



デジタルカメラと画像圧縮

石川 安 則

1. は じ め に

昨今におけるデジタルカメラ(以下DSCという)隆盛の要因の1つに、静止画像圧縮符号化方式の標準となっているJPEG規格の認知と普及があげられる¹⁾。大容量のフラッシュメモリが安価に供給されるようになり、DSCシステムの急激な価格低下を促したが、これにはJPEG規格による画像圧縮を利用して高画質で実用的な撮像枚数を確保できるようになったことも貢献している。JPEG規格は、再生画像品質とのトレードオフにより容易に圧縮率の制御が可能で、ユーザが希望する画質で撮像枚数を決定できる。また、JPEG規格を利用するファイルフォーマットの標準化(Exifなど)の意味も大きい。これは基本的にはJPEG規格のビットストリームをそのまま利用している。さらには、JPEG専用チップの普及、あるいはRISC CPUの高速化によるソフトウェア処理と簡単なハードウェアにより圧縮処理が可能となり、普及価格帯の製品開発に拍車がかかったといえる。

本稿では、まずDSCにおけるJPEG規格の利用形態を概観し、その問題点についてまとめた後、現在新たな静止画圧縮符号化国際標準として審議中のJPEG 2000規格の動向を解説する。さらにJPEG 2000技術の中核をなす、Wavelet変換方式について詳述する。

2. DSCにおける画像圧縮の現状と問題点

現在、DSCで用いられている画像圧縮符号化方式は、ほぼすべてJPEG規格である。JPEG規格の詳細については、多くの文献で詳細に説明されているので、ここでは概要と画質について簡単に触れるにとどめる。

JPEG (Joint Photographic Expert Group)とは、2つの国際標準化機関ISOとITU-T (旧CCITT)との共同検討グループの名称であり、ここで制定されたカラー静止画符号化方式標準を一般にJPEG方式と呼んでいる。図1にJPEG方式の基本ブロック図を示す。DCT (離散コサイン変換: Discrete Cosine Transform)は、空間領域から周波数領域への直交変換であり、JPEGでは 8×8 画素ブロック毎のDCTにより、 8×8 の周波数成分(DCT係数マトリクス)に変換される。一般に自然画像の直交変換により、周波数分布が低域へ集中する指数関数特性を示すが、DCTはその中でも低域集中度が優れているとされる。次段の量子化特性を低域重視型とし、符号ビット配分を低域に集中させることにより、圧縮効率を高めることができる。エントロピー符号化は、量子化後のビット列の統計量に基づく符号割当てを最適化する手法の1つである。エントロピー符号は、一般に可変長符号であり、同一の量子化特性を

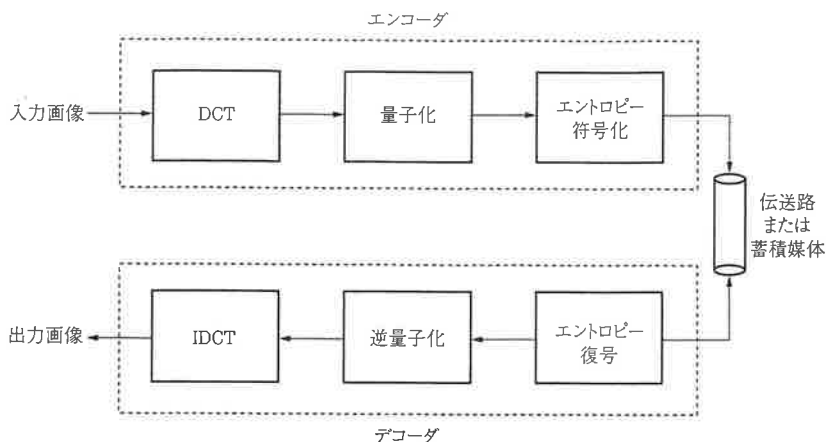


図1 JPEG方式の基本ブロック図

用いても圧縮対象画像により符号量は変化する。JPEG方式では、エントロピー符号に符号化損失を伴わない可逆符号化を採用しており、再生画質とのトレードオフによる1画像全体の符号量制御は、量子化特性の制御により行われる。

