

超高精細画像における DCT 係数の分布と QM-coder を用いた符号化に関する検討

黒木 祥光[†] 上繁 義史[†] 太田 諦二[‡]

[†]鹿児島工業高等専門学校

[‡]九州工業大学工学部

あらまし 静止画像符号化の国際標準 JPEG では、必須機能として DCT (Discrete Cosine Transform) を用いた非可逆符号化を採用している。エントロピー符号化として 2 値算術符号である QM-coder を用いる際、DCT 係数の 2 値化とマルコフモデル条件付けは統計モデルによって行われる。一般に DCT 係数の AC 成分は正規分布で近似されると言われている。しかし、Reininger 等は数種類の画像に対して KS(Kolmogorov-Smirnov) 統計量を測定した結果、画素値の変化が極端に大きな画像を除いた場合、低周波成分の分布は正規分布よりラプラス分布に近いことを示している。本論文では、超高精細画像における DCT 係数の分布に関して考察し、その分布に基づいた統計モデルを提案する。

Distributions of the DCT coefficients for SHD images and a coding scheme by using the QM-coder

Yoshimitsu KUROKI[†] Yoshifumi UESHIGE[†] Teiji OHTA[‡]

[†]Kagoshima National College of Technology

[‡]Faculty of Engineering, Kyushu Institute of Technology

Abstract This paper describes a coding scheme of the DCT coefficients for the Super High Definition (SHD) images by using the binary arithmetic coder called QM-coder. A Gaussian distribution generally approximates the distributions of the AC coefficients. However, Reininger and Gibson indicate that high-energy AC coefficients are not Gaussian, but Laplacian, by computing the Kolmogorov-Smirnov test statistic. We establish a new statistical model that specifies binary decision sequence and probability conditioning procedure conformed to the distributions of the DCT coefficients. Simulation results show that proposed model provides better compression performance than the one employed in the JPEG.

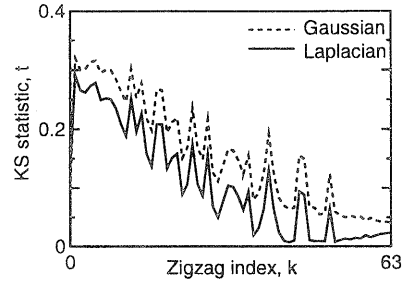
1 まえがき

本研究で使用する超高精細画像は $4,096 \times 4,096$ の画素を有し、各画素は RGB それぞれ 12 ビットで量子化された画像である。超高精細画像の符号化において、著者等はその高品質性を維持する手法が最優先と考え、可逆符号化 [1]-[3]、あるいは

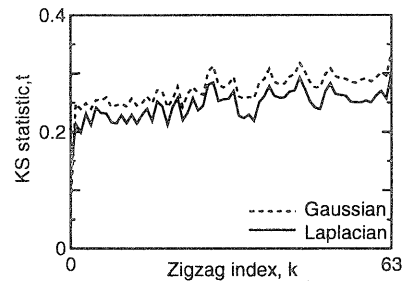
予測符号化に基づいた高画質非可逆符号化 [4] に関して検討してきた。一方、静止画像符号化の国際標準 JPEG では、必須機能として DCT (Discrete Cosine Transform) を用いた非可逆符号化を採用している [5]。高画質符号化では、一般に DCT 等の変換符号化に比べ、予測符号化の方が有効と言

われているが、量子化表の設定により高品質な符号化も可能である。また、JPEGでは低解像度の画像から徐々に高解像度の画像を再現する、いわゆるプログレッシブモードもDCTを用いており、超高精細画像をデータベース等に蓄積して検索する場合を想定すると、DCT方式も有効な手法と考えられる。従って、本研究では超高精細画像におけるDCT係数の符号化法に関して検討する。

