

JPEG 準拠のビットストリームに 可逆変換可能な画像符号化方式

小池 弘幸^{†1} 松田 一朗^{†1}
 青森 久^{†1} 伊東 晋^{†1}

本稿では、JPEG 標準との互換性を考慮した、新しいタイプの画像符号化方式を提案する。この方式で符号化された圧縮データは、再量子化処理を伴わずに JPEG baseline 方式に準拠したビットストリームに変換できる。このため、JPEG に対応した画像アプリケーションであれば独自のプラグイン等を必要とせず、またフォーマット変換に伴う画質劣化を気にせずに圧縮データを利用することが可能となる。イントラ予測やレート歪最適化といった最新の技術に基づいて符号器を実装した結果、提案方式によって JPEG 2000 を上回る符号化効率を達成可能であることが示された。

An Image Coding Scheme Allowing Lossless Conversion to JPEG Compliant Bitstreams

HIROYUKI KOIKE,^{†1} ICHIRO MATSUDA,^{†1}
 HISASHI AOMORI^{†1} and SUSUMU ITOH^{†1}

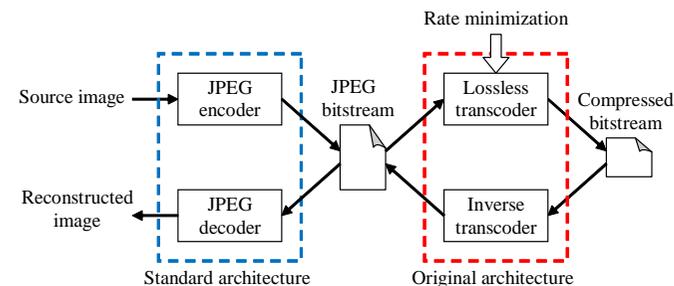
This paper proposes a new type of image coding scheme which provides some sort of compatibility with the JPEG standard. In this scheme, compressed data generated by our encoder can be converted to JPEG compliant bitstreams without additional quantization. Therefore, image applications which support JPEG format require no extra decoder nor plug-in for decoding the bitstreams. It allows us to use the image contents without any quality degradation which were usually caused by lossy transcoding. We have implemented our encoder based on modern coding techniques such as intra prediction and rate-distortion optimization. As a result, it is demonstrated that the proposed scheme attains better coding performance than the JPEG 2000 standard.

^{†1} 東京理科大学 理工学部
 Faculty of Science and Technology, Tokyo University of Science

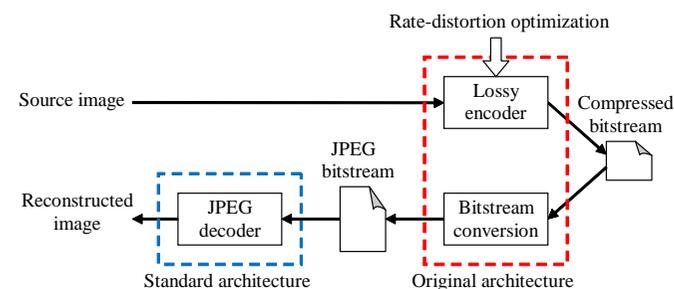
1. はじめに

JPEG baseline 方式¹⁾ は、画像符号化の国際標準方式として過去 20 年に渡り広く利用されており、現在ではほぼ全ての画像アプリケーションが JPEG 画像の入出力に対応している。JPEG 2000²⁾ や JPEG XR³⁾ といった、より高能率な符号化方式も標準化されているが、それらの画像形式に対応するアプリケーションが不足していることが普及の妨げとなっており、これまで撮影されたデジタル画像データの多くは、古い JPEG 方式で保存されているのが現状である。

そこで我々は、画質を一切劣化させずに既存の JPEG 画像を更に圧縮する、ロスレス再符号化方式について検討を行ってきた⁴⁾。この方式は、図 1 (a) に示すように JPEG 本来



(a) Lossless re-encoding method⁴⁾



(b) Proposed scheme

図 1 符号化処理及び復号処理の構造

Fig. 1 Structures of encoding and decoding processes.

