

JPEG2000はWG1の元、1997年11月に標準化が開始されました。後で詳しく述べますが、JPEG2000には数多くのPartが存在しており、最も基本のコアとなるPart-1は2001年1月に標準化が完了しています^(参考文献1)。

一方、同じSC29委員会には、動画像符号化の標準化作業を担当するWG11があります。このグループは別名、MPEG(Moving Pictures Experts Group)と呼ばれており、現在までにMPEG-1、MPEG-2、MPEG-4、MPEG-7を標準化してきました。現在の関心は、MPEG-4 Simple Profileの2倍の圧縮効率をめざすAVC(H.264)に移行しています。

1.1.2 JPEG2000標準化の背景

JPEG2000の標準化が開始された1997年では、現在のADSLのような高速のネットワーク環境は存在していませんでした。インターネットの利用も大部分が電話線を利用したダイヤルアップであったため、画像の伝送には非常に多くの時間を要するという問題がありました。狭帯域で画像を伝送するために、圧縮率を上げてJPEGファイルのサイズを小さくすることが頻繁に行われました。しかしJPEGでは、圧縮率を上げていくとブロック形状の境界が顕著に見えるほか、モスキート・ノイズという画質劣化が大きくなります。JPEG2000が生まれた背景には、このような圧縮率の向上への期待が大きく働いていました。

もちろん、圧縮率の向上以外の点でも、JPEGよりも優れた代替技術が望まれていました。それらについては、「1.2.1 JPEG2000の要求条件」で説明します。

1.2 JPEG2000の特徴

1.2.1 JPEG2000の要求条件

JPEG2000はJPEGにはない多くの機能を持っています。また、JPEGよりも優位な点も多いです。それらについて以下に説明します。以下の諸条件は、JPEG2000を標準化する際に、要求仕様書(Requirements Draft)の中で定義されていたものです。

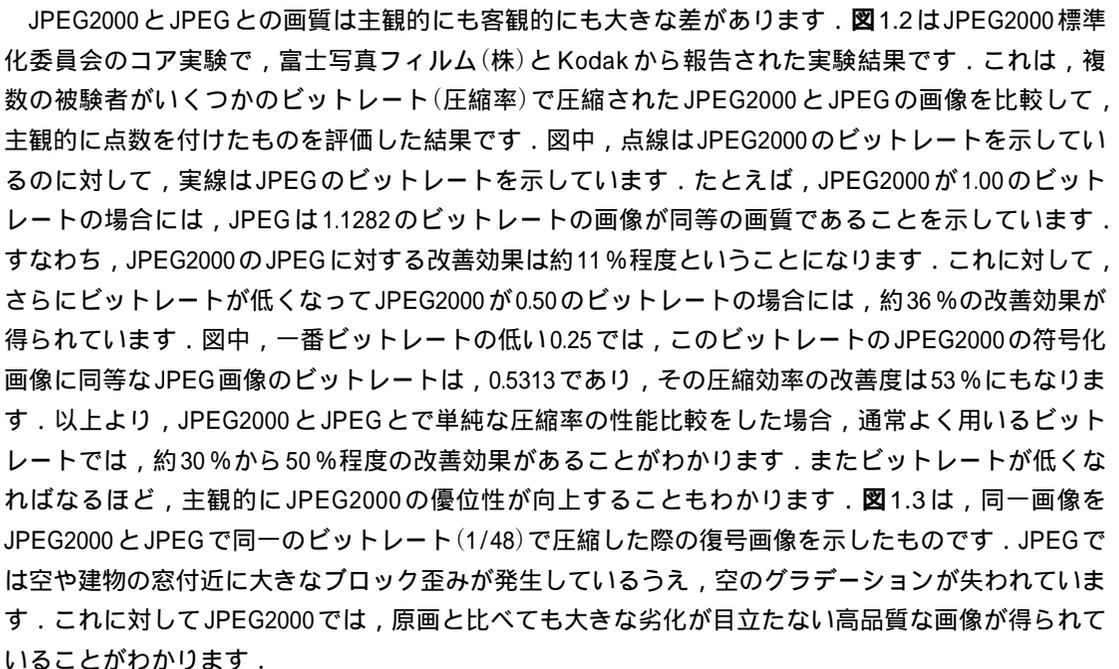
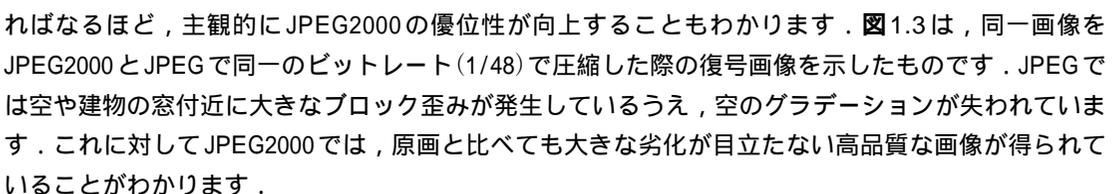
- ・ **高い圧縮効率**：すべての圧縮率においてJPEGを上回ること。とくに高圧縮時(0.25 bpp*以下)にJPEGで顕著になるノイズ・歪みが目立たない利点があること。現在DSC(Digital Still Camera)で使われているJPEGの圧縮率は5分の1～10分の1程度が多いので、この圧縮率でもJPEG2000がJPEGを大きく上回ることが要求されています。
- ・ **解像度**：医用画像や衛星写真画像のような、超高解像度画像に対しても対応できること。一般にレントゲン写真をスキャンした画像は、数千画素×数千ラインもの解像度になることが多いです。
- ・ **ビット深度**：医用画像や衛星写真画像、業務用画像などのビット深度が大きい画像に対しても対応できること。現在DSC(Digital Still Camera)で使われているJPEG圧縮では8ビットまでしか対応できません。将来的には10ビットや12ビットの画像を圧縮するニーズが高まると予想されます。この

