

静止画像符号化方式

石川 孝明[†]

1. ま え が き

インターネットが普及し、膨大な量の画像や動画がデジタルデータとして流通している。多くの人がパソコンやスマートフォンを常用し、ブラウザを開けば文字と画像で美しくレイアウトされたコンテンツが表示され、スマートフォンのカメラで気軽に写真を撮り、友人と共有できる。この一連の流れの中で、一般のユーザが符号化という言葉を意識する瞬間は存在しない。符号化の存在を感じる人は、一部の研究者や業界関係者、そして本稿の読者だけと推察される。それほどに符号化技術は広く社会に浸透しているといえる。

本稿では、符号化技術の集合体である静止画像符号化方式について解説する。JPEG, JPEG 2000, そしてJPEX XRの3つの方式について構造を概説し、特徴となる処理部や機能について述べる。最後に、国際標準化について述べる。

2. JPEG 符号化方式

JPEG (Joint Photographic Experts Group) は、ほとんどのデジタルカメラで採用されている画像を圧縮するための符号化方式である。JPEGの特徴として、高い圧縮率と演算量の低さがある。性能と処理負荷のバランスの良さから、後継となる方式のJPEG 2000やJPEG XRが作られた現在においても、依然として幅広いアプリケーションに採用されている。

2.1 要素技術

JPEG符号化方式は、以下に示す複数の要素技術に支えられている。各要素を組合せることで高い圧縮率を実現している。

- ・ 色空間変換
- ・ DCT (Discrete Cosine Transform)
- ・ DCT係数の量子化
- ・ 予測符号化

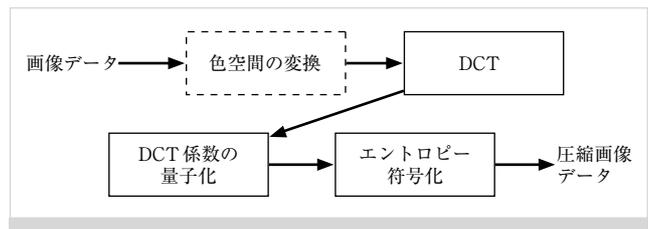


図1 JPEG画像の符号化手順

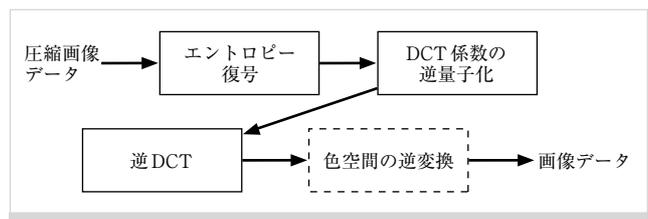


図2 JPEG画像の復号手順

- ・ ランレングス符号化
- ・ ハフマン符号化

JPEG符号化方式の符号化器と復号器の処理フローを、図1と図2に示す。デジタルカメラと画像閲覧ソフトを例にすると、まず、デジタルカメラに組込まれた符号化器が、撮影と同時に圧縮された画像データを生成する。その後、パソコンに取り込まれた圧縮画像データは、復号器を実装したソフトウェアによって画像データに戻されディスプレイに表示される。

2.2 圧縮率を高める工夫

図1と図2に示した処理フローには4つの処理部しか示されていない。しかし、各処理部には高効率な符号化を実現するために、入力画像の特徴や人間の視覚特徴を利用するさまざまな工夫が施されている。以下にそれらを示す。なお、他の符号化方式にも類似した処理部が含まれている。

2.2.1 色空間変換と間引き

デジタルカメラで撮影した瞬間の画像は、RGBの3色で表現されるが、JPEG符号化では始めにRGB色空間をYCbCr色空間に変換する。これは、人の視覚特性が輝度成

[†] 早稲田大学 国際情報通信研究センター

"Basic Technologies of Image Coding (11): JPEG Still Image Coding and Standardization" by Takaaki Ishikawa (Global Information and Telecommunication Institute, Waseda University, Tokyo)