のビットストリームと、圧縮後のビットストリームとを相互に変換する、可逆タイプのトランスコーダとみなすことが出来る。JPEG 画像のロスレス再符号化は、与えられたビットストリームより抽出した量子化済みの DCT 係数を、新たに量子化することなく再度可逆符号化することで実現される。この際、最新の符号化技術によって DCT 係数の冗長性を除去し、更に高効率なエントロピー符号化手法を導入することで、出力されるビットストリームのサイズを大幅に削減することが可能である。実際に、H.264 のイントラ予測と算術符号化を組み合わせた我々の検討方式では、元の JPEG ビットストリームに対して符号量を 18~28%削減することに成功している⁴⁾。このタイプのロスレス再符号化方式は、一般的なデジタルカメラのように撮影直後に JPEG による非可逆符号化が施され、原画像の情報が失われてしまっている場合に有効であるといえる。しかし原画像に対して直接 JPEG 2000 などの高効率な非可逆符号化方式を適用する場合と比較すると、一旦 JPEG ビットストリームを経てからロスレス再符号化を行った際の最終的なレート歪特性は、標準の JPEG エンコーダによる符号化条件の制約が足枷となり、一般に大きく劣る結果となる。

この問題を解決するため、本稿では、JPEG 標準との互換性を考慮した、新しいタイプの非可逆符号化方式を提案する。図 1 (b) に示すように、この方式の符号化処理は、一般的な非可逆符号化方式と同様に原画像を対象として行われるが、復号処理は JPEG に準拠したビットストリームへの変換処理を介して、既存の JPEG デコーダにより実行される。このため、JPEG ビットストリームへの変換時に新たな歪が生じないことを前提とすれば、原画像に対する再生画像の歪は符号化側で完全に制御可能である。このことは、量子化ステップサイズを始めとする様々な符号化パラメータを、レート歪最適化の手法に基づいて柔軟に決定できることを意味し、符号化効率の向上が期待できる。一方、既存のアプリケーションは内蔵の JPEG デコーダを通して画像データを劣化なく取り込むことが可能である。したがって、この符号化方式はある意味において JPEG 標準との互換性と高い符号化性能を兼ね備えた、新しい符号化の枠組みを提供するものであるといえる。

2. ブロック適応イントラ予測

JPEG baseline 方式は、DCT ベースの非可逆符号化アルゴリズムを採用しており、基本的な符号化処理はブロック単位で独立に実行される。このため、画像信号のブロック間相関は、DCT 係数の DC 成分を除いて全く利用されていない。提案方式ではこの欠点を克服するため、動画符号化の標準方式である H.264/AVC⁵⁾ のために開発された、ブロック適応イントラ予測の手法を導入している。具体的には、JPEG 標準で用いられる DCT のサイ

ズが 8×8 画素に固定されていることから、H.264/AVC⁵⁾ の High profile に規定されている 8×8 イントラ予測の手順に従って、ブロック毎に予測画像 $\hat{x}(i, j)$ ($i, j = 0, 1, \dots, 7$) を算出している。図 2 は、 8×8 イントラ予測で用いられる画素の位置と 9 通りの予測モード (8 方向+DC モード) を表している。

一般に、原画像 $x(i, j)$ の代わりに予測残差信号 $x(i, j) - x'(i, j)$ を符号化対象とすることで符号化効率の改善が期待できるが、提案方式においてビットストリームの可逆変換を実現するためには、最終的な再生画像 $x'(i, j)$ が JPEG 標準との整合性を満たしている必要がある。すなわち、 $x'(i, j)$ に相当する DCT 係数が JPEG 標準で規定された逆量子化器によって表現可能な値でなければならない。この条件を満足するため、提案方式では原画像の DCT 係数を対象に量子化すると共に、DCT 領域において予測残差信号の算出を行っている。

$$r(i, j) = Q[y(i, j), \Delta q(i, j)] - \hat{y}(i, j) / \Delta q(i, j), \quad (1)$$