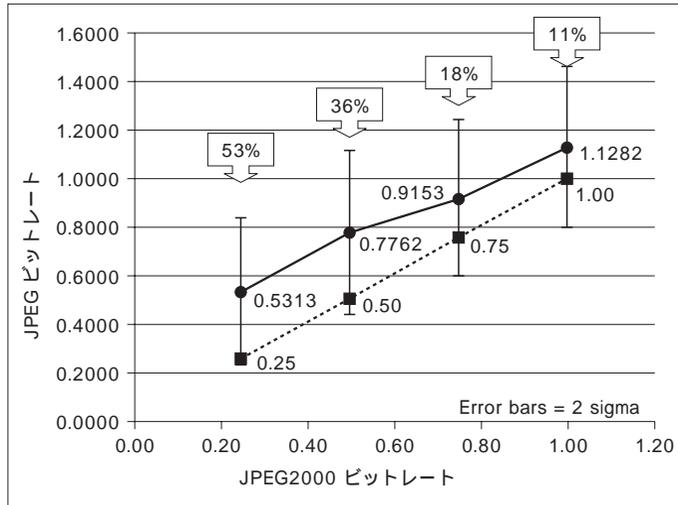
* bits per pixel, ビットレートともいう。

ときにはJPEG2000が有力な選択肢になると期待されています。

- ・**複数コンポーネントの圧縮**：複数コンポーネント(例：RGB，YCbCrなど)から構成されるカラー画像を圧縮できること．1コンポーネントのビット精度もすべてのアプリケーションをカバーするのに十分な精度を持っていること．これまではRGBのように3コンポーネントが主でしたが，最近ソニー(株)が開発した4色CCD(RGBE．Eはエメラルド)のように3コンポーネント以上の画像も，今後出現することが予想されます．
- ・**可逆・非可逆圧縮**：可逆圧縮が実現できること．また同一の手法で非可逆圧縮も可能であり，編集やビットレート変換などの操作が容易にできること．一度，画像を可逆圧縮しておけば，欠損のない画像をいつでも復元することができます．さらに用途に応じて，可逆・非可逆変換を自由に行いネットワーク越しに配信したりすることが容易にできます．
- ・**プログレッション**：画質プログレッション，解像度プログレッションなどのプログレッシブな機能を備えていること．またこれらを使い分けることで非常に柔軟な画像のハンドリングが可能なこと．
- ・**ランダムアクセス**：画面中の特定領域のみを重視して圧縮率を高く設定するなどのROI(Regions-Of-Interest)が実現できること．画像の特定部分に対して，符号化コードストリームを全部パースせずに直接アクセスできること．
- ・**エラー耐性**：コーデック側で強力なエラー耐性能力を持っていること．無線環境下やインターネットでのエラーやロスにも対応できること．

