

# 再帰型イントラ予測を用いた JPEG 画像のロスレス再符号化

## Lossless Re-encoding of JPEG Images Using Recursive Type Intra Prediction

橋本 峻弥† Syunya HASHIMOTO    小池 弘幸† Hiroyuki KOIKE    松田 一朗† Ichiro MATSUDA    青森 久† Hisashi AOMORI    伊東 晋† Susumu ITOH

### 1. はじめに

JPEG[1]は、その策定以来 20 年近くを経た現在も静止画像の符号化方式として幅広く利用されており、既存の画像データの多くは JPEG 形式で保存されている。しかしこれらの JPEG 画像を、JPEG 2000[2]などのより高性能な方式で再符号化しようとしても、非可逆符号化特有の画質劣化が重畳されるため、一般に本来の符号化性能を発揮することは困難である。そこで我々は、JPEG 画像を一切劣化させることなく更に圧縮する、ロスレス再符号化方式について検討している[3]。この方式では、JPEG 画像から抽出した量子化済み DCT 係数の冗長度を削減するため、H.264[4]と同様なブロック単位のイントラ予測手法を適用していた。本稿では、上記再符号化方式の性能を更に改善するため、画像毎に設計した予測器をブロック内の予測値に対して再帰的に適用する、新しいイントラ予測方式の導入を図る。

### 2. JPEG 画像のロスレス再符号化

図 1 に提案方式の再符号化手順を示す。まず、JPEG の符号化データをハフマン復号し、量子化済みの DCT 係数  $y(i, j)$  を抽出する。次に、 $y(i, j)$  の冗長度を削減するため、ブロック間の相関を利用したイントラ予測を適用する。但し、DCT 領域では近傍ブロックの構造的なテクスチャ情報を利用した予測が難しいため、局所復号された隣接ブロックの再生値  $x'(i, j)$  を用いて、画素領域で予測値  $\hat{x}(i, j)$  を算出する。更に  $\hat{x}(i, j)$  に DCT を適用し、量子化ステップサイズ  $\Delta Q$  の逆数を乗じた上で  $y(i, j)$  との差分を求めることにより、予測誤差  $r(i, j)$  を得る。一般に  $r(i, j)$  の値は実数となるが、 $y(i, j)$  が整数値であることから取り得る値は離散的となり、再量子化することなく可逆符号化可能である。提案方式では、ブロックのクラス分類を用いた確率モデルの反復推定手法[3]に基づいて、予測誤差  $r(i, j)$  を多値算術符号化している。

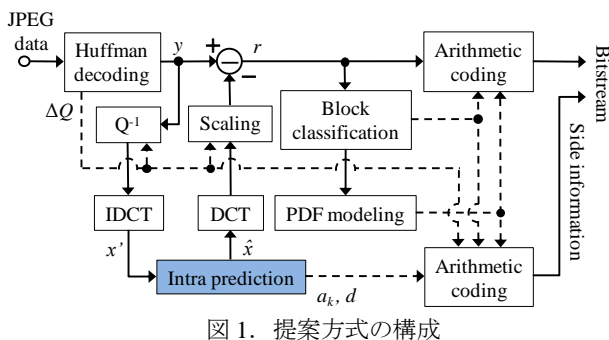


図 1. 提案方式の構成

### 3. 再帰型イントラ予測

文献[3]では、上述のイントラ予測において、H.264 の High Profile に規定された「8×8 イントラ」と呼ばれる 9 通りの予測モード[4]をブロック毎に適応選択していた。本稿ではより自由度の高い予測を実現するため、筆者らが開発したブロック適応時空間予測[5]の技術を導入する。ここで図 2 の  $B_n$  を予測対象ブロックとしたとき、予測に利用可能な再生値  $x'(i, j)$  は、同図の緑色で示した領域のみ存在する。そこで、注目画素  $p_0 = (i, j)$  を  $B_n$  内でラスタ走査させ、再帰的に以下の式を実行することで予測値  $\hat{x}(p_0)$  を逐次算出する。

$$\hat{x}(p_0) = \sum_{k=1}^K a_k \cdot \tilde{x}(p_k) \quad (1)$$

$$\tilde{x}(p_k) = \begin{cases} \hat{x}(p_k) & (p_k \in B_n) \\ x'(p_k) & (\text{otherwise}) \end{cases} \quad (2)$$