JPEGでは、量子化されたDCT係数をエントロピー符号化する際、ハフマン符号、もしくは2値算術符号であるQM-coderが用いられる。ハフマン符号はコンパクト符号であり、且つテーブルルックアップ方式で符号化が可能であるため、JPEGでは必須なエントロピー符号化となっている。しかし、情報源の統計的性質が異なる場合、新たなテーブルを作成し、サイド情報として付加する必要がある。動的適応化は困難である。一方、QM-coderは動的適応性が高く、符号化しながらダイナミックに符号化パラメータを変化させることが容易であるため、任意の情報源に対し、常に高効率な符号化が可能である。また、既に符号化済の情報によって符号化すべき情報源を分離するマルコフモデル符号化も容易に実現され、更なる符号化効率の改善が期待できる。従って、QM-coderはJPEGのみならず2値画像符号化の国際標準JBIG (Joint Bi-level Image experts Group)[6]においても採用されている。しかし、QM-coderは2値算術符号であるため、多値情報源を符号化する際、2値化する必要がある。2値化の手法として、レベルプレーン展開等の手法があるが、各判定における2値シンボルの出現確率は異なると考えるのが自然である。また、マルコフモデル符号化を行う場合、各判定を更に独立した情報源に分離し、個々の情報源毎に“状態”を割り当て、符号化パラメータを記憶する必要がある。QM-coderは2値シンボルの出現確率の初期値をそれぞれ0.5とし、符号化しながら2値シンボルの出現確率を学習していくため、一つの状態に割り当てられるシンボル数が少ない場合、学習が充分に行われず、符号化効率の低下が生じる。同時に、状態毎に留意すべき記憶容量の増大も問題となる。この解決策として、2値シンボルの出現確率が近い状態を選出し、統合する手法が考えられる。JPEGでは、レベルプレーン展開に準じた2値化と、マルコフモデル条件付けを記述した統計モデルを用意



(a) 原画像



(b) 512×512 画素にダウンサンプリングした画像

図 1: 画像 PM001 の DCT 係数に対する KS 統計量

している。JPEGの統計モデルは高効率なマルコフモデル符号化及び状態統合の手法であるが、ハフマン符号との親和性を考慮しているため、その最適性には疑問が残る。

本研究では、まず、超高精細画像におけるDCT係数の分布に関して考察し、その分布に基づいた統計モデルを提案する。なお、マルコフモデル条件付けはJPEGと同じ手法を採用する。

2 DCT 係数の分布

DCT係数のAC係数は一般に正規分布を用いて近似される[7]。しかし、Reininger等の報告によると、256×256画素、256階調の画像に対しKS統計量を測定した結果、特に変化の激しい画像を除いた場合、低周波成分の近似として、ラプラス分布の方が正規分布より適切であるという結論を得ている[8]。超高精細画像PM001の輝度信号成分(4,096×4,096画素、4,096階調)に対するDCT係数のKS統計量を図1に示す。原画像はRGB信号から成るカラー画像であるが、本報告では次式で

算出される輝度信号 Y のみを用いる。ここで、 Y は小数值をとり得るが、四捨五入し、整数化した。

$$Y = 0.30R + 0.59G + 0.11B \quad (1)$$

比較のため、(a) 原画像、(b) 512×512 画素にダウンサンプリングした画像、双方の測定結果を示す。図 1 の点線は正規分布の、実線はラプラス分布の KS 統計量をそれぞれ示している。横軸はジグザグスキャンの順序 k である。DCT のブロックサイズは JPEG に準拠し、 8×8 画素としたため、 $k = 1, 2, \dots, 63$ である。KS 統計量 t は集合 $X = (x_1, x_2, \dots, x_M)$ の分布関数 $F_X(\cdot)$ と、それを近似するための関数 $F(\cdot)$ との近似度を表す一つの指標であり、次式で与えられる [9]。

$$t = \max_{i=1,2,\dots,M} |F_X(x_i) - F(x_i)| \quad (2)$$

従って、 t が小さいほど近似度が高いといえる。また、JPEG では、DC 係数 ($k = 0$) のみ直前ブロックとの差分を符号化の対象としているため、直前ブロックとの差信号を測定の対象とした。なお、危険率 10% の両側検定におけるパーセント点 ϵ は超高精細画像の場合 $\epsilon = 2.39 \times 10^{-3}$ 、 512×512 画素の場合 $\epsilon = 1.91 \times 10^{-2}$ である。図 1 より、(a)、(b) 双方において、DC 係数を除く全ての係数でラプラス分布の方が正規分布より KS 統計量が小さい。また、原画像では高周波になるほど KS 統計量は減少する傾向であるのに対し、 512×512 画素の画像では高周波になるほど増加する結果となった。

3 DCT 係数の統計モデル

JPEG では DC 係数および AC 係数二種類の統計モデルを用意している。著者等は先の研究において、予測符号化に対し、(1) 予測誤差を絶対値と符号を表すレベルプレーン展開によって 2 値化する、(2) 予測誤差分布がラプラス分布で近似される、以上二つの条件に基づき考察した結果、絶対値を表すプレーンにおいて、2 値シンボルの出現確率は常に等しいとの結論を導出し、絶対値を表すプレーンの状態を一つに統合する手法を提案した [10]。また、JPEG の Lossless モードに適用し、新たな統計モデルの設計も行った [2]。図 1 より、ラプラス分布の KS 統計量は危険率 10% の両側検定のパーセント点 ϵ に比べかなり大きな値であり、近似に