1.2.2 JPEGとの差異

JPEG2000とJPEGとの画質は主観的にも客観的にも大きな差があります． 図1.2はJPEG2000標準化委員会のコア実験で，富士写真フィルム(株)とKodakから報告された実験結果です．これは，複数の被験者がいくつかのビットレート(圧縮率)で圧縮されたJPEG2000とJPEGの画像を比較して，主観的に点数を付けたものを評価した結果です．図中，点線はJPEG2000のビットレートを示しているのに対して，実線はJPEGのビットレートを示しています．たとえば，JPEG2000が1.00のビットレートの場合には，JPEGは1.1282のビットレートの画像が同等の画質であることを示しています．すなわち，JPEG2000のJPEGに対する改善効果は約11%程度ということになります．これに対して，さらにビットレートが低くなってJPEG2000が0.50のビットレートの場合には，約36%の改善効果が得られています．図中，一番ビットレートの低い0.25では，このビットレートのJPEG2000の符号化画像に同等なJPEG画像のビットレートは，0.5313であり，その圧縮効率の改善度は53%にもなります．以上より，JPEG2000とJPEGとで単純な圧縮率の性能比較をした場合，通常よく用いるビットレートでは，約30%から50%程度の改善効果があることがわかります．またビットレートが低くなればなるほど，主観的にJPEG2000の優位性が向上することもわかります． 図1.3は，同一画像をJPEG2000とJPEGで同一のビットレート(1/48)で圧縮した際の復号画像を示したものです．JPEGでは空や建物の窓付近に大きなブロック歪みが発生しているうえ，空のグラデーションが失われています．これに対してJPEG2000では，原画と比べても大きな劣化が目立たない高品質な画像が得られていることがわかります．



データ引用：WG1N1583 by Fuji-Film, Kodak

図1.2 JPEG2000とJPEGとの主観評価結果



(a) JPEG2000(1/48)

(b) JPEG(1/48)

図1.3 JPEG2000とJPEGとの画質差(48分の1の圧縮率)

1.2.3 JPEG2000の要素技術

JPEG2000はデコーダの規格であり、エンコーダの規定は存在しません。したがって、エンコーダにはある程度の自由度が許されていることとなります。図1.4は、典型的なエンコーダの構成例を示したものです。

各要素技術の詳細については、後述します。エンコーダの構成としては、前段部のDCレベルシフト、ウェーブレット(Wavelet)変換、量子化、中段部のEBCOT*(ビットモデリング、算術符号化)、そして後段部のレート制御、ヘッダ生成、パケット(Packet)生成に分けると理解しやすいでしょう。なお、DCレベルシフトはオプション扱いです。

この中で、既存の画像圧縮手段とJPEG2000が大きく異なるユニークな技術は、なんといっても中段部のEBCOTであるといえます。また後段部のレート制御、ヘッダ生成、Packet生成も、EBCOTと深く関係していますので、その意味ではウェーブレット変換と量子化以外は、EBCOTがJPEG2000を特徴付けているといっても過言ではないでしょう。また、図中で量子化を迂回している経路が存在していますが、これは可逆圧縮を使う場合に使われます。

DCレベルシフト

DCレベルシフトは後段のウェーブレット変換を効率的に行うため、圧縮率の向上につながります。原理的にはRGB信号は正の値(符号なしの整数)を持っていますので、原信号のダイナミックレンジを半分にするレベルシフトを行うことで、圧縮効率が向上することを目的としています。したがって、YCbCr信号のCbやCrのように符号(正負両方あり)の整数値を持つ信号を原信号とする場合には、このレベルシフトは行いません。