DSCにおいて、撮影品質モード毎にメモリへの記録枚数を保証するためには、上記の量子化特性を制御して固定長の符号量を得る必要がある。固定長符号量制御方式は、種々の方式が提案・利用されているが、いずれもJPEG規格の枠内で行うための制約がある²⁾。

表1は、DSCの代表的な圧縮仕様を示すものである。撮像素子であるCCD (Charge Coupled Device)の画素数が100万画素を超えるいわゆるメガピクセル機では、データ出力画素サイズは、1280×960画素程度(SXGA)が主に用いられるが、CCDが30万画素程度の普及タイプで主流の640×480画素程度(VGA)の画素サイズのモードも使用される。ほとんどが3段階の撮影品質モードを有し、圧縮率は1/3～1/24程度である。4MBのメモリカードを用いて、最高品質で3～6枚、高品質で10～15枚、普通品質で20～40枚程度の記録が可能なのが多い。

写真1～3は、JPEG方式による圧縮再生画像品質を比較したものである。写真1はオリジナル(非圧縮、24bit/画素)、写真2と3は圧縮率約1/24および1/48の再生画像である。図示

はしていないが、1/6程度の圧縮では歪みはほとんど検知できず、1/12程度でも背景部がわずかにザラつく程度だが、1/24程度になると背景部と滑らかな果物の表面にブロック状の歪みが見えてくる。また、1/48程度まで圧縮するとブロック歪みが大きく発生し、品質は非常に低下する。ここに掲載した再生画像は、同一の条件下で圧縮率による画像品質の比較のために、高品質の画像データを用いてJPEG圧縮を行ったものであり、実際のDSC製品においては、光学系のローパスフィルタやフォーカス精度などによるMTF特性の低下や、撮影条件が必ずしも最適でないことなどから、オリジナル画像品質が低い場合が多い。さらに通常は圧縮処理の前後の画像処理において圧縮率ごとの補正を行って画質補償をしており一般にJPEG圧縮による歪みは、製品レベルでは検知しにくい画質設計がされている。

ここで、DSC製品の観点から見たJPEG方式の問題点をまとめると以下になる。

- (1)圧縮率が高くなると、ブロック歪みやモスキート歪みなど視覚的に妨害となる人工的歪み(artifact)が大きくなる。
- (2)固定長圧縮のために特別なアルゴリズムが必要である。
- (3)DCT処理は、基本的に浮動小数点処理が必要。その他に、DSC以外の観点からの問題点として、

表1 DSCの圧縮仕様

	メーカ	型番	使用 CCD	圧縮仕様				
				モード	枚数*1	FileSize/ 1枚*2	画素サイズ	圧縮率*3
メ ガ ピ ク セ ル 機	リコー	DC-4	1/2.7inch 133万画素 (有効122万)	Fine	6	683KB	1,280×960 (SXGA)	1/5.3
				Normal	12	341KB		1/10.6
				Economy	23	178KB		1/20.2
	ニコン	COOLPIX 900	1/2.7inch 130万画素 (有効125万)	FINE	6	683KB	1,280×960 (SXGA)	1/5.3
				NORMAL	12	341KB		1/10.6
				BASIC	24	171KB		1/21
	エプソン	CP-600	1/2.7inch 130万画素 (有効 不明)	★★★	6	683KB	1,280×960 (SXGA)	1/5.3
				★★	13	315KB		1/10.6
				★(VGA)	43	95.3KB	640×480	1/9.4
	オリンパス	C-840L	1/2.7inch 131万画素 (有効125万)	SHQ	18 以上	455KB	1,280×960 (SXGA)	1/7.9
				HQ	36 以上	228KB		1/15.6
				N (VGA)	122以上	67.1KB	640×480	1/13.4
	カシオ	QV-7000SX	1/3inch 132万画素 (有効125万)	S	14	585KB	1,280×960 (SXGA)	1/6.1
				F	19	431KB		1/9.4
				N	33	248KB		1/14.5
	富士フイルム	FinePix 500	1/2inch 150万画素 (有効131万)	Fine	2	1,024KB	1,280×1,024	1/3.8
				Normal	5	410KB		1/9.4
				Basic	11	187KB		1/20.5
	オリンパス	C-1400L	1/2inch 141万画素 (有効131万)	SHQ	4	1,024KB	1,280×1,024	1/3.8
				HQ	12	341KB		1/11.2
				SQ (VGA)	49	83.6KB	640×480	1/10.7
スーパーメガ ピクセル機 (参考仕様)	リコー	RDC-500	1/2inch 230万画素 (有効219万)	Fine	9	910KB	1,792×1,200	1/6.9
				Normal	19	431KB		1/14.6
				Economy	38	215KB		1/29.3

