

ブロック適応イントラ予測に基づいた JPEG 画像のロスレス再符号化

野本 恭男[†] 萩野 崇[†] 松田 一郎[†] 伊東 晋[†]

[†] 東京理科大学 理工学部

〒278-8510 千葉県 野田市 山崎 2641

E-mail:{nomoto,hagino,matsuda,itoh}@itohws01.ee.noda.tus.ac.jp

本稿では、既存の JPEG ファイルを一切の画質劣化を生じさせることなく、更に圧縮可能な再符号化方式について述べる。この方式では、JPEG ファイル内に格納されている量子化後の DCT 係数のブロック間相関を利用するため、H.264 と同様なブロック適応イントラ予測を導入している。このときの予測値は、8×8 画素のブロック毎に画素領域で算出されるが、対応する予測誤差を DCT 領域で求めることによって、JPEG ファイル内の DCT 係数が完全再構成されることを保証している。予測誤差の適応算術符号化に必要な確率モデルは画像毎に繰り返し最適化され、付加情報として圧縮データに格納される。シミュレーションの結果、提案方式は JPEG 画像の符号量を 18-32 % 削減可能であることが確認された。

Lossless Re-encoding of JPEG Images Based on Block-Adaptive Intra Prediction

Yukio NOMOTO[†] Takashi HAGINO[†] Ichiro MATSUDA[†] Susumu ITOH[†]

[†] Faculty of Science and Technology, Tokyo University of Science

2641 Yamazaki, Noda-shi, Chiba 278-8510, Japan

E-mail:{nomoto,hagino,matsuda,itoh}@itohws01.ee.noda.tus.ac.jp

This paper describes a kind of transcoding scheme which compresses existing JPEG files without any loss of quality. In this scheme, an H.264-like block-adaptive intra prediction technique is employed to exploit inter-block correlations of quantized DCT coefficients stored in the JPEG file. This prediction is performed in a spatial domain of each block composed of 8×8 pels, but the corresponding prediction residuals are calculated in a DCT domain to ensure lossless reconstruction of the original coefficients. Moreover, probability models of DCT coefficients used for arithmetic coding of the prediction residuals are iteratively optimized for each image and then encoded as side-information. Simulation results indicate that the proposed scheme can reduce coding rates of JPEG images by 18-32 %.

1 はじめに

静止画像符号化の国際標準方式である JPEG¹⁾ は、その登場以来十数年に渡って自然画像用の記録方式として広く利用されてきた。近年、JPEG2000²⁾ や HD-Photo³⁾ のように、より性能の優れた符号化方式も開発されているが、デジタルカメラ等の画像機器には現在でもなお JPEG 方式がほぼ独占的に採用され続けている。画像符号化技術の世代交代を妨げる要因の一つとして、これまで JPEG 方式で蓄積されてきた膨大な量の画像データの存在が挙げられる。例えば、一旦 JPEG

で符号化された画像を JPEG2000 等で再符号化することは可能であるが、非可逆符号化方式の場合は更なる画質低下が避けられないため、JPEG2000 本来の符号化性能を発揮することが極めて困難となる⁴⁾。

このような観点から、既存の JPEG ファイルの画質を保ったまま符号量だけを削減するロスレス再符号化方式が近年開発されている^{5) 6) 7)}。これらの方式は、いずれも JPEG ファイルに格納されている DCT 係数を、算術符号等の高効率なエントロピー符号化技術を用いて再符号化しようとするものである。この考え方自体は新しいものでは

なく、JPEG 標準にも QM-coder と呼ばれる 2 値算術符号器を用いたエントロピー符号化モードが規定されている¹⁾。しかし、JPEG baseline 方式のハフマン符号を算術符号に置き換えただけでは符号化効率の改善量に限界があるため、最近の方式では、DCT 係数のブロック間相関に着目しているのが特徴である^{5) 6) 7)}。但し、DCT 領域の係数のまま上記の相関を利用しようとすると、隣接ブロックの同一周波数の DCT 係数しか参照することができず、必ずしも十分な効果が得られているとはいえない。

