

画像のガウス混成モデルに基づく DWT 符号化法

片岡康文

山根延元

森川良孝

岡山大学 工学部

700-8530 岡山市津島中 3-1-1

Phone. 086-251-8128

E-mail. yamane@cne.okayama-u.ac.jp

あらまし 画像の DCT 符号化における符号化能率の改善法として、ガウス混成モデルに基づく適応的のエントロピー符号化法が提案されている。本論文では、この方法を離散ウェーブレット変換 (DWT) 符号化に適用する。ガウス混成モデルを DWT 係数に整合させるため、まずその性質に見合う係数ベクトルの構成法について検討を行う。次に、DWT の高空間分解能による強い非定常性に適応するため、非定常混成モデルを採用する。更に、この非定常モデル化に伴い、変則的なモデル化が行われることを避けるため、正則化法を導入する。最後に計算機シミュレーション実験結果を示し、提案法は従来の SPIHT 法に比べ高符号化能率であることを明らかにする。

キーワード 離散ウェーブレット変換, ガウス混成モデル, エントロピー符号化

DWT Coding Method for Images based on Gaussian Mixture Distribution Model

Yasufumi KATAOKA

Nobumoto YAMANE

Yoshitaka MORIKAWA

Faculty of Engineering, OKAYAMA University

3-1-1 Tsushima-naka, Okayama-shi 700-8530, Japan

Phone. (+81)86-251-8128

E-mail. yamane@cne.okayama-u.ac.jp

Abstract In this paper, an adaptive entropy coding method based on the Gaussian mixture distribution model, proposed as a performance improving method for the DCT coding for images, is adopted to the discrete wavelet transform(DWT) coding. To modify the mixture model adequate for the DWT coefficients, construction of the coefficient vector is studied. Unstationary mixture model is employed to adapt the model to fine spatial regularization of high-frequency coefficients and, then, the regularization method is introduced to suppress the irregular modeling caused by the unstationary modeling. Simulation result shows that proposed method improves coding performance compared with conventional SPIHT method.

key words discrete wavelet transform, Gaussian mixture distribution model, entropy coding

1 はじめに

本論文では、静止画像の離散ウェーブレット変換 (DWT) 符号化法 [1] として、ガウス混成モデルに基づく適応エントロピー符号化法を提案する。本論文では画像の局所領域に対応する DWT 係数をまとめたベクトルに対し、ガウス混成モデルを当てはめる。本論文では、各 DWT 係数ベクトル毎に平均符号長最小となるガウス分布用 PCM 符号化器のレート割当てを行った場合のシミュレーション結果を示し、有効性を明らかにする。

2 ガウス混成モデル

提案法では、画像の局所領域における DWT 係数をまとめて後の図 2 に示すようにベクトル化し、このベクトルを \mathbf{x} で表す。このとき、各成分の平均は零、各成分は互いに無相関で分散 σ^2 をもつと仮定する。すなわち、

$$E[\mathbf{x}] = \mathbf{0} \tag{1}$$

$$E[\mathbf{x}\mathbf{x}^T] = \text{diag} [\sigma^2(1), \sigma^2(2), \dots, \sigma^2(N_i)] \tag{2}$$

但し、 N_i は図 2 に示すタイプ V_i のベクトルの次元数を示す。タイプ V_i のベクトル \mathbf{x} を発生する情報源として、図 1 に示す混成情報源を仮定する。但し、 $\{s_{i,k}\}_{k=1,2,\dots,C}$ はベクトル \mathbf{x} を独立同分布に発生するガウス定常過程を示し、 C 個の異なる共分散行列 $\{\Lambda_{i,k}\}_{k=1,2,\dots,C}$ で特徴づけられるものとする。すなわち、確率密度関数 $p(\mathbf{x}|s_{i,k})$ が次式で与えられる。

$$p(\mathbf{x}|s_{i,k}) = ((2\pi)^{N_i} |\Lambda_{i,k}|)^{-1/2} \exp\left(-\frac{1}{2}\mathbf{x}^T \Lambda_{i,k}^{-1} \mathbf{x}\right) \tag{3}$$