* 1: 使用メディア容量は、8MB (C-840L, QV-7000X, RDC-500), 2MB (FinePix 500), 4MB (その他)

* 2: メディアのヘッダも含む * 3: 画素サイズ/(FileSize/1枚)

(4)可逆圧縮(Lossless(無歪み)圧縮, =画質的には非圧縮)が全く別の方式である。

(5)Lossy(有歪み)圧縮から Lossless 圧縮へ連続的に階層処理が困難。

(6)階調数が8ビット(基本モード), ないし 12ビット(拡張モード)に固定され汎用性がない。

(7)付加情報の柔軟性(これはDSCでも, 日付・場所・撮影条件などの付加に必要)。

等があげられる。これらの問題点を解消すべく, 現在標準化が進められている新しいカラー静止画符号化方式がJPEG 2000 と呼ばれる方式である。これについて次章で詳しく述べる。

3. JPEG 2000 と Wavelet変換符号化方式

3.1 JPEG 2000 の要求仕様と技術課題

現 JPEG 規格制定後, その後のあたらない要求を満たすべく 1997 年 8 月に ISO の新規課題(JTC 1.29.14; ISO/IEC 15444)として承認され, その名の通り 2000 年の IS 化を目指して検討が開始された。

JPEG 2000 の要求仕様は, 前項に述べた諸機能のほかに多くの項目があり, 多様な要求項目を満たす仕様となっている。また期待されるアプリケーション毎に, 必須機能・オプション機



写真1 オリジナル画像(非圧縮, RGB24bit/画素)
画素サイズ: 1280 × 1024



写真3 JPEG圧縮画像(圧縮率: 約 1/48)



写真2 JPEG圧縮画像(圧縮率: 約 1/24)

能の整理が行われ、これらがまとめられて“Requirements & Profiles”というISO文書が作成されている(現在Version 4.0)。将来的にはこれを基に各種サービス毎のプロファイルが作成

される。DSCに特有の要求機能として、“Digital Photography”の項目が設けられ、定サイズ性とサイド情報の埋め込み、符号化のリアルタイム性が必須機能とされている。また低ワークメモリの要求などEncoder complexityの評価が必須項目とされている。

上記のような多くの要求をすべて統一的に評価する事は困難であるため、第1次評価として、まず低ビットレートでの圧縮性能を重視する事が合意され、1画素あたり1/16, 1/8, 1/4, 1/2, 1, 2 bitでの画質(主観評価とSNR)、およびLossless圧縮におけるビットレートの評価が行われた。最終的に24件ものアルゴリズム提案があったが、97年11月会合(シドニー)での評価の結果、米国アリゾナ大(UA)/SAIC (Science Applications International Corporation) 共同提案のWTCQ (Wavelet Trellis Coded Quantization) 方式が最も高い評価を受け、参照モデル方式として位置づけられた。以降の会合での審議において、この参照モデルをベースに各機能ブロックの性能改善や簡略化の提案が行われ、現在これら提案技術をVM (Verification Model) と呼ばれ