分(Y)に対して強く反応し、色差成分(CbおよびCr)に対して弱く反応する特性を利用するためである。Y成分はそのまま残し、色差成分(CbおよびCr)を間引きすることで、視覚的な影響を最小限に抑えつつ、情報量を大幅に削減できる¹⁾。なお、色空間変換と信号の間引きは、本来のJPEG符号化方式では定められておらず、JPEGで圧縮したデータを扱うためのファイルフォーマットに定められている²⁾。

2.2.2 DCT

YCbCrの各成分を8×8画素ごとのブロックに区分して、Discrete Cosine Transform (DCT)を施す。DCTは、入力となる信号をコサイン波の成分量を表す係数値の集まりとして表現する。各周波数成分には、空間的な変化の乏しい波に対応する係数から変化の激しい波に対応する係数が存在し、最も変化の乏しい波に対応する係数をDC成分と呼び、それ以外をAC成分と呼ぶ。すなわち、64個の画素値は1つのDC成分と63個のAC成分に変換される。

自然画像にDCTを適用すると、DC成分が大きな値を持ち、AC成分が小さな値になる性質がある。さらに、人間の視覚特性には小さな波の変動を敏感に検知できないという性質があるため、AC成分の情報がある程度失っても劣化として知覚されにくい。このような特性を利用して係数を量子化する。

2.2.3 エントロピー符号化

量子化された係数群から冗長な情報を取り除き、より少ないビット数で画像を表現する仕組みがエントロピー符号化である。DCT係数のDC成分は、隣接するブロックのDC成分と近い値を取る傾向がある。これは、局所的な画素値の変化が急激には変化しない自然画像において顕著に現れる。この特性を利用して、DC成分を予測符号化する。一方、AC成分はランレングス符号化とハフマン符号化を組合せてビット列へと変換する。

2.2.4 EOB (End of Block)

AC成分の符号化では、DCTで得られた係数値を効率よく扱うために、ジグザグスキャンとEOBを使う³⁾。図3(a)にDCT後の係数の状態を概念図として示す。2次元平面において、DC成分は左上に現れ、それ以外の数値はAC成分である。係数値を量子化すると、絶対値が小さな数値は0に置き換わるため、図3(a)の破線で括られた領域には0が集中することになる。

この特性を圧縮処理で利用するために、まず図3(b)のように、係数値をジグザグスキャンする。ジグザグスキャンは、2次元の数値を1次元のデータに変換するための処理である。量子化済みの2次元DCT係数のブロックに適用することで、1次元データとして並べられる係数値はある係数以降がすべて0になることが多い。そこでJPEG符号化では、これ以降は0が並んでいることを示す印としてEOBを導入している。EOBを導入することで、係数に0が並んでいることを少ない情報量で表現可能になる。

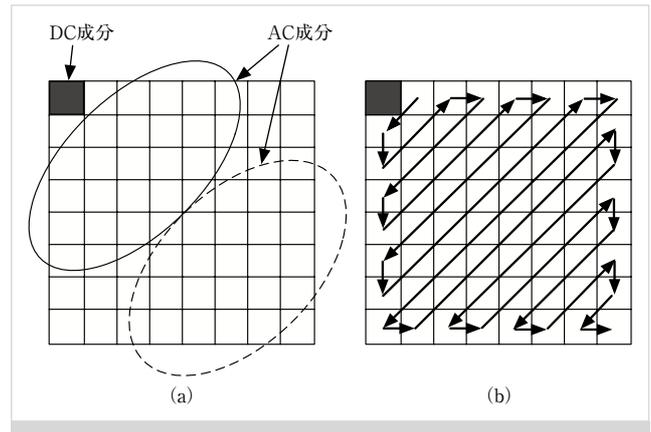


図3 係数成分とジグザグスキャン

3. JPEG 2000 符号化方式

前章で示したJPEG符号化方式と比較して、JPEG 2000符号化方式は符号化効率の向上に加えて、圧縮画像データの機能性を重視した符号化方式である。また、JPEG符号化方式の欠点である8×8画素に区分することで生じるブロックノイズの発生を本質的に解決し、低ビットレートでも画像が激しく崩れないなどの点で改善されている。

