

ラプラス分布に基づく JPEG 統計モデルの状態割り当て

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

〒 899-5193 鹿児島県始良郡隼人町真孝 1460-1

Tel(Fax):0995-42-9022

E-Mail:kuroki@kagoshima-ct.ac.jp

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。本研究では JPEG の統計モデルにより条件付けられた各コンテキストの KS(Kolmogorov-Smirnov) 統計量を測定し、AC 係数の分布は正規分布よりもラプラス分布に近いことを示す。また、ラプラス分布に基づいた統計モデルの状態割り当てについて考察し、符号量の低減を試みる。

キーワード 統計モデル, QM-coder, JPEG, DCT

State assignment of the JPEG statistical model based on a Laplacian distribution

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

[†]Kagoshima National Colledge of Technology

[‡]Faculty of Engineering, Kyushu Institute of Technology

1460-1 Shinko, Hayato-cho, Aira-gun; Kagoshima 899-5193, Japan

Tel(Fax):0995-42-9022

E-Mail:kuroki@kagoshima-ct.ac.jp

Abstract

In the lossy mode of Joint Photographic Experts Group (JPEG), Discrete Cosign Transform (DCT) coefficients are entropy-coded either Huffman coder or binary arithmetic coder called QM-coder. In the case where QM-coder is employed, DCT coefficients are converted into binary sequences and then context-modeled referring the statistical model. This paper presents that the distributions of AC coefficients are not Gaussian, but Laplacian through Kolmogorov-Smirnov (KS) test statistic. Furthermore we propose a state assignment scheme based on a Laplacian distribution to achieve higher compression performance than JPEG.

key words Statistical model, QM-coder, JPEG, DCT

1 まえがき

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 [1]. JPEG では、量子化された DCT 係数をエントロピー符号化する場合、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 では必要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、コンテキスト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)[2] 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場合、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、状態毎に用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している.

著者等は [9], [10] において、DCT 係数の分布がラプラス分布で近似可能であることを示した. また、ラプラス分布に基づいた新たな統計モデルを作成し、符号量と状態数の低減を行った. しかし、提案したモデルは DCT 係数を単純なレベルプレーン展開で

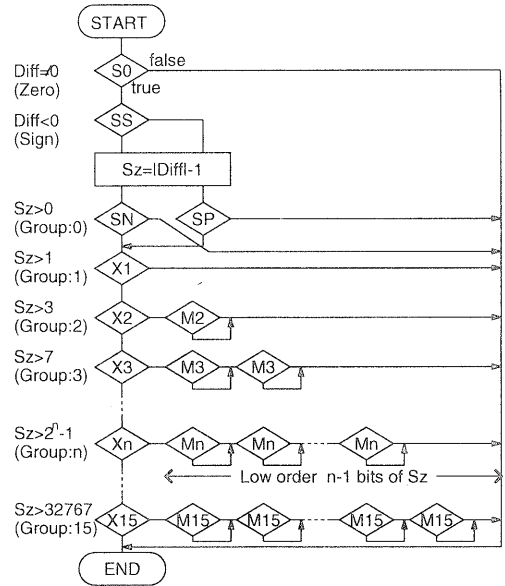


図 1: JPEG における DC 係数の 2 値化

2 値化するため、係数が大きな場合、あるいは量子化ステップが細かい場合、2 値系列が長くなり、符号化に多大な時間を必要とする. 一方、著者等は先の研究で、JPEG 可逆モードで使用される統計モデルに関して考察した結果、2 値化の手法およびマルコフモデル条件付けは JPEG と同じであるが、付加ビットに対する条件付けを変更することにより、符号量の低減を行っている [3]. 本研究では [3] の手法を DCT 係数の符号化に適応し、その有効性を検証する.

2 JPEG 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. それぞれの統計モデルの概略を以下に述べる.

2.1 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 1 に示す [4]. 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

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

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、 $S_2 = |Diff| - 1$ 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 $faulse$ 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。

2.2 AC 係数

JPEG では、AC 係数を図 2 に示すフローチャートに従い 2 値化している [4]。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 2 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 状態を割り当てるが、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 $Qe = 0x5A1D$ ($0x:16$ 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 Kx により、さらに 2 状態を割り当てている。