ここで、 $Q[\cdot, \Delta q]$ はステップサイズ Δq の線形量子化器を表し、その構造は JPEG 標準に準拠したものとなっている。また、 $y(i, j)$ と $\hat{y}(i, j)$ はそれぞれ原画像と予測画像の DCT 係数である。量子化ステップサイズを格納した配列 $\Delta q(i, j)$ は量子化テーブルと呼ばれ、その要素は JPEG 標準において色信号毎に自由に設定可能である。一般に、上式によって得られる予測残差信号 $r(i, j)$ は実数となるが、式 (1) の量子化器 $Q[\cdot, \Delta q]$ の出力が整数であるため、その値を再量子化することなく可逆符号化することが可能である。

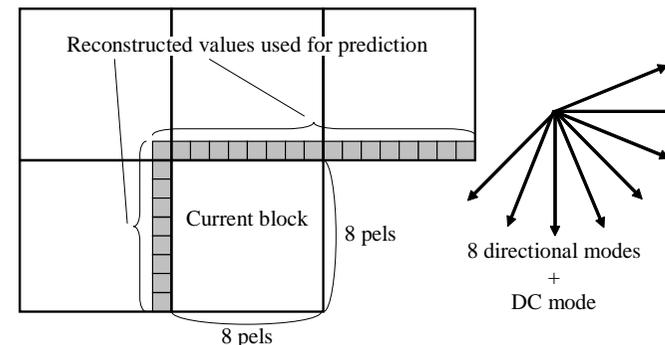


図 2 ブロック適応イントラ予測
 Fig.2 Block-adaptive intra prediction.

3. 予測残差信号の符号化

一般的な自然画像から得られる 2 次元 DCT 係数の確率密度関数 (PDF) は、以下の一般化ガウス関数⁷⁾ によってモデル化出来ることが知られている。

$$f(y | \mu, \sigma, c) = \frac{c \cdot \eta(\sigma, c)}{2\Gamma(1/c)} \cdot \exp\{-\eta(\sigma, c) \cdot |y - \mu|^c\}, \quad (2)$$

$$\eta(\sigma, c) = \frac{1}{\sigma} \sqrt{\frac{\Gamma(3/c)}{\Gamma(1/c)}}, \quad (3)$$

ここで、 $\Gamma(\cdot)$ はガウス関数、 μ 及び σ は各々 DCT 係数の平均値と標準偏差を表している。また、 c は関数の尖度を制御する形状パラメータである。本稿では、式 (1) によって算出された予測残差信号に関しても、上記の確率モデルが適用可能であると仮定する。但し、標準偏差 σ と形状パラメータ c は、周波数成分 (i, j) だけでなく画像信号の局所的な性質にも依存すると考えられる。そこで、全てのブロックを M 種類のクラスに分類すると共に、クラス毎 ($m = 1, 2, \dots, M$) に 2 種類のルックアップテーブル $\sigma_m(i, j)$ と $c_m(i, j)$ を参照することで、標準偏差と形状パラメータに適切な値を割り当てる。一方、ブロック適応イントラ予測の結果、予測残差の平均値は DC 成分も含めて 0 であると仮定する。以上より、当該ブロックが m 番目のクラスに属し、予測画像 $\hat{x}(i, j)$ の DCT 係数 $\hat{y}(i, j)$ が与えられた時、量子化出力 $q = Q[y(i, j), \Delta q] \in \mathbb{Z}$ の条件付き生起確率は次式によって見積もれる。

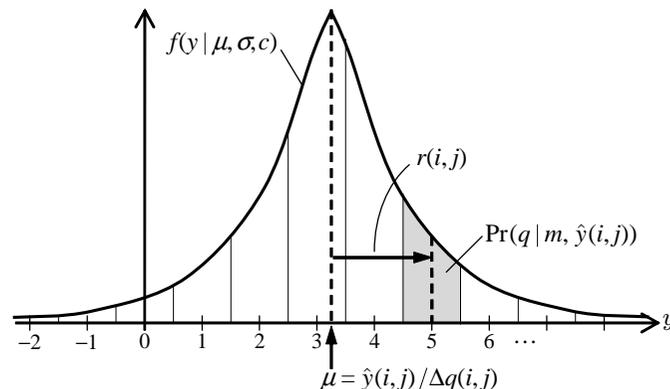


図 3 量子化済み DCT 係数の生起確率 ($q = 5$).