ウェーブレット変換

JPEG2000ではウェーブレット変換(p.85参照)を採用した大きな理由としては以下が挙げられます。

- (1) 圧縮効率の向上が実現可能
- (2) 多重解像度の表現が容易に実現可能
- (3) Watermarkなどのセキュア技術が容易に実現可能
- (4) エラー耐性が容易に実現可能

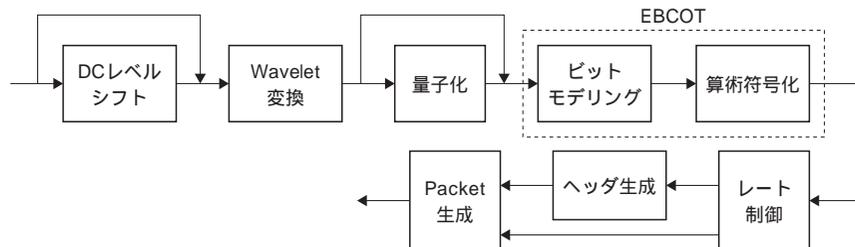


図1.4 JPEG2000 エンコーダのブロック図

* Embedded Block Coding with Optimized Truncation
エントロピー符号化の方式。

(5) 規格化当時、DCTの研究はすでに衰退しており、大学・企業の研究対象はウェーブレット変換符号化に移行

JPEG2000 Part-1で用いられているウェーブレット変換は、離散ウェーブレット変換(Discrete Wavelet Transform : DWT)であり、フィルタバンクで構成されます。

量子化

量子化手段とは、その前段部であるウェーブレット変換係数のビット精度を下げるための手段として用います。量子化を行った時点で絶対値が小さくなりますので、情報圧縮効果があります。JPEG2000ではスカラ量子化という手法が採用されています。量子化の際のステップサイズは、規格で定められた定義式によって導出されます。

EBCOT

EBCOTはエントロピー符号化を行う部分で、JPEG2000の特徴を最も象徴的に表している部分です。しかしビット単位の処理を行うために計算負荷が多いうえ、消費メモリも多いという問題点もあります。その反面圧縮効率は非常に高いものがあります。

レート制御

EBCOTの処理後、符号語が生成されます。レート制御部では、目標符号量に合わせるために、符号語の一部を切り捨てる操作を行います。どの符号語を切り捨てるかで画質が変動するため、レート制御はノウハウが活かされる部分です。

Packet生成

JPEG2000の符号化コードストリームは、Packetという情報の塊の単位で生成されています。このPacketは、Packetヘッダ(データ長や後述の符号化パスの情報などを含む)とPacket本体の二つから構成されます。符号化コードストリームがPacketごとに管理されているために、後述のプログレッションなどのJPEG2000ならではの機能可以实现できます。

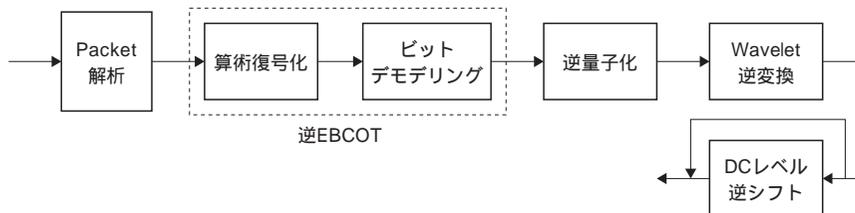


図1.5 JPEG2000デコーダのブロック図

次に、JPEG2000デコーダの典型的な構成例について、図1.5に図示します。図1.5は図1.4のエンコーダの対称形になっていることが容易にわかると思います。なお、エンコーダでDCレベルシフトが未使用の場合には、デコーダでもDCレベル逆シフトも使われません。

Packet解析

JPEG2000の符号化コードストリームをPacket毎に解析して、前記のPacketヘッダとPacket本体に分離します。JPEG2000デコーダは、Packetヘッダに記述された情報を参照しながら、Packet本体に記録されている実際のデータを逆EBCOT部に送出します。