そこで本稿では、ブロック適応イントラ予測に基づいたロスレス再符号化方式を提案する。この方式では、ブロック毎に H.264 のイントラ予測モード⁸⁾と同様の手法によって空間領域で予測値の算出を行う。これにより、DCT 領域の複数の周波数成分に分散したブロック間相関を利用することが可能となる。但し、元の JPEG ファイルに格納されている DCT 係数の完全再構成を可能にするため、予測誤差の算出は DCT 領域で行われる。これにより、JPEG の復号器と同一の逆量子化および逆 DCT の操作を用いる限り、再生画像上でも完全再構成が保証されることとなる。

2 ブロック適応イントラ予測

図 1 に、提案する再符号化方式の構成を示す。まず、入力された JPEG ファイルに対してハフマン復号の処理を施し、量子化済みの DCT 係数 $y_{k,l}(i,j)$ を抽出する。ここで、 k, l は 8×8 画素で構成されるブロックの位置を表し、 $i, j = 0, 1, \dots, 7$ は各ブロックの 2 次元 DCT 係数の周波数を示している。次に、JPEG ファイルのヘッダ情報に含まれる量子化マトリックス $Q(i,j)$ を用いてこれらの係数を逆量子化し、逆 DCT を適用して各ブロックの再生値 $x_{k,l}(i,j)$ を算出する。更に、図 2 に○印で示

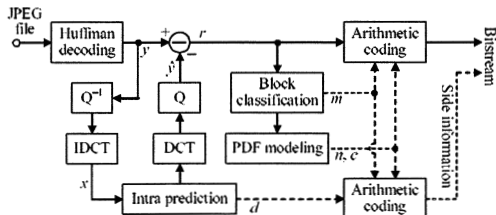


図 1 再符号器の構成

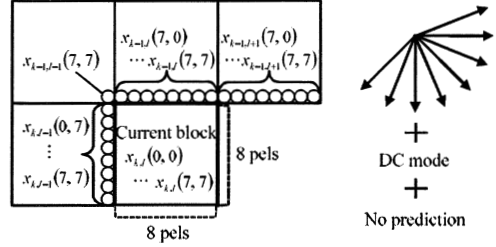


図 2 ブロック適応イントラ予測

した隣接ブロックの再生値を使用し、当該ブロックの内部の画素に対して空間方向のイントラ予測を適応的に実行する。このブロック適応イントラ予測は、H.264 で規定されている 4×4 イントラ予測を JPEG の DCT サイズである 8×8 画素のサイズに拡張したものであり、「予測なし」を含めた計 10 種類の予測モードをブロック単位で適応選択している。その後、得られた予測値を DCT 領域に変換し、量子化マトリックス $Q(i,j)$ を用いて量子化することで DCT 係数 $\hat{y}_{k,l}(i,j)$ を得る。逆量子化 (Q^{-1}) および量子化 (Q) の際に同じ量子化マトリックス $Q(i,j)$ が用いられているため、DCT 係数 $y_{k,l}(i,j)$ と $\hat{y}_{k,l}(i,j)$ は各周波数で同一スケールの整数値となる。従って、予測誤差の DCT 係数は両整数値の差分として次式により算出される。

$$r_{k,l}(i,j) = y_{k,l}(i,j) - \hat{y}_{k,l}(i,j) \quad (1)$$

上記のブロック適応イントラ予測によってブロック間の相関が取り除かれていれば、元の DCT 係数 $y_{k,l}(i,j)$ を直接符号化するより、予測誤差 $r_{k,l}(i,j)$ を符号化した方が効率的となる。但し、各ブロックで用いられる予測モード $d(k,l) \in \{0, 1, \dots, 9\}$ は、付加情報として符号化データに含める必要がある。

