2. 画 像 情 報 論

2.1~2.4 谷 本 正 幸[†](正会員)

2.1 ま え が き

画像情報はマルチメディアの中核である。画像情報圧縮は、情報量の非常に多い画像の伝送や蓄積を適正なコストで表現するための重要な基盤技術である。

現在、DCT(離散コサイン変換)、動き補償予測などの情報圧縮技術を駆使する H.261, MPEG 1, MPEG 2 などの国際標準化方式¹⁾が定められ、これらを用いてテレビ電話、テレビ会議、ビデオ CD、多チャンネルテレビ放送などのデジタル映像サービスが実現される等、画像情報圧縮技術は目ざましい成果を上げている。本文では通信、放送、蓄積メディアの幅広い分野での利用が始まった画像情報圧縮に焦点を絞り、そのアルゴリズムと、それを実行する画像用 LSI の近年の進展について述べることにする。

2.2 圧縮アルゴリズム

2.2.1 圧縮性能

過去約 20 年間に発表された動画像符号化装置の符号化レートの例を図 1 に示す。同じ品質で厳密に比較することは難しいが、放送、CATV, CCTV 用やテレビ電話、テレビ会議用のレートの下限に注目すれば、この 20 年間に 1 桁程度以上の低下が実現されている。

これらの装置で用いられる情報圧縮方式である動き補償予測や DCT などの波形符号化は、いずれも多数のデータをひとまとめにし、その重複部分を削除することによって圧縮を行う。このため、フレームメモリを用いた画像データの蓄積と、対象物の動きに合わせた統合(動き補償)が、圧縮性能の向上に大変効果的である。

一方、画像のデータ量は空間解像度、時間解像度、

図 1 動画像符号化装置の符号化レート

振幅分解能の積になる。このため、ひとまとめにするデータが同じであっても、圧縮による劣化をこれら 3 つにどのように分配するかによって画質に差が生じる。符号化パラメータの値の選び方や適応制御に関するノウハウの蓄積も、性能向上に大きく寄与する。

今日、波形符号化は成熟した技術であると考えられているが、動き補償の精度向上、動き補償予測誤差信号に整合した符号化などにより、一層の性能向上の可能性が残されている。

2.2.2 機 能

近年、デジタル HDTV 放送、インタラクティブテレビ、パッケージメディアなどにおいて画像情報圧縮の実用化が現実となるに従い、高い圧縮性能だけでなく、用途に応じた種々の機能の実現が望まれるようになってきている。例えば、現行テレビと HDTV のような古いシステムと新しいシステムが共存するときに

[†] 名古屋大学 工学部

"2. Fundamental of Image Techniques" by Masayuki Tanimoto (Faculty of Engineering, Nagoya University, Nagoya)

は、互換性が保たなければならない。古いシステムの符号器出力を新しいシステムの復号器で復元できるフォワード互換や、新しいシステムの符号器出力を古いシステムの復号器で復元できるバックワード互換などの機能が必要となる。

画像情報を何らかの基準に基づいて階層化すると、利用者が解像度等の品質を選択できるようになる。このような品質を選択できる機能をスケーラビリティ(品質可変性)という。階層化されたデータの一部を復号化すれば、必要な品質の画像を再生することができる。スケーラビリティの例としては、空間解像度、時間解像度、 S/N 、複雑さなどがある。

この他、パッケージメディアでは、ランダムアクセス、早送り再生、スロー再生、巻き戻し再生、ポーズなどの機能が要求される。このような機能を実現するため、MPEG では前後の画像と無関係に符号化する I-picture (Intra-coded Picture) が周期的に挿入される。ATM 通信では、階層化されたデータに優先順位をつけ、重要なデータにはセルロスに対して強い耐性を与える。地上系のデジタルテレビ放送では、重要な情報ビットの送信電力を高め、放送局から距離が離れるのに従って品質がゆっくり低下する Graceful Degradation を実現する。

2.2.3 原理

圧縮アルゴリズムの方式原理においても重要な進展が見られた。DCT は直交変換の中で最も優れた電力集中度を持つ KL 変換によく似た基底を持つ。しかし、DCT による周波数分析は遮断特性が緩やかで、阻止域減衰量も充分でない。また、基底の長さがすべてそろっていて次のブロックの基底との重なり合いもないため、低レートではブロック歪みが発生する。このため、LOT (Lapped Orthogonal Transform)²⁾、ウェーブレット変換³⁾等の新しい直交変換による符号化が考え出された。