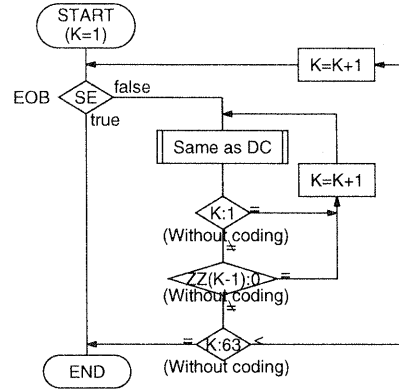


図 2: JPEG における AC 係数の 2 値化

表 2: 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	$Qe = 0x5A1D, MPS = 0$ 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	$X_i (i = 2 \sim 15)$ } Kx によって
付加ビット	M_i } さらに 2 状態

3 各コンテキストにおける 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 [5]。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 [6]。また、ITE 標準画像 (512×480 画素および 512×472 画素、256 階調) [9]、超高精細画像 ($4,096 \times 4,096$ 画素、4,096 階調) [10] における DCT 係数の分布も同様にラプラス分布の方が近似度が高い。従って、DCT 係数の分布がラプラス分布で近似可能と仮定し、統計モデルを作成した [9], [10]。提案したモデルは DCT 係数をレベルプレーン展開で 2 値化した際、各プレーンにおける 2 値シンボルの出現確率が均一になることに着目したものである。しかし、提案したモデルでは DCT 係数を単純なレベルプレーン

表 3: 正規分布とラプラス分布の KS 統計量
(a) DC 係数

		正規分布	ラプラス分布
原画像	PA022	0.147	0.205
	PM001	0.095	0.137
512 × 512	PA022	0.110	0.073
	PM002	0.079	0.132
肌色チャート†		0.160	0.100
ヘアーバンドの女性‡		0.072	0.132
天気予報†		0.084	0.122
スイスの山村‡		0.087	0.117
チューリップ‡		0.048	0.091

(b) AC 係数 ($k \leq Kx$)

		正規分布	ラプラス分布
原画像	PA022	0.437	0.415
	PM001	0.324	0.291
512 × 512	PA022	0.403	0.370
	PM002	0.253	0.216
肌色チャート†		0.295	0.255
ヘアーバンドの女性‡		0.289	0.255
天気予報†		0.321	0.298
スイスの山村‡		0.096	0.050
チューリップ‡		0.210	0.149

(c) AC 係数 ($k > Kx$)

		正規分布	ラプラス分布
原画像	PA022	0.474	0.462
	PM001	0.339	0.300
512 × 512	PA022	0.444	0.421
	PM002	0.267	0.240
肌色チャート†		0.348	0.305
ヘアーバンドの女性‡		0.326	0.291
天気予報†		0.357	0.337
スイスの山村‡		0.102	0.052
チューリップ‡		0.265	0.207

ン展開で 2 値化するため、係数が大きな場合、あるいは量子化ステップが細かい場合、2 値系列が長くなり、符号化に多大な時間を必要とする。一方、著者等は先の研究で、JPEG 可逆モードで使用される統計モデルに関して考察した結果、2 値化の手法およびマルコフモデル条件付けは JPEG と同じであるが、付加ビットに対する条件付けを変更することにより、符号量の低減を行っている [3]。[3] の手法も予測誤差がラプラス分布で近似されると仮定している。

そこで、JPEG 統計モデルの各コンテキストの分布に関して検討する。なお、EOB、ゼロ、符号の各判定はそれぞれ独立に状態割り当てがなされているため、ここでは議論しない。上記判定以降、即ち S_z

の分布に対し、正規分布とラプラス分布を候補とし、それぞれの KS 統計量を測定する。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 [7]。

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

従って、 t が小さいほど近似度が高いといえる。JPEG では、DC 係数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する。AC 係数において低周波、高周波を分離するパラメータ Kx はデフォルトの 5 を用いる。各画像における KS 統計量を表 3 に示す。画像「PA022」と「PM001」は超高精細画像であり、比較のため 512 × 512 画素にダウンサンプリングした画像に対しても測定を行った。画像「肌色チャート」以下は ITE 標準画像であり、† が 512 × 480 画素、‡ が 512 × 472 画素である。表 3 より、DC 係数は正規分布、AC 係数はラプラス分布の方が近似に適していることがわかる。また KS 統計量の値は、DC 係数 < 低周波の AC 係数 < 高周波の AC 係数である。また、エッジが多く高周波成分までエネルギーが分布している画像ほど KS 統計量（特に AC 係数）が小さい。エッジが多い画像では、高周波成分のエネルギーが高く、係数間の分布に変化が少ないためと考えられる。