逆量子化

量子化の逆の操作を行います。JPEG2000はデコーダの規格なので、この逆量子化手段が規格書で定義・記述されています。

ウェーブレット逆変換

ウェーブレット変換の逆の操作を行います。

DCレベル・逆シフト

DCレベル・シフトの逆の操作を行います。

1.3 JPEG2000の構成

1.3.1 JPEG2000のPartとは？

JPEG2000には現在(2004年6月時点)でPart-1からPart-12の12個のパートが存在しています(Part-7は欠番)。この中で、Part-1がベースラインのコーデックという位置付けであり、ファイル・フォーマットのファイル拡張子は“.jp2”と決められています。Part-1の標準化は2001年1月の投票で完了しています。Part-2はPart-1の拡張技術と拡張ファイル・フォーマット(.jpx)に関するパートです。この拡張技術としては、SSOWT(Spatial Segmented Overlapped Wavelet Transform：タイル境界をオーバーラップさせながらウェーブレット変換を行う手法で、Part-1の対称拡張方式のウェーブレット変換に比べて、タイルの境界が目立たなくなる長所がある。ただし複雑度が上がるデメリットがある)や、TCQ(Trellis Code Quantization：ピタビ復号で著名なトレリス符号を利用して量子化を行う手段で高画質化が実現できる一方、Part-1で規定されているスカラー量子化に比べて複雑度が大きいデメリットがある)などが、Part-2の技術として存在しています。またPart-2のファイル・フォーマットである.jpxはPart-1の.jp2を大幅に拡張した形式のフォーマットです。

Part-3は動画像符号化規格であるMotion-JPEG2000です。ただし、MPEG規格とは異なり、静止画の連続画像という扱いです。これはそのアルゴリズムにも現れており、MPEGでは必須の技術であ

る動き補償予測(符号化効率を向上させる手法)を, Motion-JPEG2000では採用していません。Motion-JPEG2000のファイル・フォーマットは“.mj2”と定められています。これら以外のPartについての説明は,「1.3.2 JPEG2000の新規格」で説明します。

表1.1 JPEG2000のPart

Part	タイトル	内容
1	JPEG2000 Image Coding System Core Coding System	JPEG2000の中核になる技術を扱う。
2	JPEG2000 Image Coding System Extensions	Part-1の拡張方式に関する技術 (例)TCQ(Trellis Code Quantization, SSWT : Spatially Segmented Wavelet Transformなど)を扱う。JPXも定義。
3	Motion-JPEG2000	Part-1をベースに静止画の連続動画圧縮に関する技術を扱う。
4	Conformance Testing	Part-1で定義される規格内でのビットストリームの検証を行う。
5	Reference Software	Part-1の符号化効率の検証, およびビットストリームの検証を行う目的で公開される公式の参照ソフトウェア。
6	Compound Image File Format	自然画像, 図表画像(文字を含む)などの統計量の異なる複数種類の画像を含んだ画像に特化したファイル・フォーマットを取り扱う。

1.3.2 JPEG2000の新規格(New Part)

JPEG2000の新規格は, Part-8からPart-12までを意味しています(2004年6月時点)。しかし, 標準化の進行にしたがって, 新たなPartが加わる可能性もあります。表1.2は現時点での新規格(すべて規格化中)の内容を示したものです。これらの規格は, 圧縮以外の技術に関する規格であり, 圧縮技術としてはJPEG2000 Part-1を使うことが前提になっています。

表1.2 JPEG2000の新Part

Part	タイトル	内容
8	JPSEC : Secure JPEG2000	セキュリティ用途にJPEG2000を応用するための規格。符号化コードストリームへのアクセスや鍵情報の伝送方法, Watermark, 暗号化技術との融合技術の規格化。
9	JPIP : Interactivity tools, APIs and protocols	JPEG2000のファイルに, クライアントサーバ間でアクセスするなどのツール, API, プロトコルの規格化。
10	JP3D : 3-D and floating point data	JPEG2000で, 3次元データ, 浮動小数点精度のデータを入力とした場合の規格化。
11	JPWL : Wireless	無線環境下でのロバストなエラー耐性などの技術の規格化。
12	ISO Base Media File Format	MP4とMJ2の共通部分の規格化。