3 予測誤差の適応算術符号化

上述の手順により DCT 領域で算出された予測誤差の確率密度関数を高精度にモデル化するため、本方式では次式に示す 32 種類の一般化ガウス関数を用意している⁹⁾。

$$P_n(t) = \frac{c_n \cdot \eta(c_n, \sigma_n)}{2\Gamma(1/c_n)} \cdot \exp\left\{-|\eta(c_n, \sigma_n) \cdot t|^{c_n}\right\},$$

$$\eta(c_n, \sigma_n) = \frac{1}{\sigma_n} \sqrt{\frac{\Gamma(3/c_n)}{\Gamma(1/c_n)}} \quad (n = 1, 2, \dots, 32) \quad (2)$$

但し、 $\Gamma(\cdot)$ はガンマ関数、 σ_n は予測誤差の標準偏差、 c_n は関数 $P_n(t)$ の尖度を制御する形状パラメー

タである。本稿では、32通りの σ_n の値を予測誤差のアクティビティの範囲を網羅するように予め定めている。一方、形状パラメータ c_n の値は実際の確率密度関数の分布形に合わせて画像毎に変化させることが可能である。更に、符号化対象画像は8×8画素のブロック単位で M 通りのクラスに分類されており、それぞれのクラスには周波数成分 (i, j) と確率モデルの分散値 σ_n^2 との対応関係を記述した表(分散マップ)が割り当てられている。ここで、 $P_n(t)$ は、量子化前のDCT係数の確率密度関数をモデル化したものであるから、符号化対象となる予測誤差の量子化レベル r の生起確率は次式によって表せる。

$$\Pr(r|n, Q) = F_n((r+0.5) \cdot Q) - F_n((r-0.5) \cdot Q),$$

$$F_n(x) = \int_{-\infty}^x P_n(t) dt \quad (3)$$

但し、 $Q = Q(i, j)$ は量子化ステップサイズ、 $n = n_m(i, j)$ は m 番目クラスの分散マップの値をそれぞれ表している。実際の予測誤差 $r = r_{k,l}(i, j)$ の値は、(3)式によって算出される生起確率に従って算術符号化される。本稿では高速な多値算術符号器として知られるレンジコーダ¹⁰⁾を使用する。

4 パラメータの最適化手順

提案方式では、以下に記載するパラメータを付加情報として符号化する必要がある。

- 元のJPEGファイルから抽出された量子化マトリックス $Q(i, j)$
- 各ブロックで適用された予測モード
 $d(k, l) \in \{0, 1, \dots, 9\}$
- ブロック毎に割り当てられたクラス番号
 $m(k, l) \in \{1, 2, \dots, M\}$
- m 番目のクラスに割り当てられた分散マップ
 $n_m(i, j) \in \{1, 2, \dots, 32\}$
- 確率モデル $P_n(t)$ の形状パラメータ
 $c_n \in \{0.2, 0.4, \dots, 3.2\}$

実際の符号化に先立ち、量子化マトリックス $Q(i, j)$ を除く各パラメータは、次式のコスト関数が最小となるよう画像毎に最適化される。

$$J = - \sum_{k,l} \sum_{i,j} \log_2 \Pr(r_{k,l}(i,j) | n_{m(k,l)}(i,j), Q(i,j))$$

$$+ B_{side} \quad (4)$$

但し、右辺第1項は予測誤差の符号量、第2項(B_{side})は上記の付加情報に要する符号量である。具体的な最適化手順を以下に示す。

- (1) 各ブロックの初期予測モード $d(k, l)$ を決定する。この段階では、コスト関数 J の代わりに予測誤差の2乗和 $S_{k,l}(d)$ を最小化している。

$$d(k, l) = \underset{d \in \{0, 1, \dots, 9\}}{\operatorname{argmin}} S_{k,l}(d)$$

$$S_{k,l}(d) = \sum_{i,j=0}^7 \left\{ r_{k,l}(i, j) |_{mode=d} \right\}^2 \quad (5)$$