4 提案する統計モデル

表 3 より、AC 係数の KS 統計量は低周波と高周波の二種類に統合することにより、増加する傾向にある。しかし、この場合も正規分布と比較し、ラプラス分布の方が近似度が高い。また、[3] の統計モデルは、超高精細画像の予測誤差等ラプラス分布と多少異なる分布に対しても、符号化の改善がみられる。従って、DCT 係数の符号化にも [3] のモデルを採用し、符号化効率の改善を試みる。提案する DC 係数の状態割り当てを図 3 に示す。

ラプラス分布で近似される情報源に対し、グループ幅が 2 のべき乗で増加している場合、グループ判定の統合は不可能である。しかし、付加ビットはグループ番号に関わらず、LSB (Least Significant Bit) から何ビット目であるかによって 2 値シンボルの出現確率が変化する [3]。従って、図 1 に示した JPEG の統計モデルでは、グループ毎の付加ビットに状態を割り当てていたが、提案するモデルでは

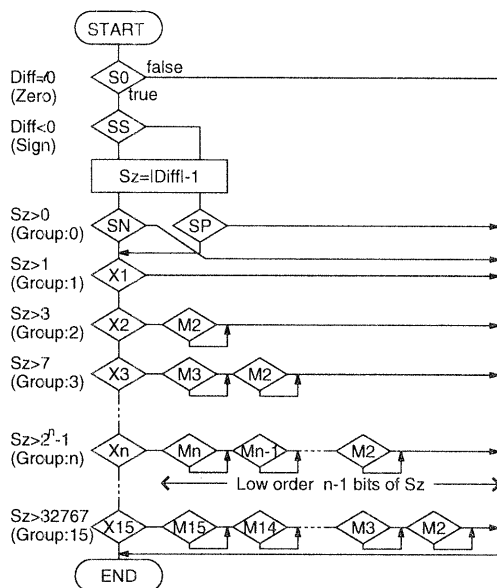


図 3: 提案する DC 係数の状態割り当て

表 4: 符号量の比較 (bit)

		JPEG	提案モデル
原画像	PA022	3166976	3166480
	PM001	4795600	4790288
512 × 512	PA022	127280	127408
	PM002	201696	201536
肌色チャート		88112	88024
ヘアーバンドの女性		117280	117016
天気予報		120480	120176
スイスの山村		329576	329440
チューリップ		295648	295456

付加ビットの下位から何ビット目であるかによって状態を割り当てる。また、AC 係数に対しても同様の割り当てを行う。

5 シミュレーション

JPEG の統計モデルおよび提案モデルを用いた場合の符号量を表 4 に示す。符号量にヘッダ情報は含まれていない。量子化表は 256 階調の画像に対しては JPEG Annex K に記載してある表を使用し、4,096 階調の画像に対しては JPEG Annex K の表の全係数を 16 倍した。表 4 より、唯一 512 × 512 画素の画像「PA022」のみで増加しているものの、

その他全て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わかる。また、量子化表を変化させた場合も同様の結果となった。

6 むすび

本論文では、JPEG の統計モデルにより条件付けた場合の DCT 係数の分布と QM-coder を用いた符号化に関して検討した。各コンテキストにおける KS 統計量を測定し、AC 係数の分布は正規分布よりもラプラス分布に近いことを示した。また、係数毎に求めた KS 統計量に比べ、低周波と高周波の二種類に統合した場合、特にエッジの少ないならかな画像で増加する傾向がみられた。更に、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。提案したモデルはコンテキストの条件付け、2 値化の手法は JPEG と同様であり、付加ビットに対する状態割り当てのみを変化させたものである。符号化シミュレーションの結果、ほぼ全ての画像で符号量を低減することができた。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] ISO/IEC 10928-1 "Digital compression and coding of continuous-tone still images : Requirements and guidelines", Feb. 1994.
- [2] ISO/IEC 11544, "Coded representation of picture and audio information -Progressive bi-level image compression", Dec. 1993.
- [3] 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.
- [4] W. B. Pennebaker and J. L. Mitchell, "JPEG still image data compression standard", Van Nostrand Reinhold, 1993.
- [5] A. N. Netravali and J. O. Limb, "Picture coding: A review," *Proc. IEEE*, vol. 68, pp. 366-406, Mar. 1982.

- [6] 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.
- [7] 竹内 啓, 確率分布と統計解析, 日本規格協会, 1975.
- [8] 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法", PCSJ95, pp.209-210.
- [9] 黒木祥光, 太田諦二, 上繁義史: "QM-coder を用いた DCT 係数の符号化", PCSJ99, pp. 23-24.
- [10] 黒木祥光, 太田諦二, 上繁義史: "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", 情報処理学会研究報告, vol. 99, no. 82, pp. 7-12.