Fig. 3 Occurrence probability of a quantized DCT coefficient ($q = 5$).

$$\Pr(q | m, \hat{y}(i, j)) = \int_{q-0.5}^{q+0.5} f(y | \mu, \sigma, c) dy, \quad (4)$$

$$\mu = \hat{y}(i, j) / \Delta q(i, j), \quad (5)$$

$$\sigma = \sigma_m(i, j), \quad (6)$$

$$c = c_m(i, j). \quad (7)$$

これらの式は、一般化ガウス関数の分布中心を予測値の位置にシフトした上で、図 3 に示すような等間隔の量子化区間で積分することを意味する。このようにして、取り得る全ての量子化レベルについて算出された確率値は、予測残差 $r(i, j)$ を効率的にエントロピー符号化するために使用される。実際のエントロピー符号化器として、本稿では高速な多値算術符号化器として知られるレンジコーダ⁸⁾を採用している。

4. 符号化パラメータのレート歪最適化

提案方式では、付加情報として表 1 に示す符号化パラメータが必要となる。先に提案したロスレス再符号化方式⁴⁾では、量子化テーブル $\Delta q(i, j)$ の値が JPEG のエンコーダによって決められており、再符号化処理の間は一切変更できない。従って、原画像と再生画像との間の歪は不変であり、符号化パラメータの最適化はレート最小化問題として扱われていた。これに対し提案方式では、量子化テーブル及び量子化方法自体を変更することにより、符号

表 1 提案方式に必要な符号化パラメータ

Table 1 Parameters required for the proposed scheme.

| Description | Number of parameters | Allowed values |
|-------------------------------------|-------------------------|----------------|
| Quantization table $\Delta q(i, j)$ | $8 \times 8 \times N_q$ | 1, 2, ..., 255 |
| Standard deviation $\sigma_m(i, j)$ | $8 \times 8 \times M$ | See Eq.(13) |
| Shape parameter $c_m(i, j)$ | $8 \times 8 \times M$ | See Eq.(14) |
| Prediction mode d | N_b | 1, 2, ..., 9 |
| Classification label m | N_b | 1, 2, ..., M |

N_q : The number of quantization tables ($N_q = 1$ for monochrome images)

N_b : The number of 8×8 blocks

化器側で発生歪を自由に制御することが可能である。このことは、レート歪最適化の概念に基づいて各種符号化パラメータを柔軟に決定できることを意味する。提案方式におけるパラメータの最適化手順は、以下に述べるように大きく2つの段階からなり、1つ目の段階では量子化器の設計、2つ目の段階では残りのパラメータの最適化が行われる。これらの手順は、全てのパラメータが収束するまで交互に繰り返される。

4.1 ECSQ の量子化ステップサイズの最適化

この段階では、量子化ステップサイズを更新することにより、以下に示すコスト関数の最小化を図る。

$$J(\lambda) = D + \lambda \cdot R, \quad (8)$$

ここで、 D は原画像と再生画像との間の平均2乗誤差 (MSE) を測度とする歪であり、 R は再生画像を生成するのに必要な符号化レートを表す。提案方式では、ラグランジュの未定乗数に相当する正定数 λ を、レートまたは品質を制御する唯一のパラメータとして、符号化開始時に設定している。この種のレート歪最適化問題は、JPEG に準拠した符号器の性能を向上させることを目的として、既にいくつかの検討が行われている^{9),10)}。しかしながら、JPEG は量子化済みの DCT 係数にランレングス符号化を適用しているため、組み合わせ最適解の導出には、動的計画法のような複雑なアルゴリズムが必要であった¹⁰⁾。一方、提案方式では DCT 係数を個別に算術符号化しているため、ステップサイズ Δq で量子化された DCT 係数の歪 $D_{i,j}(\Delta q)$ と符号化レート $R_{i,j}(\Delta q)$ は、周波数成分毎に独立して求めることが出来る。

$$D_{i,j}(\Delta q) = \sum_{\text{for all blocks}} \{y(i,j) - \Delta q \cdot Q[y(i,j), \Delta q]\}^2, \quad (9)$$

$$R_{i,j}(\Delta q) = \sum_{\text{for all blocks}} -\log_2 \Pr(Q[y(i,j), \Delta q] | m, \hat{y}(i,j)). \quad (10)$$