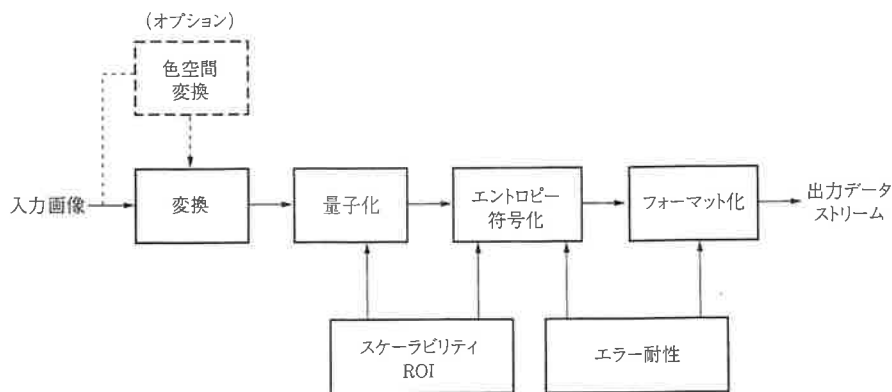


図2 JPEG2000のブロック構成(エンコーダ)

る検証モデルへの組み込み・評価を行って方式の絞り込みを行っている。

JPEG 2000の機能ブロックを図2に示す。各機能ブロックにおける技術課題は、以下のようになまとめられる³⁾。

(1) 変換方式

Wavelet変換とDCTの提案がある。DCTは、現JPEGとの互換性を重視する立場からは必然の帰結であるが、圧縮性能、Lossy/Lossless統一符号化方式への適合性、解像度Scalabilityの実現手段などの点でWaveletが採用される方向である。現在は、Wavelet変換フィルタの選定が主要課題である。

(2) 量子化

WTCQ方式で提案されたTCQ (Trellis Coded Quantization)は、統計量から推定される量子化ステップ値に対して、与えられたビットレート下で量子化歪みを最小(MSE)とするように、4つの量子化器をトレリス線図に基づき時系列的に最適な量子化器を切り替える適応的な方法であり、ベクトル量子化器を用いている。特に低ビットレートでのレート・歪み特性が優れているが、アルゴリズムの複雑さが性能に見合うものかは議論がある。ビットレートにより、スカラー量子化器と切り替えるなどの選択枝も有る。

(3) エントロピー符号化

エントロピー符号化には、ブロック符号(ハフマン符号)と非ブロック符号(算術符号)の提案が

ある。原理的な符号効率には算術符号が優れているが、処理速度の点ではハフマン符号が有利であるとされる。また、エントロピー符号化には、コンテキストモデル(出現確率予測モデルやブロック符号化における符号化順序制御)も含まれるため、評価を複雑にしている。全体としては、実装時のコストを重視して符号化順序制御を有しないブロック符号の要望が強いが、今後の評価結果により結論の予測は現状では難しい。なお、JPEG 2000標準のポリシーはパテントフリーとしているが、エントロピー符号化に関しては依然として問題を抱えており、技術的評価に少なからず影響がある。

(4) ROI (Region of Interest)

JPEG 2000の機能性を特徴づける仕様の1つが、ROIである。1枚の画像のうち、ある特定領域を優先的に高精度とした符号化・復号を行う機能や、ランダムに特定領域にアクセスして復号できる機能を言う。当初符号化効率の点から、64×64程度のタイリング毎にROIを設定していたが、ランダムアクセス性や実装におけるメモリ量の問題が提起され、32×32、ないしもっと小さい領域毎や、ストライプ・ライン単位のROIを設定する提案も行われている。