この混成モデルの構成法として、トレーニング系列の画像を用い、共分散行列 $\{\Lambda_{i,k}\}_{k=1,2,\dots,C}^{i=0,1,2,3}$ を最尤推定法に基づく逐次的なアルゴリズムにより決定する Wornell の方法 [3] が知られている。このガウス混

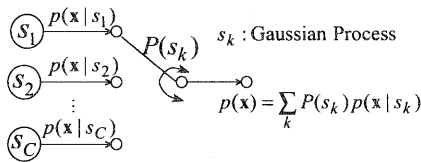


図 1: ガウス混成モデルの説明図

成モデルを DWT 係数列に適用する際には、DWT 係数ベクトルをどのように構成するかを決定することが重要な課題となる。次章では、DWT 係数ベクトルの構成法について述べる。

3 DWT 係数列のベクトル化とガウス混成モデルの構成

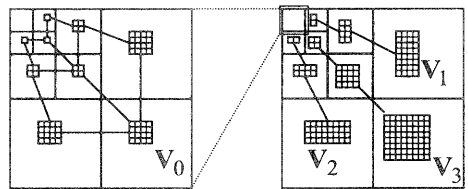
本章では、まず、DWT 係数列の統計的性質を調べて係数ベクトルを構成する方法の検討を行い、次に、混成モデルの構成法を示す。

3.1 DWT 係数列のベクトル化

図 3 に後の実験で用いる画像 “boats” に 3 階層の DWT を施した画像を示す。但し、DWT 係数の絶対値の対数を 30 倍して示している。LH 成分は垂直方向の、HL 成分は水平方向の高周波成分をもっているため、それぞれ水平方向、垂直方向のエッジが現れている。HH 成分は主に斜め方向のエッジが現れている。また、最高周波数成分では急峻なエッジの信号電力が 2 画素程度の幅で局在しており、非常に強い非正常性をもつことがわかる。

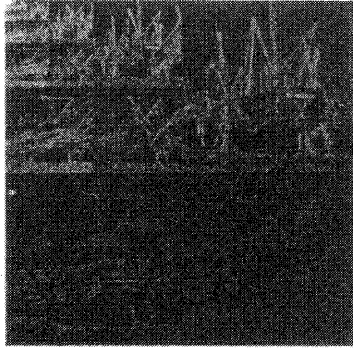
図 2 に提案法における DWT 係数のベクトル化の様子を示す。このベクトル化は以下に示す 3 つの条件を満たすように決定したものである。

- (1) 帯域の決定：ベクトル化は、画像の局所領域に対応する係数列を成分に選んで行う。図 3 から分かるように LH_i 成分、HL_i 成分および HH_i 成分の係数電力は、それぞれ、階層間で強い相関をもっているため、各成分毎に階層にわたってベクトル化を行う。但し、全階層にわたってベクトル化を行えば、ベクトルサイズが非常に大きくなるため低周波成分と高周波成分に分けてベクトルを構成する。

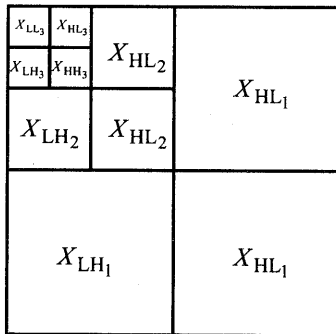


(a) 低周波成分(4~6階層) (b) 高周波成分(1~3階層)