また、上式により得られた $S_{k,l}(d(k, l))$ の値を $M = 24$ レベルに量子化することで、ブロック毎にクラス番号の初期値 $m(k, l) \in \{1, 2, \dots, M\}$ を割り当てる。

- (2) 各クラス(m)に属するDCT係数について、周波数成分毎に32種類の確率モデル $P_n(t)$ ($n = 1, 2, \dots, 32$)を試し、分散マップ $n_m(i, j)$ の値を最適化する。
- (3) 確率モデル毎に形状パラメータ c_n の値を $\{0.2, 0.4, \dots, 3.2\}$ の範囲で変化させ、最適な値を選択する。
- (4) 各ブロックについて M 通りの分散マップをそれぞれ試し、最適なクラスを再選択する。
- (5) 各ブロックについて、コスト関数 J が最小となる予測モード $d(k, l)$ を決定する。
- (6) M 種類あるクラスの内1つを削除すると仮定し、コスト関数 J の値を評価する。この時、いくつかのブロックでは2番目に最適なクラス(分散マップ)が代用されるため、(4)式の第1項は増加するが、削除したクラスに関する付加情報は第2項から除外される。このトレードオフ関係に基づき、 J の減少量が最も大きなクラスを実際に削除する。
- (7) 上記のクラス削除処理を、コスト関数 J が減少しなくなるまで繰り返す。これにより、各JPEG画像の性質に応じて自動的にクラス数 M の値を設定することが可能となる。
- (8) (2)～(7)の手順を、全てのパラメータが収束するまで反復的に実行する。通常、パラメータが収束するまでに要する反復回数は25～35回程度である。

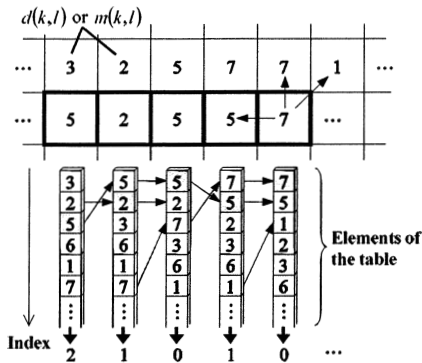


図3 Move-To-Front 法による付加情報の符号化

最適化後の各パラメータを符号化する際は、それぞれの統計的性質に応じた工夫により、付加情報の削減を図っている。例えば、ブロック毎に必要なとなる $d(k,l)$ と $m(k,l)$ については、その値を直接符号化するのではなく、MTF (Move-To-Front)¹¹⁾ と呼ばれる手法により並び替えられたテーブルの位置情報 (インデックス値) を符号化対象としている。MTF は図3のように、上、左、右上のブロックで選択された値がテーブルの上位に現れるよう逐次並び替えを行っており、パラメータ $d(k,l)$ および $m(k,l)$ の値がブロック間で相関を有する場合、得られるインデックス値は比較的小さな値となることから、効率の良い算術符号化が期待できる。

5 特性評価

モノクローム (Y 信号)、カラー各 9 種類の JPEG 画像を対象として、符号化実験を実施した。表1の右端に示した「JPEG (baseline)」は、Independent JPEG Group (IJG) の符号化ツール cjpeg を使用し、画像品質を制御する Q ファクタを 75 に設定して得られた JPEG 画像の符号化レートである。それ以外の方式については上記の JPEG 画像を再符号化することで得られた符号化レートを表している。ここで、「StuffIt」¹²⁾ は Smith Micro (旧 Allume Systems) 社の商用圧縮ツールの名称であり、評価に使用した version 9.0 では、JPEG ファイルのロスレス再符号化機能を有していることが知られている。但し、「StuffIt」が使用している JPEG ファイル用の圧縮アルゴリズムの詳細は非公開である。「PackJPG」¹³⁾ はフリーで入手可能なプログラム