(5) スケーラビリティ

スケーラビリティの機能は、現JPEGの拡張機能としてある程度実現されている。すなわち、プログレッシブ再生機能としては、DCT係数の

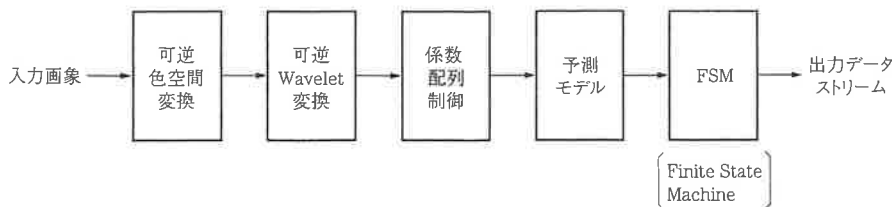


図3 Wavelet変換符号化方式のブロック図(CREW方式)

周波数方向または階調方向の順次伝送により実現され、またハイパーキカル符号化では入力画像の空間解像度を、水平・垂直方向に1/2ずつ段階的にサブサンプルしていくことにより、異なる解像能力を有する端末間での伝送を容易とする機能である。現JPEG方式では、それぞれ異なるデータ構造を持たせてこれらの機能を実現しているが、JPEG 2000においては統一的なデータ構造を用いて、空間解像度と周波数または階調方向の解像度を独立にスケラブルとする方式が提案されている。特にWavelet変換は、前述のように本質的に解像度スケラブルな構造を持っており、実現を容易としている。

(6) その他

誤り耐性(Error Resilience)機能は、移動体通信のプロファイルでは必須機能と位置づけられている。現在のところ、JPEGのRST(Restart)マークを拡張した再同期機能に絞った議論が行われている。

符号化効率向上のための色空間変換に関しては、JPEGではカラー変換は標準化範囲外とされ、現在ISO/IECの別の課題として検討が行われている(マルチコンポーネント符号化, WD 14493)。JPEG 2000では、これを標準内で規定することになっているが(オプションとなる予定)、現状では議論はあまり活発ではない。しかし、端末間でのカラーマネジメント・色再現範囲の確保など、符号化における色空間変換の問題は、今後専門家を含めた議論が望まれる。

3.2 Wavelet変換符号化

前項において、JPEG 2000の評価で、DCT符号化に対してWavelet変換を利用した符号化の方

が特に低ビットレートの画質が優れることが述べられた。ここでは、JPEG 2000の要求仕様に良く合致する機能を有するWavelet変換符号化方式の1つについてアルゴリズムを説明する⁵⁾。全体ブロック図を図3に示す。

(1) 可逆色空間変換

JPEG標準の規定としては色空間変換は標準範囲外であるが、一般に普及しているJPEG符号化ツールでは、RGB空間のまま符号化すると圧縮効率が上がらないため、YCbCr空間などへの変換後に符号化する。YCbCr変換としては、ITU-R勧告BT. 601ないしBT. 709の規格を利用する人が多い。変換に際しては、一般に有限語長演算に起因する系統的な演算誤差が発生するが、JPEGにおけるDCT演算自体が系統的誤差を有するため、いずれにしても符号化・復号を繰り返すと誤差が蓄積する。しかし、以下に示す色空間変換は系統的誤差は発生しない。

$$\begin{aligned} Y_r &= \left\lfloor \frac{R + 2G + B}{4} \right\rfloor \\ U_r &= R - G + 2^{\text{depth}} - 1 \\ V_r &= B - G + 2^{\text{depth}} - 1 \end{aligned} \quad (1)$$

ただしここでdepthは、入力画素精度(ビット数)をあらわし、8ビットなら数値レンジは、符号なしで0～255、符号つきで-128～+127となる。演算記号 $\lfloor x \rfloor$ は、切り捨て(x に等しいかそれ以下の整数)を表す。上記演算は、すべてシフト演算のみで行えることに注意すべきである。

(1)式の結果、 Y_r は入力精度のレンジを保持し、 U_r 、 V_r は入力精度+1ビットのレンジとなる。この変換係数は、ITU-R BT. 601の変換を近似