通常の直交変換では N 個のデータから N 個の変換係数を得るが、LOT では N 個ずつ重ね合わせて取る $2N$ 個のデータを線形変換して N 個の変換係数を得る。逆変換を行うと N 個の変換係数から $2N$ 個のデータが得られる。これを N 個ずつ重ね合わせて加算すると、元のデータが復元できる。LOT では基底の両端が滑らかに変化し、しかも互いに重なり合うため、ブロック歪みの発生が抑えられる。

ウェーブレット変換は、基本ウェーブレットと呼ばれる基本波形を相似変換および平行移動したものを基底とする直交展開である。平行移動した基底は互いに重なり合う。基底の長さは DCT や LOT とは異なり、

高周波成分の基底は短く、低周波成分の基底は長い。このためウェーブレット変換では、ブロック歪みだけでなく、物体のエッジ部近傍に発生するモスキート雑音を抑圧できる。

画像情報圧縮は本質的には画像の表現手段である。信号を解像度の異なる空間に分割して表現することを多重解像度表現 (Multiresolutional Representation) という。画像の多重解像度表現の代表はピラミッド⁴⁾である。画像を低域通過フィルタリングしてサブサンプリングし、縮小画像を得る。これを繰返して得られるピラミッドは、全サンプル数が原画像のサンプル数より必ず増加するので、画像表現法としては冗長である。

このような従来のピラミッドの問題点であった冗長性は、相補ピラミッド⁵⁾によって解決された。相補ピラミッドは、各階層のサブサンプリング点の集合が元の画像のサンプリング点と一致する相補サブサンプリングによって作られるピラミッドである。

前述のウェーブレット変換も、冗長性のない画像の多重解像度表現のひとつである。従来のピラミッドではサンプル数が増加するが、フィルタに関する制限は緩やかである。これに対して、ウェーブレット変換や相補ピラミッドではサンプル数は増加しないが、フィルタには制約条件がつく。

新しいタイプの直交変換である LOT、ウェーブレット変換や相補ピラミッドは、いずれも古くから知られているサブバンド方式の一種とみなすことができる。図 2 にサブバンド関連方式の相互関係を示す。LOT は直交性を保ったまま基底の長さを直交変換の 2 倍に拡大したものである。相補ピラミッドは、ピラミッドの持つ冗長性を除去して得られる多重解像度表現である。ウェーブレット変換は、直交変換の持つ直交性とピラミッドの持つ多重解像度表現の性質を合わせ持つものであり、冗長性も除去されている。LOT、ウェーブレット変換、相補ピラミッドは、いずれもサブバンド方式から導くことができる。

サブバンド方式に関して、符号化利得の一般式が導かれ、これを用いていろいろなサブバンドフィルタの比較が行われた⁶⁾。また、符号化利得を最大にするサブバンド方式が明らかにされた⁷⁾。

ウェーブレット変換や LOT を用いた符号化方式によっても、QMF や CQF を用いたサブバンド方式に比べて圧縮性能の大幅な向上は実現されていない。ウェーブレット変換の持つ連続微分可能な性質や正則性などの条件も、現在のところ画像符号化では大きな利点とはなっていない。

図2 サブバンド関連方式の相互関係

自己相似性を利用するフラクタル符号化⁸⁾において、自然画像の圧縮に注目すべき進展があった。この方式は、送信側では画像を重ねる合わないレンジセルに分割し、レンジセルが同一画像内の別の大きなドメインセルの縮小画像でよく近似できるようなドメインセルと縮小画像をレンジセル毎に見つけ出す。ドメインセルの位置と画像を決定するパラメータが、圧縮された画像情報となる。受信側では任意の初期画像に対して、この情報に従ってレンジセルをドメインセルの縮小画像で置き換え、画像を変換する。こうして得られた画像に対して再び同様の変換を行い、これを繰り返すことにより極限画像に収束させ、再生画像を得る。この方法を基にして、JPEG アルゴリズムと同等程度以上の符号化性能が実現された。