このことは、各周波数成分の量子化ステップサイズを、下記のような分割された最適化問題により個別に決定できることを意味する。

$$\Delta q(i,j) = \operatorname{argmin}_{\Delta q \in \{1,2,\dots,255\}} \{D_{i,j}(\Delta q) + \lambda \cdot R_{i,j}(\Delta q)\}. \quad (11)$$

但し、式 (10) の予測値 $\hat{y}(i,j)$ は、隣接ブロックを参照するブロック適応イントラ予測と DCT によって求められるため、厳密には他の周波数の量子化ステップサイズの影響を受けて微小に変化する場合がある。そこで、提案方式では量子化テーブルの初期値を $\Delta q(i,j) = \lfloor 2\sqrt{\lambda/0.85} + 0.5 \rfloor$ ($i,j = 0,1,\dots,7$) とおき、各反復段階においてこれらの値

を徐々に更新させることとする。

さらに上記の最適化においては、量子化代表値が JPEG 準拠であるという条件の下、量子化方法あるいは量子化の閾値を変更することも可能である。そこで、本稿では DCT 係数の量子化方法に関してエントロピー拘束付きスカラー量子化器 (ECSQ)¹¹⁾ の概念を導入する。提案方式で用いられる ECSQ の手順は次式で表せる。

$$Q[y, \Delta q] = \operatorname{argmin}_{q \in \mathbb{Z}} \{(y - \Delta q \cdot q)^2 - \lambda \cdot \log_2 \Pr(q | m, \hat{y}(i,j))\}. \quad (12)$$

4.2 他のパラメータの最適化

上記の手順により、一旦 ECSQ の量子化ステップサイズが決定されると、原画像と再生画像との間の歪は固定となり、式 (8) の右辺第1項は無視できる。従って、ブロック毎の予測モード (d) やクラス番号 (m)、クラス毎の確率モデル ($\sigma_m(i,j)$ と $c_m(i,j)$) といったパラメータは、式 (8) の R が最小となるよう最適化される。更に、クラス数 (M) についても、ロスレス再符号化方式と同様の手法⁴⁾により、レート最小化の手順の中で画像毎に自動設定することが可能となる。なお、本稿においてパラメータ $\sigma_m(i,j)$ と $c_m(i,j)$ の値は、以下に示す組み合わせから選ばれるものとする。

$$\sigma_m(i,j) \in \{0.1 \cdot (5/4)^n \mid n = 0,1,\dots,31\}, \quad (13)$$

$$c_m(i,j) \in \{0.5, 1.0, 1.5, 2.0\}. \quad (14)$$

5. 実験結果

提案方式の符号化性能を評価するため、複数のモノクロ画像を用いてシミュレーション実験を行った。図4は、提案方式、JPEG 2000 (JasPer version 1.900.1¹²⁾, 9/7 wavelet), JPEG XR (Reference software version 1.8¹³⁾), JPEG baseline 方式 (IJG library version 6b¹⁴⁾) の各方式によるレート歪特性を比較した結果である。これより、広範囲の符号化レートにおいて、提案方式が比較方式を上回る符号化性能を達成している様子を確認できる。現行の標準方式の中で最も高性能とされる JPEG 2000 と比較した場合、PSNR の改善量は最大 1.2 dB に達している。図5に、提案方式と JPEG baseline 方式で用いられた量子化テーブルの例を示す。JPEG baseline 方式のエンコーダは、JPEG の規格書で推奨されている量子化テーブルのを、レートに応じたスケールリングを施して用いるのが一般的である。この推奨量子化テーブルは、人間の視覚特性に基づいて決定されたものであり、レート歪特性の観点からは必ずしも最適なものではない。一方、提案方式によって決定された量子化テーブルは比較的平坦な値となっていることが分かる。これは、量子化誤差は全ての周波数