しており、符号化効率の低下は概ね 0.1 dB 以下とされる⁶⁾。

(2) 可逆 Wavelet 変換

Wavelet 変換は、タイル(符号化の単位ブロック)毎に、以下の Wavelet 変換フィルタを用いて行う。フィルタは、2 タップのローパスフィルタと 10 タップのハイパスフィルタ(いずれも 1/2 ダウンサンプル)の組合わせである。 $s(n)$ はローパスフィルタ出力、 $d(n)$ はハイパスフィルタ出力である。最初に水平方向に $s(n)$ と $d(n)$ を求め、それぞれに対して垂直方向に同じフィルタをかける。

$$s(n) = \left\lfloor \frac{x(2n) + x(2n+1)}{2} \right\rfloor$$

$$d(n) = x(2n) - x(2n+1) + p(n)$$

$$p(n) = \left\lfloor \frac{3s(n-2) - 22s(n-1) + 22s(n+1) - 3(n+2) + 32}{64} \right\rfloor \quad (2)$$

逆変換は次のようになる。

$$x(2n) = s(n) + \left\lfloor \frac{d(n) - p(n) + 1}{2} \right\rfloor$$

$$x(2n+1) = s(n) - \left\lfloor \frac{d(n) - p(n)}{2} \right\rfloor \quad (3)$$

$s(n)$ は、入力画素値精度(レンジ)に対して出力精度は等しいが、 $d(n)$ は 1 ビット増える。したがって、水平・垂直ともローパス成分(LL)は入力レンジに等しく、水平ローパス・垂直ハイパス(LH) および水平ハイパス・垂直ローパス(HL)成分は 1 ビット増え、水平・垂直ハイパス(HH)成分は 2 ビット増やすことにより、可逆変換精度が確保できる。このように、変換を整数演算のみによる実装を行えば系統的演算誤差は発生せず、非可逆圧縮における誤差はすべて量子化制御に委ねられる。LL成分はさらに、解像度が 4×4 程度になるまで階層的に繰り返し同じ変換を行う。図 4 は、4 階層の変換例を示す。

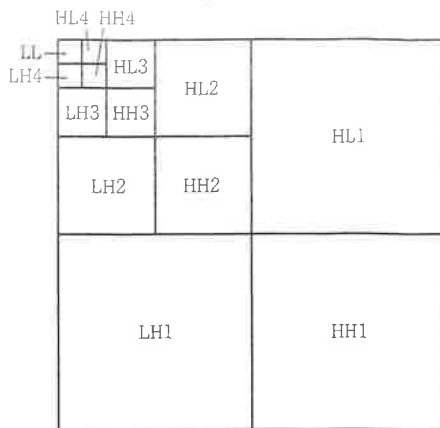


図 4 4 階層の Wavelet 変換

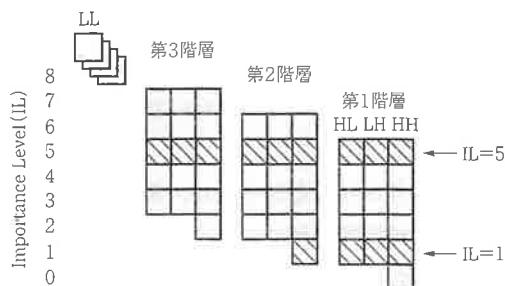


図 5 係数列列(1) ~ Normal Alignment ~

(3) 係数列列と量子化

Wavelet 変換係数値は、まず bit-significance form という形式に変換される。例えば、“17”を 8 ビットで表現すると、次のようになる。

$$\begin{array}{cc} \text{head} & \text{tail} \\ 001 & 0111 \\ \hline 0 & \\ \text{sign} & \end{array} \quad (4)$$

すなわち、符号ビット(sign)+振幅値(magnitude)の形式とし、MSB (Most significant bit) から見て最初の“1”までを head bits とする。次に sign bit, 最後に head bits 以下のビット(tail bits)の順に並び替える。ただし、値が“0”の時は sign bit は省略される。