であり、文献7) で提案されたアルゴリズムを実装したものである。「JPEG (QM-coder)」は DCT 係数のエントロピー符号器に JPEG 標準で規定された QM-coder¹⁾ を用いた場合に相当する。なお、提案方式においてカラー画像を再符号化する場合は、Y, C_b , C_r 信号に対して共通の分散マップ (M 種類) を割り当てるものとする。また、表中で太字で示した数値は、各画像について最も優れた符号化レートであることを示している。これより、多くの画像において提案方式は最も優れた符号化性能を達成しており、「JPEG (baseline)」に比べてモノクローム・カラー共に平均で 22 % 符号量を削減することに成功している。しかし、特にサイズの大きなカラー画像において提案方式が StuffIt に劣る場合も見受けられる。提案方式では、Y, C_b , C_r 信号間の相関を全く利用しておらず、この点で改善の余地があるものと考えられる。

最後に、画像「Camera」と「Lena」について求めた符号化レートと SN 比の関係を図4に示す。同図には、上記のロスレス再符号化方式と共に代表的な直接符号化方式として JPEG 2000²⁾ および HD-Photo³⁾ の特性を併せて示した。これより、提案方式と StuffIt の符号化性能はほぼ同等であり、他のロスレス再符号化方式に対して明確な利得を有していることがわかる。また、画像「Lena」に関しては、提案方式の符号化性能が直接符号化方式である HD-Photo の性能に迫っている様子を確認できる。HD-Photo はマイクロソフト社によって開発され、次世代標準 JPEG-XR への採用が検討されている最新の方式であるが、ロスレス再符号化の技術を用いることで旧規格との互換性を保持したまま最新の符号化方式に対抗できる性能を実現する可能性を秘めているといえる。

Table 1 符号化レートの比較 (bits/pel)

(a) モノクローム画像

画像	サイズ	提案方式	StuffIt	PackJPG	JPEG (QM-coder)	JPEG (baseline)
Camera	256×256	1.014	1.063	1.111	1.217	1.319
Couple		0.812	0.827	0.890	0.987	1.047
Bridge		1.706	1.729	1.760	1.851	2.085
Airplane	512×512	0.797	0.784	0.850	0.955	1.034
Lena		0.860	0.862	0.931	1.062	1.131
Mandrill		1.706	1.744	1.771	1.898	2.103
Balloon	720×576	0.394	0.381	0.442	0.529	0.577
Barb		0.996	1.003	1.097	1.235	1.321
Barb2		1.050	1.050	1.112	1.251	1.366
Average		1.037	1.049	1.107	1.221	1.331

(b) カラー画像

画像	サイズ	提案方式	StuffIt	PackJPG	JPEG (QM-coder)	JPEG (baseline)
Girl	256×256	0.935	0.949	1.007	1.116	1.214
Couple		0.920	0.940	1.010	1.116	1.212
Swiss		2.108	2.158	2.214	2.336	2.611
Airplane	512×512	0.896	0.881	0.959	1.078	1.183
Lena		0.858	0.855	0.939	1.072	1.153
Mandrill		1.899	1.928	1.974	2.108	2.357
KodakNo.05	512×768	1.610	1.607	1.746	1.884	2.050
KodakNo.15		0.781	0.781	0.847	0.991	1.060
KodakNo.23		0.617	0.588	0.670	0.774	0.834
Average		1.181	1.187	1.263	1.386	1.519