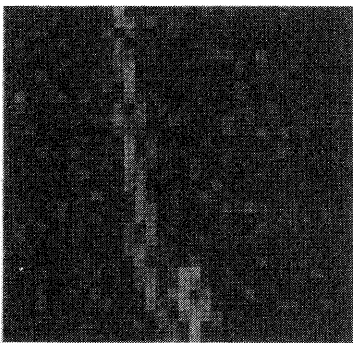
図 2: DWT 係数のベクトル化



(a) 係数画像



(b) 帯域分割の説明図



(c) (a) の HL_1 成分の拡大画像

図 3: DWT 係数画像の例

(2) 形状の決定：上で述べたように LH_i 成分には水平方向の、 HL_i 成分には垂直方向のエッジがそれぞれ現れるため、これらに見合った形状を選ぶ。

(3) ベクトルサイズの決定：DWT 係数の高周波成分は強い非定常性をもつため、適応性の観点からはベクトルサイズは小さい方が望ましい。一方、適応化のための付加情報のレートはベクトルサイズを小さく選べば比重が大きくなる。従来の DCT 符号化法では、ブロックサイズを概ね 8×8 画素に選んだ場合の符号化特性が最適になることが知られており、提案法においても、各ベクトルのサイズは次元数が 64 に近い値を取るように選ぶ。

タイプ V_i のベクトルの次元数 N_i は、それぞれ、 $N_0 = 63$, $N_1 = N_2 = 42$, $N_3 = 84$ である。

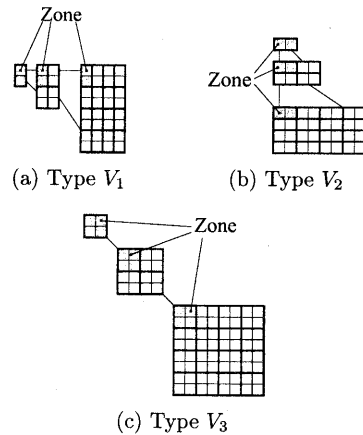


図 4: DWT 係数ベクトルのゾーン分割

3.2 共分散行列の決定

前節で述べたガウス混成モデルは、DWT 係数ベクトルの各要素の分散が全て異なるという意味で非定常なモデル化が可能になっている。先の図 3 に見られるように、画像の急峻なエッジの信号電力は、特に最高周波数成分においては 2 画素幅程度の細い帯状に分布しており、ベクトル内でのエッジの位置のずれに適応化を図るためには非定常なモデルが必要となると考えられる。しかし、トレーニング画像列を用い

て、混成モデルの共分散行列の集合 $\{\mathbf{A}_{i,k}\}_{k=1,2,\dots,C}^{i=0,1,2,3}$ を決定する場合、非定常モデルではトレーニング画像の高周波成分に含まれる雑音の影響により、不要なクラスが用意され、モデルの性能を低下させることになる。そこで、提案法では各タイプの DWT 係数ベクトルを、図 4 に示すようにゾーンに分割し、各ゾーン内の DWT 係数は広義定常過程と仮定してモデル化を行う。このとき、第 k 番目のクラスに属するタイプ V_i の DWT 係数ベクトルの共分散行列 $\mathbf{A}_{i,k}$ は次式で与えられる。

$$\mathbf{A}_{i,k} = \text{diag}[\sigma_{i,k}^2(z_i(1)), \sigma_{i,k}^2(z_i(2)), \dots, \sigma_{i,k}^2(z_i(N_i))] \quad (4)$$

但し、 $z_i(n)$ は第 i 番目のベクトルの第 n 成分が属するゾーンの番号を示す。

上式の共分散行列の集合 $\{\mathbf{A}_{i,k}\}_{k=1,2,\dots,C}^{i=0,1,2,3}$ の決定は Wornell の方法 [3] と同様な逐次的アルゴリズムを用いて行う。但し、提案法ではトレーニング系列に多種多様な画像列を用いて汎用的な混成モデルを構成するため、事前確率 $P(s_{i,k})$ を一定と仮定する。

以下に、共分散行列集合の決定法の手順を示す。

[共分散行列の決定手順]