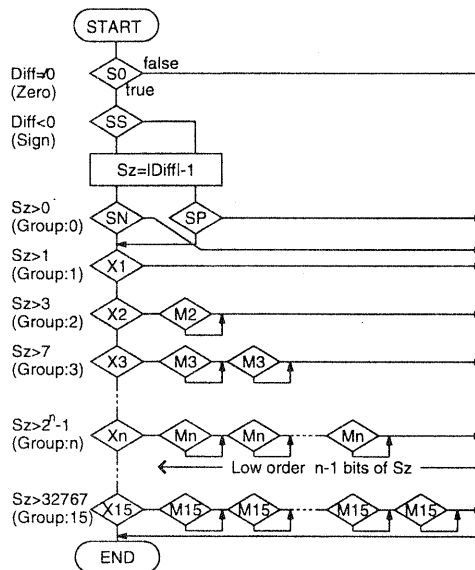


図 2: JPEG における DC 係数の 2 値化

表 1: JPEG の DC 係数に対する統計モデル

判定項目	判定の状態と条件付け
ゼロ判定	5 ステート
サイン判定	5 ステート
グループ 0	5×2 (正/負) ステート
グループ 1	X_1
グループ 2 ~ 15	$X_i (i = 2 \sim 15)$
付加ビット	M_i

用いる最適な関数とは言い難い。しかし、一般に使用されるガウス分布に比べ、ラプラス分布は DC 係数で多少劣るものの、全ての AC 係数において高い近似度を示している。従って、文献 [1] において考察した手法による符号量の低減が期待される。以下、DC 係数、AC 係数それぞれに対し、ラプラス分布に基づいた統計モデルを作成する。

3.1 DC 係数

JPEG における DC 係数の 2 値化のフローチャートを図 2 に示す [11]。まず、直前ブロックの DC 係数との差分 $Diff$ に対し、ゼロ判定 ($Diff \neq 0$?) が行われる。判定結果が false の場合、QM-coder はシンボル 0 を符号化し、true の場合は 1 を符号化する。もし、判定が false であれば、 $Diff = 0$ と

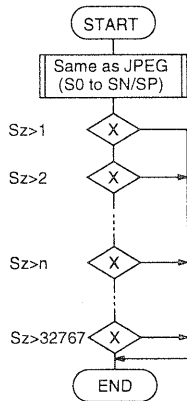


図 3: 提案する DC 係数の 2 値化

表 2: 提案する DC 係数の統計モデル

判定項目	判定の状態と条件付け
ゼロ判定	} JPEG と同様
サイン判定	
グループ 0	
$S_z > 1$ 以降全て	

決定されるため、符号化を終了する。この判定は、コンテキスト S_0 によって条件付けられる。続いて、コンテキスト SS に条件付けられたサイン判定 ($Diff < 0?$) が行われる。更に、 $S_z = |Diff| - 1$ に対し、グループ判定が行われる。グループ 3 以降の判定で false を出力した場合、グループ内のどの値であるかを示す 2 進数 (グループ n であれば $n-1$ ビット) を符号化する。コンテキストによる条件付けの手法を表 1 に示す。ゼロ判定 (S_0)、サイン判定 (SS) およびグループ 0 判定 (SN/SP) は直前ブロックの DC 差分に対し、それぞれ大 (正)、小 (正)、ゼロ近傍、小 (負)、大 (負) の 5 ステートの状態を割り当てる。グループ 1 以降の判定 (X_1, X_2, \dots, X_{15})、および付加ビット (M_2, M_3, \dots, M_{15}) は判定毎に一つの状態を割り当てる。

JPEG の 2 値判定はレベルプレーン展開に準じた手法ではあるが、グループ 2 以降、即ち DC 差分の絶対値が 3 以上の場合はグループ判定と付加ビットで表現している。しかし、グループ化せず、DC 差分の絶対値が 2 (グループ 1 に相当) 以降の判定もレベルプレーン展開で 2 値化すると、ラプラス