上式の  $a_k$  は、図 2 のように配置された参照画素群  $p_k (k = 1, 2, \dots, K)$  に対する予測の重み(予測係数)を表している。また、 $\hat{x}(p_k)$  は参照画素の値を表しており、(2) 式のように  $p_k$  がブロック  $B_n$  の内部に位置するときは既に算出済みの予測値  $\hat{x}(p_k)$  を再帰的に使用し、それ以外ではブロック  $B_n$  の外部に存在する再生値  $x'(p_k)$  を用いる。但し、参照画素  $p_k$  がブロック  $B_n$  の右側に存在する場合には、上方の再生値をコピー (パディング) した値を  $x'(p_k)$  とみなす。

この予測手法は、ブロック適応時空間予測[5]において時間方向の参照画素数を 0 とした場合に相当し、予測係数  $a_k$  の値を変化させることで様々なタイプの予測を実現できる。そこで画像毎に  $D = 60$  個の予測器を用意し、それらをブロック毎に適応選択した際の予測誤差電力を目的関数として、非線形最適化のアルゴリズムとして知られる準ニュートン法に基づいて、それぞれの予測係数を決定している。更に、得られた予測係数を  $\Delta a = 1/32$  の精度で量子化し、確率モデルの反復推定手順の中で徐々に不要な予測器を削除する

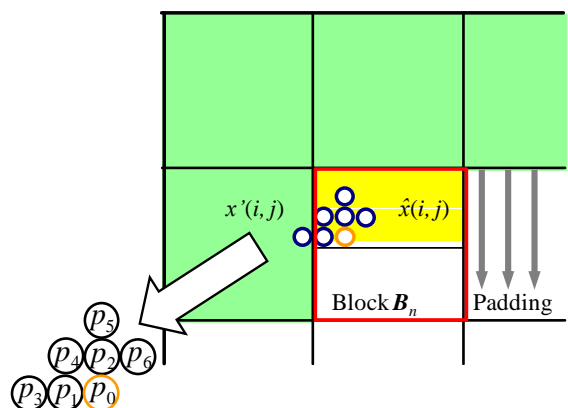


図 2. 再帰型イントラ予測

†東京理科大学 理工学部 電気電子情報工学科

ことで、画像毎に適切な数 $D$ の予測器を導出している。最終的に、 $D$ 組の予測係数およびブロック毎の予測器の選択情報 $d \in \{1, 2, \dots, D\}$ は付加情報としてビットストリームに多重化される。