1. 初期分割：トレーニング画像の系列に DWT を施し、図 3 に示す各タイプ $V_{0,1,2,3}$ 毎にベクトル化し、各タイプのトレーニングベクトルとする。各トレーニングベクトルをそれぞれ C 個のクラス $s_{i,k}$ に初期分割とする。
2. 共分散行列の推定：次式により式 (4) の $\mathbf{A}_{i,k}$ の第 m 対角要素 $\sigma_{i,k}^2(m)$ を推定する。

$$\hat{\sigma}_{i,k}^2(m) = \frac{\sum_{\mathbf{x} \in s_{i,k}} \sum_{n|z_i(n)=m} x^2(n)}{\sum_{\mathbf{x} \in s_{i,k}} \sum_{n|z_i(n)=m} 1} \quad (5)$$

3. 再分類：各タイプ毎、トレーニングベクトルを次式の出現確率 $\hat{p}(\mathbf{x}|s_{i,k})$ を最大とするクラスに再分類する。

$$\hat{p}(\mathbf{x}|s_{i,k}) = \frac{1}{\sqrt{(2\pi)^{N_i} |\mathbf{A}_{i,k} + \mathbf{I}\delta^2|}} \times \exp\left(-\frac{1}{2} \sum_n \frac{x^2(n) + \delta^2}{\sigma_{i,k}^2(z_i(n)) + \delta^2}\right) \quad (6)$$

4. 再分類によりベクトルが移動しなくなるまで手順 2,3 を繰り返す。

手順 3. において δ^2 は、非定常モデルの採用に伴って不要なモデル化が行われることを防ぐために導入した正則化定数を示す。低レート符号化時の高周波成分の係数には零ビット割当てが行われ、このような高周波信号は雑音と見なすことができる。この雑音の影響は正則化定数を適切に選ぶことにより抑圧できる。

4 提案法

本章では提案法の構成と処理手順について述べる。

図 5 に提案法の構成を示す。本法では、画像の DWT 系列に前章で述べたベクトル化を行い、各ベクトルの統計的性質に応じたガウス分布用 PCM 符号化器を割当てる。

以下に、提案法各部の構成を示す。

(1) ガウス分布用 PCM 符号化器：量子化器として、ガウス分布に対する最適量子化器とほぼ同等の性能を有するミッドステップ型の最適一様量子化器を用い、その出力をガウス分布用エントロピー符号化器により符号化する。文献 [2] の従来の DCT 符号化法では、エントロピー符号化器としてハフマン符号を用いればほぼエントロピー・レートで符号化できることが知られている。

(2) レート割当表：前章で述べた DWT 係数ベクトル \mathbf{x} の共分散行列を $\text{diag}[\hat{\sigma}^2(0), \hat{\sigma}^2(1), \dots, \hat{\sigma}^2(N-1)]$ とする。 \mathbf{x} の各要素に対するレート割当て $\{b(0), b(1), \dots, b(N-1)\}$ は次式で与えられる。

$$b(n) = \max\left[f\left(\left\lfloor \frac{\hat{\sigma}(n)}{Q\tau} + 0.5 \right\rfloor \tau\right), 1.0\right] [\text{bits}] \quad (7)$$

但し、 $f(\cdot)$ は標準偏差が \cdot のガウス過程を単位ステップ幅で量子化したときのエントロピー符号化器のレート、 Q は量子化ステップ幅を示す。また τ はレート割当ての精度を定める定数であり PCM 符号化器の個数を低減するために導入している。また、与えられた平均レートのもとで、ひずみ電力を最小とするレート割当ては、量子化出力エントロピーが 1 ビット以下となる係数に零ビットを割り当てることで近似的に実現できることが知られており、提案法においてもこの方法を採用している。

上で述べたレート割当て処理を、与えられた量子化ステップ幅 Q に対して各タイプの各クラス毎に求めてレート割当表 (CAT) を作成する。また、付加情報削減のため、レート割当てが一致したクラスは統合する。