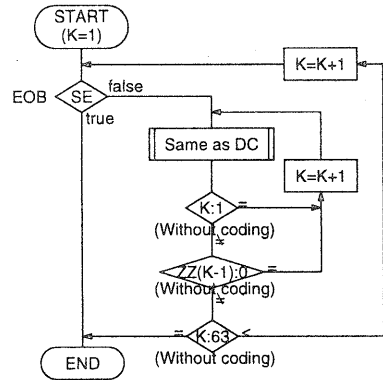


図 4: JPEG における AC 係数の 2 値化

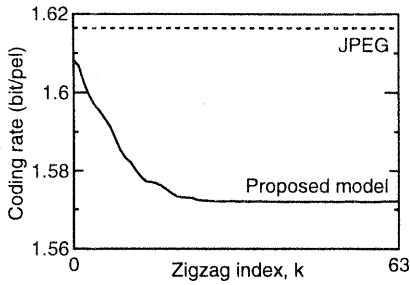
表 3: JPEG の AC 係数に対する統計モデル

判定項目	判定の状態と条件付け
EOB 判定	係数 (k) 毎に 63 状態
ゼロ判定	係数毎に 63 状態
サイン判定	$Q_e = 0x5A1D$, $MPS = 0$ に固定
グループ 0	係数毎に 63 状態
グループ 1	係数毎に 63 状態
グループ 2 ~ 15	} K_x によって さらに 2 状態
付加ビット	

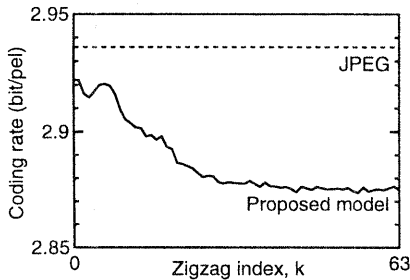
分布で近似される場合、全ての判定における 2 値シンボルの出現確率はほぼ均一となる。従って、図 3 に示すレベルプレーン展開で 2 値化し、 $S_z > 1?$ 以降全ての判定に対して、1 ステート (X) を割り当てる。従って、統計モデルは表 2 となる。

3.2 AC 係数

JPEG では、AC 係数を図 4 に示すフローチャートに従い 2 値化している [11]。ここで、ジグザグインデックス k における AC 係数の値を $ZZ(k)$ ($k = 1, 2, \dots, 63$) とする。AC 係数では、まず、コンテキスト SE で条件付けられた EOB (End of Block) 判定を行う。EOB 判定は $k = 1$ および直前の係数が 0 以外の時のみ行われる。当然、 $k = 63$ の場合は不要である。続いて、DC 係数と同様の手法により、2 値化が行われる。各 2 値判定は表 3 に示す手法により条件付けられる。EOB 判定からグループ 1 の判定に対し、係数毎にそれぞれ 63 状態を割



(a) 原画像



(b) 512×512 画素にダウンサンプリングした画像

図 5: JPEG と提案モデルにおける符号量の比較

り当てているが、サイン判定では 0, 1 の出現確率が等しいと考え、確率がほぼ 0.5 であることを意味する $Q_e = 0x5A1D$ ($0x: 16$ 進数) に固定している。グループ 2 以降の判定に関してはグループ判定、付加ビットそれぞれに対し、低周波と高周波を分離するパラメータ K_x により、さらに 2 状態を割り当てている。

図 3 より、AC 係数の 2 値化は DC 係数の場合と EOB の処理が異なるのみである。従って、DC 係数と同様、 $S_2 > 1?$ 以降全ての判定にレベルプレーン展開を採用する。JPEG では異なる成分を低周波、高周波の 2 種類に統合しているが、提案手法では、成分毎に独立していたグループ 0, グループ 1 およびそれに続く全ての判定を一つに統合し、成分間の統合は行わない。ここで、 $ZZ(k) = 0$ の判定は量子化器の特性上、他の判定とは 0, 1 の出現確率が異なるため [10]、統合は行わない。