次に、図 4 に示すそれぞれの階層の周波数バンド毎に、各々のビットプレーンを MSB から LSB (Least significant bit) の順に並べる。図 5 は典型的な配列の例を示す。入力画素精度が 4 ビット

トの例である。また、Importance level (IL)は、量子化の優先度を表すインデクスであり、ILが小さくなるに従い優先的に量子化が行われる。LL成分は、4ビットの階層すべてが同じILを有する。次のILは、第1階層のMSBのビットプレーンについてHL, LH, HHの周波数バンドの順に並べる。IL = 2では、第1階層のHL, LH, HHの次のビットプレーン、および第2階層のHL, LH, HHのMSBのビットプレーンを並べる。なお、それぞれの周波数バンド内の係数ビットの走査順序は、水平方向を主とし垂直方向が副走査である。

図5の配列では、量子化の優先度は、階層の低い方かつLSB側から行われるので、全体として解像度・階調数同時に量子化制御ができる。一方、図6に示す配列では、低い階層全体から量子化が行われるので、解像度を優先した量子

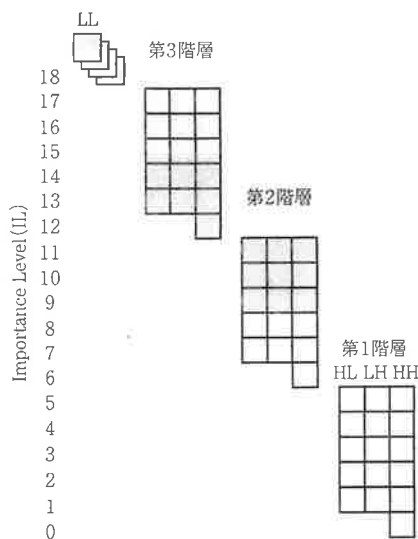


図6 係数配列(2) ~Pyramidal Alignment~

化制御となる。また、符号化時に任意のILで符号化を打ち切る事ができるので、タイル単位などで符号化順序を制御すれば、定レート・定サイズの符号化が可能となる。さらに、任意のILで復号の打ち切りができるので、復号側でも量子化制御が可能である。

(4) エントロピー符号化とコンテキストモデル
コンテキストモデルは、以下のような予測アルゴリズムに従う。

- 各周波数バンド毎のビットプレーンの全係数値が“0”となるかの予測。
- (a)でなければ、各周波数バンドの親ビットプレーン(図7参照)を含め、16ビット毎に“0”が続くか(ゼロラン)の予測。
- 次に、head bits, sign bit, tail bits毎に、各ビットプレーンの周囲のビットおよび親ビットのビットからの予測。

実際のエントロピー符号化の実行は、図8に示すFSM (Finite state machine)によって行われる。コンテキストの予測確率(Context probability)は、

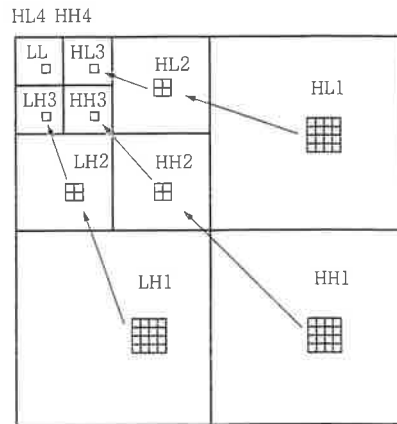


図7 係数の親子関係

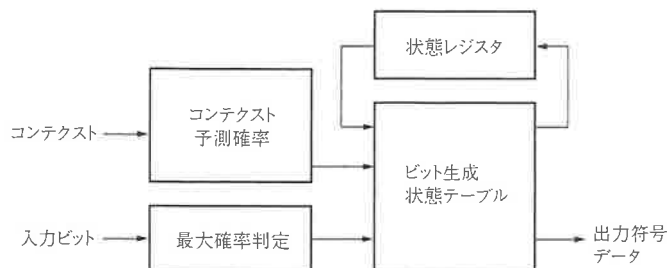


図8 FSM (Finite State Machine) のブロック図

上記のコンテキスト毎に異なり、また1つのビットが符号化されるごとに更新される。ビット生成のLUT (Lookup table)は、入力画像や、コンテキストに因らず一つの固定テーブルが用いられる。