1.3.3 互換性のための検証テスト

JPEG2000はデコーダの規格であるため, JPEG2000準拠のデコーダの検証を行う必要があります。そのために, あらかじめJPEG2000エンコーダで生成した符号化コードストリームを用いて, そのデコーダが規格に準拠しているかどうかの検証を行います。JPEG2000のPartの中ではPart-4がその役割をはたしました。

実際には、表1.3に示すように、クラス(Compliance classes)をCclass 0からCclass 2までの三つのクラスに分け、それぞれのクラスごとにパラメータを定義して、各パラメータの条件を満たす符号化コードストリームを参照用のエンコーダで作成しました。表を見てわかるとおり、Cclass 0, Cclass 1, Cclass 2の順に、解像度やその他パラメータの規模が大きくなっています。

表1.3 Part-4で規定されたCompliance classes (Cclass)

パラメータ	Cclass 0	Cclass 1	Cclass 2
W × H (Size)	128 × 128	2,048 × 2,048	16,384 × 16,384
C (Components)	1	4	256
N_{cb}	$(HW/1024+32)C=48$	$(HW/256+128)C=66,048$	$(HW/256+128)C=268,468,224$
N_{comp}	64	256	16,384
L_{body}	8192 bytes	2^{23} bytes	2^{30} bytes
M	11	15	30
P	9-71でエンコードされたデータを5-31でデコードするのに必要な最低量	16ビット固定小数点精度の実装	32ビット浮動小数点精度の実装
B	8	12	16
T_L	3	7	12
L	15	255	65,535
Progressions	CODで定義可能なすべてのprogression。1個のtileにつき先頭のprogressionのみでコードできればよい	解像度レベル数、レイヤ数、コンポーネント数によってのみ制限を受ける	解像度レベル数、レイヤ数、コンポーネント数によってのみ制限を受ける
Tile Parts	1個のtileにつき先頭のtile partのみデコードできればよい	N_{cb} または L_{body} で定義される数値までの、すべてのtile partをデコードすること	N_{cb} または L_{body} で定義される数値までの、すべてのtile partをデコードすること
Precincts	1個のsubbandにつき先頭のprecinctのみでコードできればよい	N_{cb} または L_{body} で定義される数値までの、すべてのprecinctをデコードすること	N_{cb} または L_{body} で定義される数値までの、すべてのprecinctをデコードすること

表1.3で示した各パラメータの定義について、以下に紹介します。

H, W, C: デコード時に必要な値で、水平(W)、垂直(H)はデコード画像の水平、垂直方向の最大サイズを示します。Cはコンポーネント数の最大値を示します。たとえば、Cclass 0のデコーダは、水平・垂直方向に128サンプル以下のサイズの、最大解像度レベルをデコードできる必要があります。さらに3ステージまでのウェーブレット逆変換を行える必要もあります($T_L=3$ が相当)。他方、Cclass 1のデコーダは、水平・垂直方向に2,048サンプル以下のサイズの、最大解像度レベルをデコードできる必要があります。さらに7ステージまでのウェーブレット逆変換を行える必要もあります($T_L=7$ が相当)。

N_{cb} (Code-block parsing guarantee): デコーダが符号化コードストリームを構文解析(parse)する際に、どこまでをparseできるかを示します。この値は換言すれば、parseしたデータを記憶しておく容量を示しています。デコーダは、 $N_{cb}(x) > N_{cb}$ の時点でparseを中止することが許されています。

N_{comp} (Component parsing guarantee) : デコーダのメモリ節約のために, N_{comp} を超える値のコンポーネントに対しては parse を中止することができます .

L_{body} (Coded data buffering guarantee) : デコード開始の前に, デコーダによって蓄積されるパケットのバイト数の上限値を示します . この値に達した時点でデコーダは符号化コードストリームの読みを中止することが許されます .

M (Decoded Bitplane guarantee) : デコード対象のビットプレーン数の最大値 .

