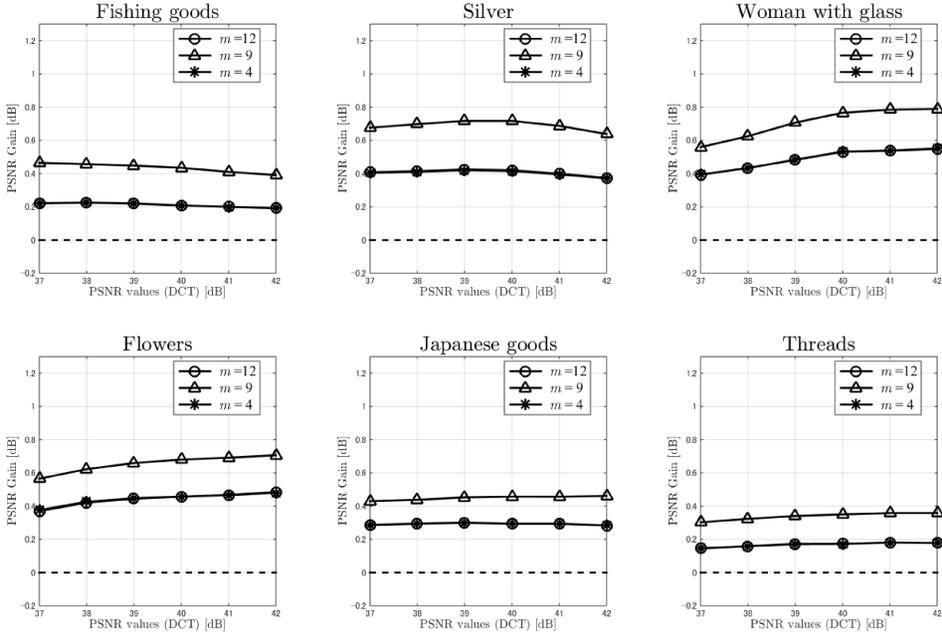


図 13 SCID 画像における 64×64 画素ブロック単位での基底系選択の効果

した基底系で十分であることが推察される。 $m = 12$ が効果的でない一方で、AT を組み合わせた基底系に対応する $m = 9$ では、 $m = 4$ と比較し約 $0.2[\text{dB}]$ の PSNR 値の向上が見て取れる。これは、図 6 で新たに加えた 5 種類の基底系の効果であると考えられる。

§5. まとめ

本稿では、ブロック単位で周波数領域への変換を行うという制約下において、ウェーブレット変換を利用する 2 つのアプローチを提案した。

提案アプローチで新たに用いた 2 つの基底系は、我々が既に提案している DCT と HT を組み合わせることによって得られる基底系を発展させてのものであり、組み合わせ式直交変換と名付けた枠組みの一例となっていることを述べた。 §3.2. で述べた直流成分に限定し DCT と HT を組み合わせた基底系は、画像内のより詳細な特徴（例えば幅が 1 画素のエッジ）の表現を狙いとしており、 §4.2. の数値実験でその画像圧縮における有効性を示した。

§3.3. では、異なる 3 種類の直交変換を組み合わせることを試みた。 HT の一般化としての側面をもつアルパート変換に着目することで、 §4.3 に示したように高精細画像の圧縮

に対しても有効な基底系の構成を実現した。

本稿で報告した2つのアプローチは、基礎研究の段階である。画像圧縮アルゴリズムとしての実装に向けては、符号化の詳細な検討や基底系選択ルールの検討などの課題が残る。しかしながら、直流成分に限定して変換を組み合わせるアプローチ、および、3種以上の基底系を組み合わせるアプローチは、多くのウェーブレット変換と組み合わせ可能であり、ウェーブレット変換を画像圧縮に応用していく端緒となり得るアイデアであると考えている。

参考文献

- [1] W.B. Pennebaker and J.L. Mitchell : “JPEG Still Image Data Compression Standard”, Van Nostrand Reinhold, New York (1993)
- [2] 亀田昌志, 白川 将寛, “画像のサブバンド符号化における符号化レートに応じた最適帯域分割”, 映情学誌, 65, 11, pp.1631–1635 (2011)
- [3] 對馬直哉, 亀田昌志, “離散コサイン変換における DCT 係数の符号情報保存に注目した高効率画像符号化”, 映像学技報, 37, 12, pp.9–12(2013)
- [4] 川村和也, 亀田昌志 : “局所特徴の保存に着目した ICA と DCT の組み合わせによる新たな静止画像符号化方式”, 映情学技報, 108, pp.25–28(2014)
- [5] 芦澤恵太, 山谷克 : “DCT とハール変換を縦横に組み合わせた新たな周波数変換方式の提案と画像圧縮におけるモスキートノイズの軽減”, 信学論 (A), J96-A, 7, pp.484–492 (2013)
- [6] 森田雅貴, 芦澤恵太, 山谷克 : “DCT と HT に基づく組み合わせ式直交変換における演算量の軽減方法とその高精細画像圧縮への応用”, 都市情報学研究, 19, pp.81–92 (2014)
- [7] B. Alpert : “A class of bases in L_2 for the sparse representation of integral Operators”, SIAM J. Math. Anal. 26, pp.246–262 (1993)
- [8] 芦野隆一 : “マルチウェーブレットの構成と応用”, 数理解析研究所講究録, 1622, pp.18–36 (2009)
- [9] W.K.Pratt, W-H.Chen, L.R.Welch : “Slant Transform Image Coding”, IEEE Trans. Commun., COM-22, pp.1075–1095 (1974)
- [10] 大見克麿, 吉田茂 : “Slant-Haar 変換符号化の画像圧縮特性についての検討”, 科学技術フォーラム 2012, I-020 (2012)