JPEG 2000符号化方式に関連する規格は、13個のパートから構成されている⁴⁾。そのうち符号化方式の中心を担う部分は、Part1: Core Coding Systemとして規格化されている。Part2以降には、動画像のファイルフォーマット、セキュリティ技術、APIや通信プロトコルへの拡張などが定義されている。

3.1 要素技術

JPEG 2000符号化方式は、以下の要素技術で構成されている。可逆符号化と非可逆符号化を同じ処理構造で実現している点や、ビットレートを指定して符号化できる点でJPEG符号化方式と違いがある。また、JPEGと比較してほぼ半分のビットレートで符号化できる。

- ・色空間変換
- ・DWT (Discrete Wavelet Transform)
- ・タイリング
- ・係数モデリング
- ・ビットプレーン展開と量子化
- ・エントロピー符号化
- ・符号順序制御

JPEG 2000符号化方式の符号化器の処理フローを、図4に示す。復号器は、JPEGと同様に符号化器の処理を逆向きにたどる。ただしJPEG 2000符号化方式には、符号化器側に符号順序を制御する処理が加えられている。そのため、エントロピー符号化が施された後のビット列を組み替えるだけで、目的の異なるアプリケーションに利用できる。

圧縮画像データの一部を復号して情報を抽出する機能を

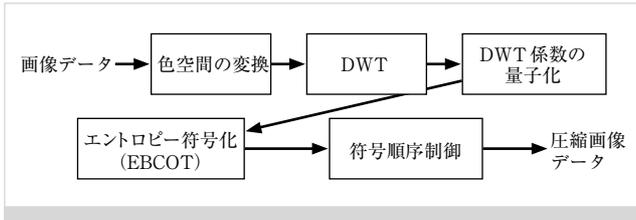


図4 JPEG 2000画像の符号化手順

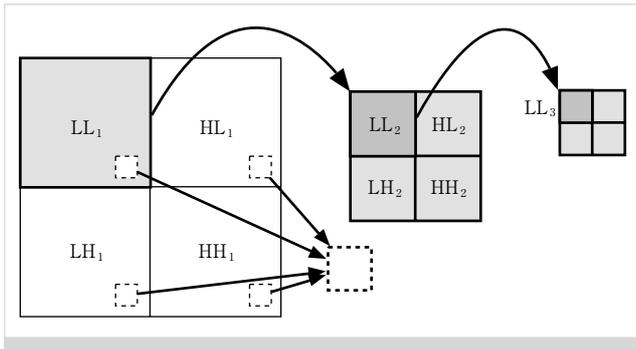


図5 入力画像のサブバンド分割

スケーラビリティと呼ぶ。JPEG 2000符号化方式には、空間スケーラビリティおよび品質スケーラビリティがあり、画像の一部の領域だけを復号することや、特定の色コンポーネントだけを復号することも可能である。

3.2 圧縮率を高める工夫

JPEG 2000符号化方式の圧縮率を高める工夫のうち、2次元ウェーブレット変換とEBCOT (Embedded Block Coding with Optimized Truncation) について述べる。

3.2.1 2次元ウェーブレット変換

JPEG 2000符号化方式では、入力画像データを区別せず2次元のウェーブレット変換を画像全体に施す。これにより、JPEG符号化方式で生じていたブロックノイズの問題を本質的に解決している。実際の処理手順は、水平方向と垂直方向のそれぞれに1次元のウェーブレット変換を施す。その結果、4つの領域に低周波数成分と高周波数成分に対応する係数が並ぶことになる。ウェーブレット変換後のDWT係数の状態を図5に示す。信号処理の分野ではこのような周波数分割をサブバンド分割と呼んでいる。