波形を符号化する従来の符号化とは異なる、新しい原理による知的符号化の研究が行われた⁹⁾。対象物の構造・形状モデルを利用する符号化をモデルベース符号化という。これは送信側と受信側に対象物に関する共通の形状モデルを持ち、送信側では対象画像を分析して形状モデルの変形情報を抽出する。受信側では変形情報によって形状モデルを変形して画像を合成・生成する。対象画像は顔画像などに限定されていることが多い。モデルの精度を向上すれば、よりリアルな合成画像が得られる。人物像に関しては、頭部全体の動作、眼の動作、身振り、手振り、指動作、表情、人の集団の動作、人物などの認識に関する研究が進んだ。

2.2.4 対象

画像情報圧縮の対象が、超高精細画像¹⁰⁾や3次元画像などへ新しい広がりを見せた。4000×4000画素以上のUDTV(Ultra Definition TV)では、1枚の絵を視距離フリーで見ようとする用途が考えられるため、劣化が検知されないことを前提とする情報圧縮が必要

になろう。

3次元画像には、2眼式や多眼式のレンチキュラレンズを用いるものや、ホログラムを用いるものなどがある。3次元物体の情報量は非常に多い。2次元画像の標本点は100万点程度であるが、3次元物体をホログラムで表すときの標本点数はおよそ100億点にもなる¹¹⁾。2眼式や多眼式の3次元画像の情報圧縮¹²⁾として、視差方向の冗長性を利用する方式や撮像系の幾何学を利用する方式が報告されている。ホログラフィ方式についても圧縮の研究が行われている¹³⁾。

2.2.5 標準化

1991年以降、画像情報圧縮の分野においてもH.261, JPEG, MPEG 1, MPEG 2などの国際標準化方式が次々に定められてきた。1993年には、放送、通信、パッケージメディアのあらゆる分野に適用できる汎用ビデオ符号化標準であるMPEG 2¹⁴⁾の草案が完成した。以下ではMPEG 2の主要点を紹介する。

MPEG 1はノンインタレース画像を符号化の対象としていた。MPEG 2ではインタレース画像を主対象とし、その符号化効率を向上させる工夫がなされている。さらに、スケーラビリティの概念が導入されている。基本となる画像フォーマットは、CCIR 601およびHDTVである。

幅広い分野に適用できるように、符号化に必要なパラメータは変数となっている。パラメータには、画像フォーマットに関するパラメータと符号化モードに関するパラメータがある。これらのパラメータはプロファイルと呼ばれる機能セット毎にまとめられている。MPEG 2の用途は非常に広いため、満たさなければならない要求条件が多い。この要求条件を満たす技術をつールという。しかし、すべてのツールを包含した装置を作ることは当面現実的でない。このため、限ら

れた数の機能を持つサブセットがプロファイルとして定義されている。プロファイルをパラメータ値の大きさに区切ったものをレベルという。

プロファイルには、片方向予測のみを用いるシンプルプロファイル、両方向予測も用いるメインプロファイル、SNR スケーラビリティの機能を持つプロファイル、空間スケラビリティの機能を持つプロファイル、すべてのツールを含むハイプロファイルがある。レベルについては、高いレベルほど解像度が高い。CCIR 601 画像を両方向予測も用いて符号化する部分はメインプロファイル・メインレベルと呼ばれる。その仕様を表 1 に示す。

画像データは MPEG 1 と同様に、上位から順に Sequence, GOP(Group of Pictures), Picture, Slice, Macro Block, Block の階層構造となっている。

Picture には I-picture (Intra-coded Picture), P-picture (Predictive-coded Picture), B-picture (Bidirectionally Predictive-coded Picture) の 3 種類がある。これら 3 種類の Picture の集まった GOP は、ランダムアクセス機能と高い符号化効率を両立させるための仕組みである。GOP はランダムアクセス可能な I-picture から始まり、次の I-picture の直前の P-picture で終わる。

符号化は、フレーム/フィールド適応動き補償予測とフレーム/フィールド適応 DCT によって行われる。フレームの状態でもクロブロックを構成し、動き補償や DCT を行うフレーム構造の符号化と、フィールドの画素から成るクロブロックを用いるフィールド構造の符号化がある。

フレーム構造の符号化では、フレーム単位の 1 ベクトルによる動き補償とフィールド単位の 2 ベクトルによる動き補償を、クロブロック毎に切り換えることができる。また、Dual-prime というインタレース画像用の特殊な動き補償を、クロブロック単位に切り換えて用いることもできる。これはインタレース構造を利用して、動ベクトルの内挿・外挿と平均化により予測効率を向上させる技術であり、インタレース画像の場合に問題であった、ゆっくりした垂直方向の動きの補償に効果がある。