写真4に Wavelet 変換方式によるサンプル画像を示す。圧縮率約 1/48 のサンプルであるが、JPEG 方式(写真3)と比較してブロック歪みがなくなり、滑らかな再生画像が得られている。

3.3 DSC用PCアプリケーションと画像圧縮

DSCの利便性を特徴づける機能の1つに、PC (Personal computer)との容易な接続性があげられる。DSCの普及は、インターネット上のWWWの爆発的な広がりとはほぼ同期しているが、ユーザが簡単に画像をホームページに貼り付けて情報発信できることが、DSCの購買意欲を促したと言える。また、いわゆる電子アルバム機能を有するアプリケーションが豊富に供給され、銀塩写真では手間のかかった画像の編集・整理といったホビータ的な作業が家庭のデスクトップ上

で容易に実現できるようになった。図9は、電子アルバム編集機能の1つの例で、サムネイル(見出し)画像の作成例を示す。圧縮された画像



写真4 Wavelet変換符号化による画像(圧縮率: 約 1/48)



図9 電子アルバムアプリケーションのサムネイル画像作成例(リコー製DIGICLIP-2使用)

データから、縮小した画像を多数生成する必要があるので、通常処理時間がかかるが、前項で説明した Wavelet 変換符号化方式を用いると JPEG 方式と比較して約 2 倍の処理速度が得られる。

4. ま と め

DSC における画像圧縮の現状と、新しい静止画像圧縮方式について解説した。現 JPEG 方式はすでに広く普及しており、成熟した技術を用いた安定した方式として今後も種々の分野で使われていくと思われる。一方 JPEG 2000 は、現 JPEG 規格制定以降の画像圧縮分野の研究成果を集積し、豊富な機能と高能率な圧縮性能を提供することを目標としているが、JPEG 2000 の基本コンセプトとして現 JPEG との共存が謳われており、アプリケーションによる棲み分けが行われると考えられる。また新しいマルチメディア符号化標準である MPEG-4 と共通インターフェイスの検討も行われており、次世代のマルチメディアシステムにおける中枢技術の一翼を担う

技術になるであろう。マルチメディア端末の 1 形態として DSC を見た場合に、MPEG-4 と JPEG 2000 の標準化に向けた動向には今後とも注意していく必要がある。

参 考 文 献

- 1) 柚木; 画像電子年報: デジタルスチルカメラ, 画像電子誌 27 [6] (1998)
- 2) 石川; 画像データ圧縮技術と画質, O plus E [214] (1997)
- 3) 小野, 上野, 渥美; JPEG2000 の動向, 映像メディア誌 52 [12] (1998)
- 4) JPEG2000 Requirements and profiles ver. 4.0, ISO/IEC JTC1/SC29/WG1 N1105R (1998)
- 5) M. Boliek, et al.; Next Generation Image Compression and Manipulation Using CREW, Proc. ICIP (1997)
- 6) M. J. Gormish, et al.; Lossless and nearly lossless compression for high quality images, Proc. IS&T/ SPIE 9th Annual Symposium (1997)

マシンビジョン用語

マシンビジョン用語委員会

マシンビジョン用語

マシンビジョン用語委員会

A 5 判 81 頁 定価 1,522 円 (消費税込み)

送料 実 費

JCEM

JCEM

社団法人 日本オプトメカトロニクス協会

〒105-0011 東京都港区芝公園 3 丁目 5 番 22 号 機械振興会館別館内

TEL (03) 3435-9321 代 FAX (03) 3435-9567

E-mail: joem@sepia.ocn.ne.jp

JAPAN OPTOELECTRO-MECHANICS ASSOCIATION 3-5-22 Shibakoen, Minato-ku, Tokyo 105-0011, Japan