左上のサブバンドは低周波数成分に対応し、人間の視覚には縮小画像のように見える。その他のサブバンドは、原信号のエッジ成分と細かな変動を表現している。低周波数成分のサブバンドを再帰的に2次元ウェーブレット変換すると、入力信号のエネルギーが少数の係数で構成される領域に集約される。その結果、JPEG符号化と同様に、人間の視覚特性を利用した量子化が可能になる。また、高周波数成分のサブバンドでは、エッジ成分が大きな値を持つ係数

として抽出されているため、効率のよい圧縮が可能になる。

なお、JPEG 2000符号化方式では、入力画像を一定サイズに区切り、それぞれを独立な入力画像として扱うこともできる。この機能はタILINGと呼ばれる。極めて高解像度な入力画像を効率よく符号化するために有効であるが、圧縮率を高めるとJPEG符号化同様にタイルの境界に視覚的なノイズが生じる。

3.2.2 EBCOT

JPEG 2000符号化方式では、2次元ウェーブレット変換で得られた係数群を効率良く符号化するために、EBCOT (Embedded Block Coding with Optimized Truncation) と呼ばれるエントロピー符号化を導入している。EBCOTでは、まず係数群を 32×32 や 64×64 などの、コードブロックと呼ばれる一定サイズのブロックに区分する。次に、ブロック内の係数値を2進数で表してレイヤ化するビットプレーン展開を施し、レイヤごとに周辺ビットの関連性を3つの符号パス (significance propagationパス, magnitude refinementパス, cleanupパス) を用いて表現する。これは、係数値そのものではなく、係数の状態を符号化することを意味している。係数モデリングの結果を算術符号化することで、高い圧縮効率を実現している。

3.3 機能性

再帰的な2次元ウェーブレット変換とEBCOTにおけるコードブロック単位の処理により、JPEG 2000符号化方式は複数のスケーラビリティを有している。符号化におけるスケーラビリティは、ビット列の一部を復号するだけでもアプリケーションに利用でき、残りのビット列があればその精度が向上する機能である。ビット列を追加的に復号することで、空間解像度や画像品質が変化する。また、図5の破線で表したように、画像の部分的なアクセスを可能にしている。

3.4 空間スケーラビリティ

図5に示したサブバンドをLL3からLL1に向かって順番に符号化すれば、縮小画像に対応する情報がビット列の前半部に集約される。例えば、分割レベル2で得られる7つのサブバンド (LL3, HLk, LHk, HHk (k=2, 3)) に対応するビット列だけを復号すれば、原画像と比較して解像度が半分の画像を再生できる。

3.5 品質スケーラビリティ

ウェーブレット係数をビットプレーン展開し、最上位ビット (MSB: Most Significant Bit) から順番に符号化することで品質スケーラビリティを実現している。例えば、ある係数値が100である時、表1のようにMSBだけを符号化

表1 ビットプレーン展開の例

MSB(64)	(32)	(16)	(8)	LSB(4)
1	1	0	0	1



すれば、64を復号できる。MSBから2つ目のビットまでを符号化すれば96を復号できるため、係数値の真値である100に近づくことがわかる。

4. JPEG XR 符号化方式

JPEG XR符号化方式は、JPEGとJPEG 2000の中間的位置づけにある方式である。符号化効率はJPEGより高く、JPEG 2000に及ばない。その一方で、スケーラビリティなどの機能性を有し、JPEG 2000よりも低い処理量で符号化でき、ハードウェアでの実装が容易であるという利点がある。JPEG 2000符号化方式がJPEG符号化方式の後継規格として利用されていない一つの理由として、JPEG 2000符号化方式の処理の複雑さが指摘されている。後に規格化されたJPEG XR符号化方式は、この欠点を解決した方式である。

Microsoft社が開発したHD Photoを基盤とするJPEG XR符号化方式は、Microsoft社のブラウザやAdobe社のFlash Playerでサポートされ、ハードウェア化も開始されているため⁵⁾、今後の普及が期待される。