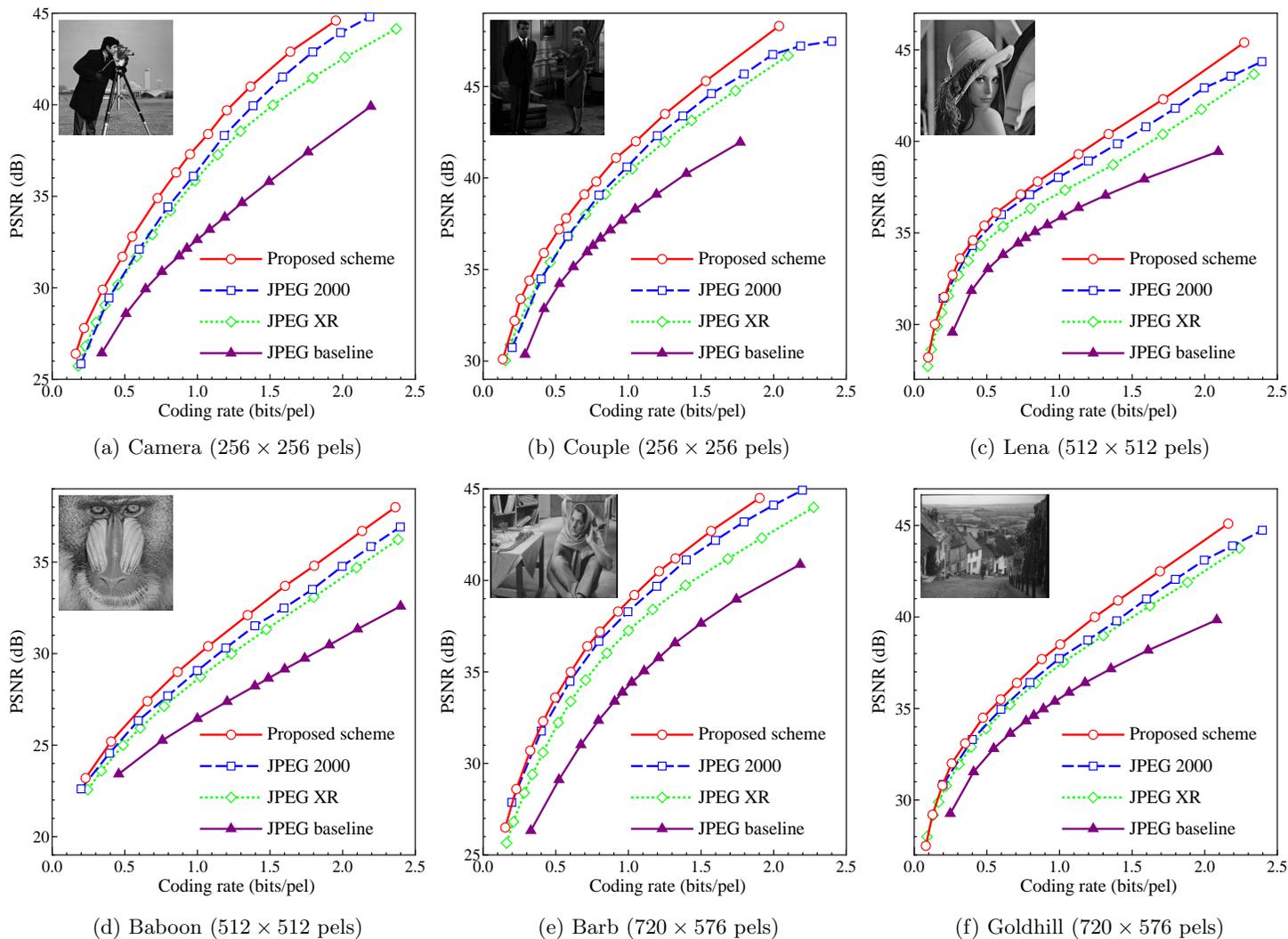


図 4 符号化特性
Fig. 4 Coding performance.

| | |
|-------------------------|-------------------------|
| 14 12 14 14 15 15 15 14 | 5 3 3 5 7 12 15 18 |
| 14 13 15 15 15 14 16 15 | 4 4 4 6 8 17 18 17 |
| 14 14 14 15 14 15 15 17 | 4 4 5 7 12 17 21 17 |
| 14 14 14 14 15 16 17 16 | 4 5 7 9 15 26 24 19 |
| 14 15 15 14 16 14 14 14 | 5 7 11 17 20 33 31 23 |
| 14 15 15 13 15 14 14 14 | 7 11 17 19 24 31 34 28 |
| 16 16 15 15 17 15 15 15 | 15 19 23 26 31 36 36 30 |
| 16 17 16 16 16 16 16 18 | 22 28 29 29 34 30 31 30 |

(a) Proposed scheme (b) JPEG baseline
0.95 bits/pel, 37.3 dB 1.76 bits/pel, 37.4 dB

図5 量子化テーブルの例 (Camera).
Fig.5 Examples of the quantization tables (Camera).