P (9-7I Precision guarantee) : 非可逆型 9-7 DWT (離散ウェーブレット変換) を用いて符号化されたコードストリームをデコードする際には, 9-7 逆 DWT を行う必要があります . Cclass 0 の場合には, 5-3I (5-3 の非可逆型逆 DWT) を用いて, デコード画像の精度を検証することになります . これは最低位の compliance class 用に使われる, 最も安価なデコーダ向けの実装になります . より高位の compliance class 用のテストで, 9-7 DWT の代わりに 5-3 DWT を用いると検証テストにパスしません . 非可逆型 9-7 DWT を用いて符号化されたコードストリームを 5-3 逆 DWT でデコードした際には, 信号に依存したノイズが発生します . たとえば, 高域であるエッジ境界にエラーが発生します . それらのエラーは信号に依存しますので, Cclass 0 用の DWT に実装上の精度を要求することはありません . そのかわり, Cclass 0 で用いる参照画像に対して 5-3I 逆 DWT を用いた際の, 9-7 DWT の精度の範囲が設定されています .

B (5-3R Precision guarantee) : デコーダが可逆型 5-3R 逆 DWT を実装する際に, B ビット/サンプル (SIZ マーカセグメントで記述) までのコンポーネントに対しては, 忠実に逆 DWT ができることが必要です . 符号化コードストリームが RCT (Reversible Component Transform) を伴っている場合には, 5-3R 逆 DWT と逆 RCT の両方を, B ビット/サンプルに対して忠実に実現できる必要があります .

T_L (Transform level guarantee) : 逆 DWT を行う際の, 最大解像度レベル数を示しています . T_L を超える解像度レベル数を含むコードストリームの場合には, デコーダは最上位から数えて T_L 個の解像度レベル数のみをデコードすればよいことになります .

L (Layer guarantee) : 符号化コードストリーム中で, デコーダがデコードを保証できる最小レイヤ数を示しています . L を超えるレイヤ数を含むコードストリームの場合には, デコーダは最上位から数えて L レイヤのみをデコードすればよいことになります .

Progressions : すべての Cclass で, デコーダは COD マーカセグメントで定義されたすべての可能な progression をデコードできる必要があります . もしもコードストリーム中で POC マーカセグメントが使われていた場合には, Cclass 0 のデコーダは, POC マーカセグメントで規定された progression の内, 最初の progression のみを, 1 タイルごとにデコードできれば良いです . Cclass 0 以外のクラスでは, ほかの制限 (たとえば前記の N_{cb} や L_{body}) にしたがって, デコード対象のパケットがスキップされます .

Tile parts : Cclass 0 のデコーダでは, 先頭以外のタイル部分は無視します (前記の N_{cb} や L_{body} による制限が達していなくとも) . Cclass 0 以外の Class では, 他の制限 (たとえば前記の N_{cb} や L_{body}) にしたがって, デコード対象のタイル部分がスキップされます .

Precincts : タイルには複数個のプレシント (precinct) を含んでいることがありますが, Cclass 0のデコーダは, 各タイルのサブバンド (subband) 中の, 先頭のプレシントのみをデコードできればよいことになります. Cclass 0以外のClassでは, すべてのプレシントをデコードできる必要があります.

1.3.4 リファレンス・コード

リファレンス・コードは, JPEG2000のアルゴリズム(デコーダと, 参照用としてのエンコーダ)をソフトウェア言語(C, C++, Javaなど)で実現したものです. 標準化の中では, 表1.1で説明したとおり, Part-5として規格化されています.

JPEGの公式ホームページ(www.jpeg.org)では, JPEG2000 code or test data(<http://www.jpeg.org/jpeg2000/testlinks.html>)で, 実際にJPEG2000のソースコードや符号化コードストリームが入手できます. JJ2000, JASPERの二つがPart-5では定義されていますが, これら以外にもKakaduと呼ばれるソフトウェアもよく使われています. Kakaduに関する情報や, 評価用のサンプル・コードは, <http://www.kakadusoftware.com/>から入手することができます. また, このKakaduの開発者であるDavid Taubman氏のホームページ(<http://www.ee.unsw.edu.au/~taubman/>)からは, JPEG2000の技術や書籍などの有益な情報も得ることができます. 同氏が大学の博士課程で開発した技術が, JPEG2000で採用されているEBCOT(Embedded Block Coding with Optimized Truncation)の基礎になっています.