4.1 要素技術

JPEG XRは以下の要素技術で構成される。JPEG 2000と同様に、可逆符号化と非可逆符号化を同じ処理構造で実現でき、空間および品質スケーラビリティを有する。

- ・色空間変換
- ・重複双直交変換
- ・可変サイズタイリング
- ・係数予測符号化
- ・ハーモニック量子化
- ・動的な係数スキャン
- ・エントロピー符号化
- ・符号順序制御

JPEG XR符号化方式の符号化器の処理フローを、図6に示す。復号処理は他の方式と同様に符号化の逆処理である。

4.2 低処理負荷で圧縮率を高める工夫

JPEG XR符号化方式は、圧縮効率を高めるために周波数変換、量子化、エントロピー符号化を施す点でJPEGとJPEG 2000と同じだが、各処理部では処理負荷を低く抑えた方式が採用されている。

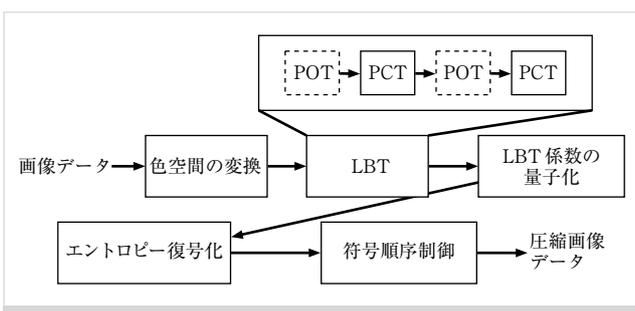


図6 JPEG XR画像の符号化手順

4.2.1 重複双直交変換

JPEG XRでは、JPEGのように入力画像を一定のブロックサイズに分割して独立に処理する。16×16画素の領域をマクロブロック (MB) と呼び、MBをさらに4×4画素に分割した領域において周波数変換を施す。JPEG XRの周波数変換は、重複双直交変換 (LBT: Lapped Biorthogonal Transform) を採用している。実際の変換処理は、POT (Photo Overlap Transform) と PCT (Photo Core Transform) の2つの処理の組合せを2回繰り返す。POTは4×4画素のブロック境界をまたぐ処理のため、高圧縮時に境界に生じるノイズを低減する役割がある。一方、PCTを2回実行することで3レベルの階層化を実現している。POTとPCTの内部は、2×2の2次元アダマール変換や係数値のスケーリング、係数の入れ替えなどの単純な演算処理で構成されるため、処理負荷を低く抑えることができる。なお、POTを常に実行する必要はない。

4.2.2 エントロピー符号化

LBTで得られる係数を予測符号化する。3レベルの階層のうち、最も低い周波数成分を表すDC係数は、上に位置するMBからの予測、左に位置するMBからの予測、上と左の両方のMBからの予測、予測なしの4つの予測モードで符号化される。また、それ以外の周波数成分を表すLP係数とHP係数については、水平方向と垂直方向の予測と予測なしの3つの予測モードがあり、最適なモードが動的に選択される。予測後の係数はランレングス符号化もしくは適応的可変長符号化される⁶⁾。適応的可変長符号化では、複数の可変長符号テーブルを予め定義されており、符号化処理の過程で適応的にテーブルを選択する。そのため、入力画像の特性が局所的に変化する場合でも、その性質に近いテーブルを選ぶことで符号化効率を高められる。また、最適な符号を符号化時に都度算出する処理ではないため、処理の複雑化を緩和し、特にハードウェア実装において優位である。

5. 国際標準化方式

前章までに解説した3つの符号化方式は、いずれも国際標準化された画像符号化方式である。方式名称であるJPEGは、Joint Photographic Experts Groupという標準化組織の名前でもある。JPEGの規格は、ISO/IEC JTC 1/SC 29とITU-Tの下部組織による共同作業で策定されている。