で均一化すべきという、古典的な最適ビット割り当て問題の結論に一致するものであるといえる¹⁵⁾。量子化テーブルの値の微小な変動については、周波数成分毎の DCT 係数の PDF の差異に起因するものであると考えられる。

6. む す び

本稿では、符号化と復号の処理が非対称な構造を持つ、新しいタイプの画像符号化方式を提案した。この方式では、符号化処理において直接原画像から圧縮後のビットストリームを生成するのに対し、復号処理では得られたビットストリームを一旦 JPEG 準拠の形式に可逆変換し、これを標準の JPEG デコーダに入力させることで再生画像を得る。すなわち提案方式は、復号処理の過程で生成される中間ファイルとして、JPEG 準拠のビットストリームを出力する、特殊な非可逆符号化方式であると捉えることができる。この中間ファイルは既存の JPEG デコーダで問題なく復号できるため、各アプリケーションは提案方式に対応したデコーダやプラグインを個別に用意する必要がない。このような符号化の枠組みは、新しい符号化方式の普及の妨げとなっていた互換性の問題に対し、一つの解決策を示すものであるといえる。シミュレーションの結果、提案方式は古い JPEG 標準との互換性を考慮しているにもかかわらず、JPEG 2000 などの後継の符号化方式を上回る符号化効率を達成していることが示された。

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 — Part1: Core Coding System," 2001.
- 3) ITU-T Rec. T.832 | ISO/IEC 29199-2, "Information Technology — JPEG XR Image Coding System Part2: Image Coding Specification," 2009.
- 4) 松田一朗, 橋本峻弥, 須田貴志, 池田悠, 青森久, 伊東晋, "イントラ予測と算術符号を用いた JPEG 画像のロスレス再符号化," 電子情報通信学会論文誌, vol.J94-D, no.8, Aug. 2011 (掲載予定).
- 5) ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC, "Advanced Video Coding for Generic Audiovisual Services," 2003.
- 6) D. Marpe, T. Wiegand and S. Gordon, "H.264/MPEG4-AVC Fidelity Range Extensions: Tools, Profiles, Performance, and Application Areas," *Proc. of 2005 IEEE International Conference on Image Processing (ICIP 2005)*, vol.1, pp.593–596, Sep. 2005.
- 7) F. Müller, "Distribution Shape of Two-Dimensional DCT Coefficients of Natural Images," *Electronics Letters*, vol.29, no.22, pp.1935–1936, Oct. 1993.
- 8) 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.
- 9) S. W. Wu and A. Gersho, "Rate-Constrained Picture-Adaptive Quantization for JPEG Baseline Coders," *Proc. of 1993 International Conference on Acoustics, Speech, and Signal Processing (ICASSP-93)*, vol.5, pp.389–392, Apr. 1993.
- 10) M. Crouse and K. Ramchandran, "Joint Thresholding and Quantizer Selection for Transform Image Coding: Entropy-Constrained Analysis and Applications to Baseline JPEG," *IEEE Trans. on Image Processing*, vol.6, no.2, pp.285–297, Feb. 1997.
- 11) G. J. Sullivan, "Efficient Scalar Quantization of Exponential and Laplacian Random Variables," *IEEE Trans. on Information Theory*, vol.42, no.5, pp.1365–1374, Sep. 1996.
- 12) <http://www.ece.uvic.ca/~mdadams/jasper/>
- 13) ITU-T Rec. T.835 | ISO/IEC 29199-5, "Information Technology — JPEG XR Image Coding System Part5: Reference Software," 2010.
- 14) <http://www.ijg.org/>
- 15) N. S. Jayant and P. Noll, "Digital Coding of Waveforms: Principles and Applications to Speech and Video," Prentice-Hall, 1984.