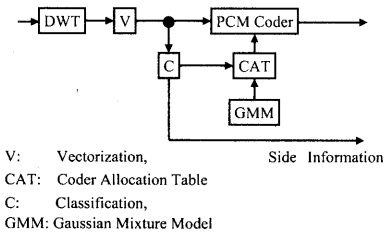


図 5: 提案法の構成図

表 1: 正則化定数 δ

量子化幅 Q	δ
0 ~ 15.0	0.1
15.0 ~ 37.0	5.0
37.0 ~ 50.0	10.0
50.0 ~ 65.0	15.0
65.0 ~	30.0

以下に、提案法の符号化処理手順を示す。本法では、予め、トレーニング系列の画像を用いて、ガウス混成モデルの共分散行列の集合を決定し、これを符号・復合器共に記憶しておく。

【符号化処理手順】

1. クラス分け：画像に DWT を施しベクトル化を行う。各ベクトルの統計的性質に応じ、全てのベクトルを混成モデルの要素過程 $s_{i,k}$ のいずれかにクラス分けする。
2. 符号化：現ベクトル \mathbf{x} が属するクラス $s_{i,k}$ に対するレート割当表を参照し、 \mathbf{x} を割り当てられたレートの PCM 符号化器に入力する。

手順 1. のクラス分けの方法として種々の方法が知られている [2] が、本論文では提案法の符号化特性の上限を調べるため、処理の負担が大きい方法ではあるが、平均符号化レートが最小となるクラス分けを行う。この方法では、上で述べた CAT の決定において零ビットの割当てが起らない程度に十分小さい量子化ステップ $Q = 1$ を選んでクラス分け用の CAT を求める。1つのベクトルを、全てのクラス分け用 CAT を用いて符号化した場合毎の平均符号化レートを求め、このレートが最小となるクラスに現ベクトルをクラス分けする。

5 計算機シミュレーション

提案法の計算機シミュレーション実験を行い、SPIHT 法 [1] との比較を行なう。

5.1 実験条件

実験に用いた画像は、SCID の標準画像 “ポートレート”, “カフェテリア”, “自転車” (2048 × 2560 画素) 及び “果物かご”, “ワインと食器”, “蘭”, “ミュージシャン”, “キャンドル” (2560 × 2048 画素) から、512 × 512 画素の輝度成分を切り出した 256 階調の濃淡画像 123 枚である。

これらの内、7枚の画像を符号化対象画像とし、残りの画像をトレーニング系列の画像として用いた。また、混成モデルの普遍性の確認のため、トレーニング系列の画像と統計的性質の大幅に異なる ISO-CCITT のテスト画像 “Barbara”, “Boats”, “Black” の輝度成分の画像 (512 × 512 画素, 256 階調) を用いる。以下に実験の条件を示す。

- DWT は SPIHT 法と同様に文献 [4] のものを用いた。
- 符号化レート：符号化レート R_{total} は次式により求める。

$$R_{total} = R + R_0 + r \quad (8)$$

ここで、 R は DWT 交流係数のレートを示し、文献 [2] と同様に、ガウス分布用エントロピー符号化器を用意して各係数を符号化した場合の平均レートを算出する。また、 R_0 は最低周波数成分の DWT 係数を DPCM 符号化したときのレートを示す。 r はクラス分け情報のレートを示し各クラスの出現頻度からエントロピーを算出する。