本稿で述べた各符号化方式の主なアプリケーションを表2にまとめる。各方式が広く利用されている背景として、国際標準化の議論に製品化を具体的に検討する団体や組織が参加していることがあげられる。

5.1 標準化の流れ

符号化の国際標準化は以下に示す手順に沿って進められる。最終的な方式を示した仕様書が完成するまでには、(4)と(5)のプロセスが数回繰り返され、その度に国際会議で



表2 応用アプリケーション

JPEG	JPEG 2000	JPEG XR
デジタルカメラ ブラウザ 画像ソフト全般	デジタルシネマ アーカイブ 免許証	ブラウザ

審議し、各国が投票権を持つ投票を経て規格が成立する。

- (1) 市場の要求や既存技術の分析
- (2) 要求機能の実現可能性の検証
- (3) 提案方式の募集
- (4) 提案方式の比較(実験と評価)
- (5) 仕様書の草稿の作成および修正
- (6) 仕様書の完成

5.2 コンフォーマンス

コンフォーマンスは、国際標準の符号化方式に対応した機器やソフトウェアを開発する時に、それらが正しく動作するかを検証するための適合試験である。実装の正しさを検証する手段が広く提供されることは、標準に準拠した製品を作るために重要な役割を果たす。また、標準化の過程においてもコンフォーマンスに関する議論を深めることで、規格書の解釈の違いから生じる問題が明らかになり、より明確な記述に改善できる。JPEG 2000符号化方式では、Part4にコンフォーマンスが規定されている。

5.3 参照ソフトウェア

参照ソフトウェアは、コンフォーマンスと同様に、標準に準拠した製品を作るために参考とするべき実装例である。例えばJPEG 2000ではPart5で参照ソフトウェアを規定しており、C言語で実装されたJasperと、Javaで実装されたJJ2000の2つが提供されている。

5.4 プロファイル

符号化方式の標準では、復号器の構成だけを規定するのが一般的である。これは、符号化器に対する制約を弱め、

独自のアルゴリズムを適用しやすくするためである。どのように入力画像が符号化されようとも、復号器が必ず画像を再生できればよいとの考え方である。符号化器に対する自由度は、符号化器を作るベンダー同士に競争をもたらし、各社、より符号化効率の高い、もしくは、高速に動作するアルゴリズムの考案を目指すことになる。

しかし、アプリケーションに応じて機能を絞り込んだり、復号器の実装規模を制限することを目的として、符号化器が生成するビットストリームに対して一定の制約が必要な場合がある。符号化方式の標準では、このような条件をプロファイルとして規定している。例えば、JPEG 2000符号化方式には2つのプロファイルが規格化されており、画像サイズやコードブロックサイズを制限している。上位のプロファイルは下位のプロファイルを包含するように規定されている。

6. む す び

本稿では、国際標準化された画像符号化方式としてJPEG, JPEG 2000, そしてJPEG XRについて解説した。

(2013年3月6日受付)

【文 献】

- 1) 末松, 山田: “画像処理工学”, コロナ社 (2000)
- 2) E. Hamilton: "JPEG File Interchange Format (JFIF)", C-Cube Microsystems (1992)
- 3) 小野, 渡辺: “国際標準画像符号化の基礎技術”, コロナ社 (1998)
- 4) 野水: “JPEG2000符号化方式解説”, トリケップス (2003)
- 5) 小高, 佐藤: “JPEG XRとそのデジカメへの適用”, 第33回VMA研究会 (July 2012)
- 6) 原: “JPEG XR画像符号化方式と性能評価”, トリケップス (2010)



石川 孝明 いしかわ たかあき 2005年, 早稲田大学大学院国際情報通信研究科修士課程修了。2007年, 同大学国際情報通信研究センター助手。2010年, 同センター招聘研究員となり, 現在に至る。画像符号化の研究に従事。ISO/IEC JTC1/SC29/WG11/Video小委員会およびWG1小委員会各委員。