表 4: 提案する AC 係数の統計モデル

判定項目	判定の状態と条件付け
EOB 判定	} JPEG と同様
ゼロ判定	
サイン判定	
$S_2 > 1$ 以降全て	

表 5: 状態数の比較

	DC	AC	計
JPEG	49	280	329
提案モデル	21	189	210

4 シミュレーション

超高精細画像およびダウンサンプリングした画像に対し、符号化シミュレーションを行う。提案する統計モデルの効果を調べるため、係数毎に提案モデルを採用した場合の符号量を図 5 に示す。横軸 k は、 k 以下の係数全てに提案モデルを使用したことを意味する。図中の点線は JPEG の符号量、実線は提案モデルによる符号量である。JPEG におけるパラメータ K_x はデフォルトの 5 を採用した。符号量にヘッダ情報は含まれていない。量子化表は JPEG Annex K に記載してある表を使用した。この表は 256 階調の画像に対する量子化表であり、4,096 階調の画像では量子化ステップが小さいと思われる。しかし、JPEG と提案した統計モデルは EOB 判定が同じであるため、荒い量子化表では、早期の判定で符号化が終了し、差異が見られない。また、超高精細画像の性質を考慮した場合、高品質な再生画像を必要とすると考え、本量子化表を採用した。なお、再生画像の SNR は原画像で 52.47(dB)、512 × 512 の画像で 49.75(dB) である。図 5 より、双方の画像で提案モデルにより、符号量が低減していることが分かる。提案した統計モデルはラプラス分布に基づいた手法であるため、KS 統計量が大きな場合、例えば原画像の低周波成分、ダウンサンプリングした画像の高周波成分、等における号量の増大を危惧したが、全ての係数を提案モデルで符号化した場合 ($k = 63$)、符号量が最小、或いは極めて最小値に近い値となることがわかった。従って、表 2 に示した DC 係数の統計モデルを全 AC 係数に適用する。従って、AC 係数の

表 6: SNR と平均符号量 (bit/pel) の比較

		画像名		
		PA022	PM001	SA001
原画像	SNR(dB)	55.57	52.47	54.37
	JPEG	1.201	1.616	1.518
	提案モデル	1.163	1.572	1.463
512 × 512	SNR(dB)	49.78	49.75	48.06
	JPEG	2.450	2.936	3.296
	提案モデル	2.380	2.875	3.202

統計モデルは表 4 となる。また、JPEG および提案した統計モデルの状態数をそれぞれ表 5 に示す。表 5 より、提案モデルの状態数は JPEG の 2/3 以下である。

JPEG および提案モデルを用い、三種類のサンプル画像に対して符号化した結果を表 6 に示す。超高精細画像、ダウンサンプリングした画像全てにおいて、提案モデルにより、2.08 ~ 3.62% 符号量が低減される。また、量子化表を変化させた場合も同様の結果となった。

5 むすび

本論文では、超高精細画像における DCT 係数の分布と QM-coder を用いた符号化に関して検討した。DCT 係数の KS 統計量を測定し、正規分布よりもラプラス分布の方が適当であることを示した。その際、KS 統計量は、超高精細画像では低周波、512 × 512 画素にダウンサンプリングした画像では高周波で増加する傾向がみられた。また、DC 係数、AC 係数双方に対し、ラプラス分布に基づいた統計モデルを提案した。JPEG の統計モデルは 329 の状態を必要とするのに対し、提案モデルでは 210 へと低減される。また、符号化シミュレーションの結果、符号量も 2.08 ~ 3.62% 減少することがわかった。今後はプログレッシブモードへの適用等について検討する予定である。

謝辞

本研究の一部は文部省科学研究費補助金奨励 A (課題番号 11750345) の補助を受けた。

参考文献

- [1] 黒木祥光, 太田諦二, “JPEG Spatial 方式における新しい統計モデル”, 信学論 (B-I), vol. J79-B-I, no. 7, pp. 509-511, July 1996.
- [2] 黒木祥光, 太田諦二, 上繁義史 “JPEG 可逆モードにおける統計モデルの再構成”, 信学論 (D-II), vol. J80-D-II, no. 11, pp. 3108-3110, Nov. 1997.
- [3] 黒木祥光, 芝浩二郎, 上繁義史, 太田諦二, “QM-Coder を用いた超高精細画像の可逆符号化に関する一検討”, 信学技法, IE98-31, July. 1998.
- [4] 黒木祥光, 太田諦二, “超高精細画像の高画質符号化”, 信学論 (B-I), vol. J78-B-I, no. 11, pp. 672-679, Nov. 1995.
- [5] ISO/IEC 10928-1 “Digital compression and coding of continuous-tone still images: Requirements and guidelines”, Feb. 1994.
- [6] ISO/IEC 11544, “Coded representation of picture and audio information -Progressive bi-level image compression”, Dec. 1993.
- [7] A. N. Netravali and J. O. Limb, “Picture coding: A review,” *Proc. IEEE*, vol. 68, pp. 366-406, Mar. 1982.
- [8] R. C. Reininger and J. D. Gibson, “Distributions of the Two-Dimensional DCT Coefficients for Images”, *IEEE Trans. Commun.*, vol. COM-31, pp.835-839, June 1983.
- [9] 竹内 啓, 確率分布と統計解析, 日本規格協会, 1975.
- [10] 黒木祥光, 太田諦二: “QM-Coder を用いた多値画像可逆符号化における状態の統合”, PCSJ95, 9.21.
- [11] W. B. Pennebaker and J. L. Mitchell, “JPEG still image data compression standard”, *Van Nostrand Reinhold*, 1993.