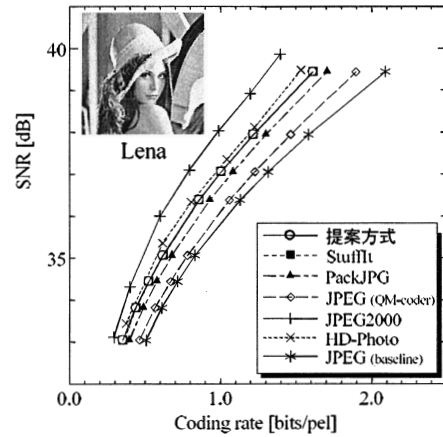
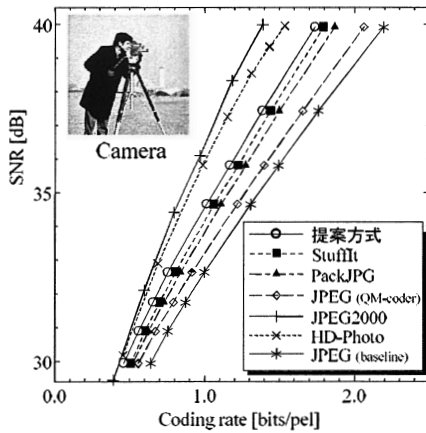


図 4 符号化特性

6 まとめ

本稿では、一旦 JPEG 方式によって符号化された画像を、画質を一切劣化させずに再符号化する方式を提案した。この方式では、H.264 と同様なブロック適応イントラ予測の技術を導入することにより、JPEG 標準では直流係数を除いて利用不可であった、ブロック間相関に基づいた冗長度削減を可能にしている。また、DCT 領域で算出した予測誤差の確率モデルを、クラス分類に基づいてブロック毎に切り替える手法を開発し、算術符号化の効率を高めている。シミュレーションの結果、提案方式の符号化効率は第三者のベンチマークテスト¹⁴⁾ ¹⁵⁾において最も優れた性能を示している StuffIt に匹敵していることを確認した。今後はイントラ予測方式の改良や色信号間予測の導入等について検討を加え、更なる性能向上を図る予定である。

参考文献

- 1) ITU-T Rec. T.81 | ISO/IEC 10918-1, “Information Technology – Digital Compression and Coding of Continuous-Tone Still Images: Requirements and Guidelines,” 1994.
- 2) ITU-T Rec. T.800 | ISO/IEC 15444-1, “Information Technology – JPEG 2000 Image Coding System Part 1: Core Coding System,” 2001.
- 3) <http://www.microsoft.com/windows/windowsmedia/forpros/wmphoto/default.aspx>
- 4) 松尾 他: “ブロックひずみの影響を考慮した JPEG2000 による JPEG 画像の再符号化方法”, 電子情報通信学会論文誌, Vol. J87-D-II, No.8, pp.1740–1744, Aug. 2004.
- 5) I. Bauermann *et al.*: “Further Lossless Compression of JPEG Images,” Proc. of Picture Coding Symposium (PCS 2004), Dec. 2004.
- 6) N. Ponomarenko *et al.*: “Additional Lossless Compression of JPEG Images,” Proc. of the 4th Intl. Symposium on Image and Signal Processing and Analysis (ISPA 2005), pp.117–120, Sep. 2005.
- 7) M. Stirner *et al.*: “Improved Redundancy Reduction for JPEG Files,” Proc. of Picture Coding Symposium (PCS 2007), Nov. 2007.
- 8) ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC: “Advanced Video Coding for Generic Audio-visual Services”, 2003.
- 9) 松田 他: “画像ごとに適応予測器と算術符号を最適化する可逆符号化方式”, 電子情報通信学会論文誌, Vol. J88-D-II, No.9, pp.1798–1807, Sep. 2005.
- 10) 奥村晴彦, “データ圧縮の基礎から応用まで”, C Magazine, Vol. 14, No.7, pp.13–35, July 2002.
- 11) J.L. Bentley *et al.*: “A Locally Adaptive Data Compression Scheme,” Communications of ACM, Vol.29, No.4, pp.320–330, Apr. 1986.
- 12) <http://my.smithmicro.com/win/index.html>
- 13) <http://www.elektronik.htw-aalen.de/packjpg/>
- 14) <http://compression.ca/act/act-jpeg.html>
- 15) <http://www.maximumcompression.com/data/jpg.php>