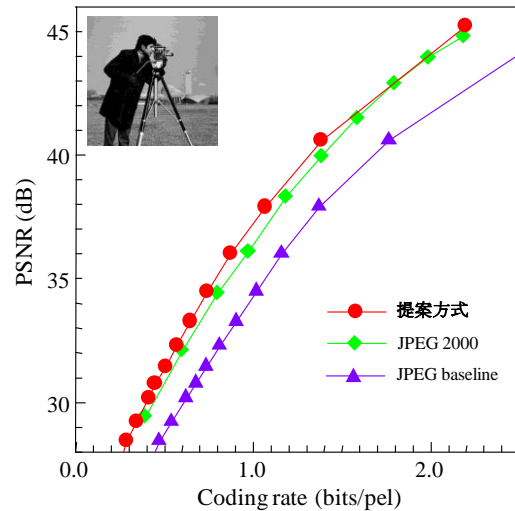
4. 特性評価とまとめ

提案方式の性能を評価するため、標準的な JPEG エンコーダ (baseline 方式) で符号化されたモノクローム静止画像を対象として再符号化を試みた。表 1 は、JPEG の画質を制御するパラメータ (quality factor) を 75 と設定したときの符号レートを比較した結果である。ここで「StuffIt」は、JPEG や MP3 形式のファイルに特化したアルゴリズムを備える Smith Micro 社の商用圧縮ツール (StuffIt Deluxe 2010[6]) で圧縮した結果であり、「 $8 \times 8$  イントラ」は、H.264 と同様なイントラ予測を用いた文献[3]の方式による再符号化結果を表している。また「提案方式」は、予測次数を $K = 6$ と設定した際の結果であり、カッコ内の数値は最終的に使用した予測器の数である。これより、多くの画像において提案方式が最も優れた符号化効率を示しており、再符号化前の JPEG baseline 方式の符号量と比較すると、平均で 33%の削減を達成していることがわかる。

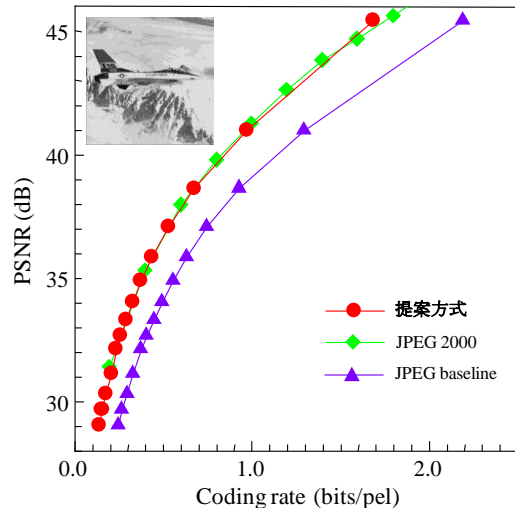
次に、JPEG baseline 方式と提案方式を組み合わせた際の、原画像に対するレート歪特性を調査した。但し、標準的な JPEG エンコーダの設定では、視覚特性に基づいた周波数成分の重みを考慮した量子化テーブルを用いているため、レート歪特性の評価では不利となる。そこで、量子化テーブルの値 $\Delta Q$ を一律に設定した条件で実験を行った。図 3 より、JPEG baseline 方式と提案方式の組み合わせによって、JPEG 2000 (実数型 9/7 wavelet)[2]に匹敵する符号化効率を実現している様子を確認できる。このとき、図中の提案方式と JPEG baseline 方式の符号化データは互いに無歪で変換できるため、JPEG に対応したアプリケーションであれば、新たな劣化を気にすることなく、再生画像を利用することが可能となる。

表 1. 符号化レートの比較 (bits/pel)

Image	提案方式	$8 \times 8$ イントラ	StuffIt	JPEG baseline
Camera	<b>0.991</b> (17)	1.009	1.050	1.308
Couple	<b>0.791</b> (14)	0.804	0.815	1.047
Noisesquare	1.111 (4)	<b>1.078</b>	1.151	1.467
Airplane	<b>0.779</b> (29)	0.792	0.781	1.034
Baboon	<b>1.695</b> (41)	1.703	1.741	2.103
Lena	<b>0.848</b> (30)	0.857	0.859	1.131
Lennagrey	<b>0.736</b> (30)	0.741	0.743	0.994
Peppers	<b>0.762</b> (33)	0.784	0.763	1.031
Shapes	0.465 (29)	0.473	<b>0.464</b>	0.707
Baloon	0.381 (39)	0.389	<b>0.378</b>	0.577
Barb	<b>0.939</b> (58)	0.995	1.001	1.321
Barb2	<b>1.017</b> (45)	1.046	1.048	1.366
Goldhill	0.914 (41)	0.927	<b>0.910</b>	1.176
average	<b>0.879</b>	0.892	0.900	1.174



(a) Camera (256x256 pels)



(b) Airplane (512x512 pels)

図 3. レート歪特性

【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] 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).

[4] ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC, “Advanced Video Coding for Generic Audiovisual Services,” 2003.

[5] 海野恭平, 青森久, 松田一朗, 伊東晋: “動画画像符号化のためのブロック適応時空間予測器の最適化”, 電子情報通信学会技術研究報告, Vol.110, No.148, pp.71-76, July 2010.

[6] <http://www.stuffit.com/>