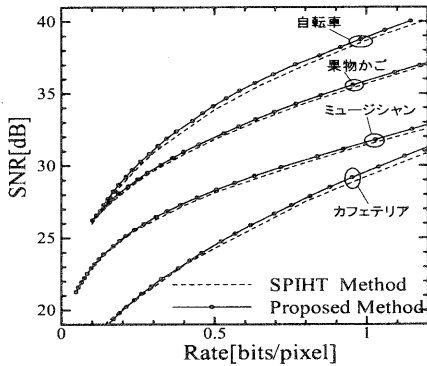
- SNR は次式で定義する。

$$SNR = 10 \log_{10} \frac{255^2}{(\text{平均二乗誤差})} \quad (9)$$

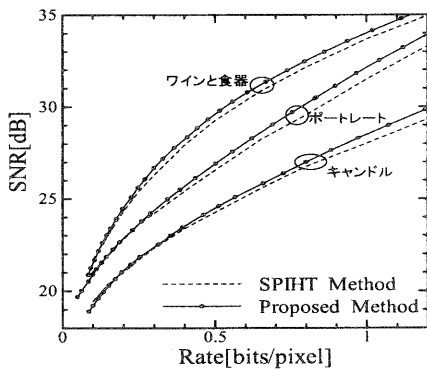
- 正則化定数 δ は、表 1 のものを用い、PCM 符号化器の割当精度の定数 $\tau = 0.05$ とする。

5.2 実験結果

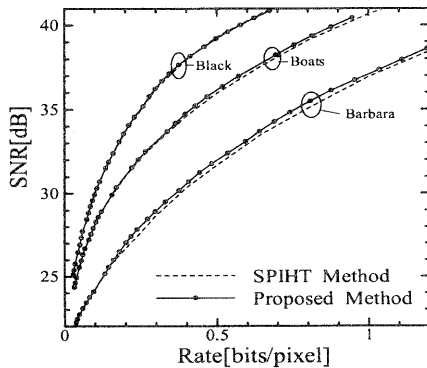
図 6 に提案法と SPIHT 法のレート-SNR 特性を示す。同図 (a), (b) には、トレーニング系列の画像と符号化対象画像の統計的性質が整合している場合における特性を示している。レートの広い範囲において提案法は SPIHT 法と同等以上の符号化特性を示し、



(a) SCID 標準画像 (1)



(b) SCID 標準画像 (2)



(c) CCITT 標準画像

図 6: 符号化特性

レートが 0.5[bits/pixel] において約 0.15 ~ 0.35[dB] だけ高い SNR が得られていることがわかる。

また、同図 (c) には、トレーニング系列と統計的性質の異なる画像における特性を示している。これらの画像においてもほぼ同等、もしくはより優れた符号化特性が得られていることがわかる。提案法で採用した普遍的混成モデルが、有効であることを示している。

特に、画像“Barbara”の低レート部分では、提案法の特長向上が大きくなっている。SPIHT法に比べ、提案法では各係数へのレート割当てを最適に行っていることに起因すると考えられる。

6 むすび

本稿では、画像の DWT 符号化において、ガウス混成モデルに基づく適応エントロピー符号化法を提案し、有効性を示した。今後の課題は、

- (1) クラス分け処理の単純化を行い、符号化処理の高速化を図ること、
 - (2) クラス数および非定常モデルの構成の最適化を行うこと、
 - (3) モデルの普遍性の向上を図ること、
- 等が挙げられる。

参考文献

- [1] Amir Said and William A. Pearlman: "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", IEEE Transactions on Circuits and Systems for Video Technology, vol. 6, pp. 243-250, June 1996.
- [2] 山根 延元, 森川 良孝, 浜田 博, 梅田 清: "画像のガウス混成モデルとエントロピー符号化法への応用", 信学論, Vol. J80-B-I, No. 7, pp. 564-572, July 1997.
- [3] Gregory W. Wornell, David H. Staelin: "Transform image coding with a new family of models", Proc. IEEE ICASSP'88, vol. 2, pp. 777-780, 1988.
- [4] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies: "Image coding using wavelet transform", IEEE Transactions on Image Processing, Vol. 2(2) pp. 205-220, 1992.
- [5] A. N. Tikhonov and V. Y. Arsenin: "Solutions of Ill-Posed Problems", Wiley (1977).