フィールド構造の符号化では、クロブロック毎の動き補償、それを上下に分割したサブクロブロック毎の動き補償、および Dual-prime を選択できる。

フレーム/フィールド適応 DCT には、フィールド別に DCT を行うモードと、2 フィールド分を交互に混ぜてフレーム構造として DCT を行うモードがある。

通信分野に適用する場合には、符号化復号遅延が短いことが望まれる。B-picture を用いなければフレーム並べ換えの遅延がなくなる。I-picture を用いずに Intra-slice を用いて周期的にリフレッシュすれば、バッファサイズが減少し、遅延を減らすことができる。

MPEG 2 のシンタクスは MPEG 1 の拡張になっており、MPEG 2 は MPEG 1 に対してフォワード互換性がある。すなわち、MPEG 2 のデコードは MPEG 1 のビットストリームを復号できる。一方、MPEG 1 の符号化アルゴリズムを B-picture を含まない構成にすれば、そのハードウェアの機能は H. 261 の符号化処理に近づく。DCT 係数のジグザグス

表 1 MPEG 2 メインプロファイル・メインレベルの仕様

項 目	内 容
画像フォーマット	CCIR 601 (720×480×29.97 Hz, 720×576×25 Hz) 以下
ビットレート	15 Mb/s 以下
色信号	4:2:0
ピクチャタイプ	I, B, P-pictures
符号化画像単位	フレーム構造およびフィールド構造
動き補償予測	
フレーム構造	16×16 フレーム予測, 16×8 フィールド予測, Dual-Prime(I, P 構造の場合)
フィールド構造	16×16 フィールド予測, 16×8 フィールド予測, Dual-Prime(I, P 構造の場合)
動き補償範囲	垂直方向 [-127.5, +128.0], 0.5 画素精度, 水平方向 [-704.0, +704.0]
バッファサイズ	1.835008 Mb 以下
互換性	MPEG 1 フォワード互換
イントラ DC 精度	8, 9, 10 ビット
イントラ VLC	MPEG 1/New Table(ピクチャレイヤで切り換え可能)
DCT 係数スキャン	MPEG 1/New Scan(ピクチャレイヤで切り換え可能)
エラー耐性	イントラクロブロック動ベクトルによるコンシールメント
VBR 動作	含まれる

キャンは MPEG 2, MPEG 1, H. 261, JPEG で共通である。また, DCT 係数の量子化に量子化マトリクスを用いることも同じである。したがって, MPEG 2 や MPEG 1 のデコーダは, H. 261 や JPEG の圧縮データを簡単な変更で復号できる。

2.3 画像用 LSI^{(15)~(18)}

画像情報圧縮の普及には, 方式の標準化と共に LSI 化が不可欠である。H. 261 や JPEG の標準化の成功によって, DCT チップを含む画像符号化専用 LSI の価格低下が加速された。また, MPEG 1 の標準化によって MPEG 1 デコーダの LSI 開発に拍車がかかった。解像度やフレームレートを限定すれば, 1 ボードや 1 チップの画像コーデックも開発されるようになってきた。デコーダについては, すでに H. 261 および MPEG 1 の 1 チップ LSI が開発されている。現在, MPEG 2 の仕様決定を受けて, CCIR 601 画像を 5~10 Mb/s 程度に圧縮する MPEG 2 対応の LSI が次々と発表されている⁽¹⁹⁾。

画像の情報圧縮アルゴリズムは大量のデータを積和演算する形のものが多いが, その処理を LSI で実現するには, 演算器の並列配置が大きな効果を発揮する。このため, 比較的高速に動作し, かつ回路規模の増大にも耐えうる CMOS の大規模ディジタル系の LSI 技術が用いられる。現在では数十万から百万トランジスタ以上の LSI が開発されるようになり, また 4 Mb や 16 Mb 等の大容量メモリも実用化されている。ISSCC で発表された DRAM の発表年とメモリー容量の関係を図 3 に示す。約 3 年毎に 4 倍の高集積化が実現されている。

図 3 ISSCC で発表された DRAM のメモリー容量

MPEG 2 メインプロファイル・メインレベル用と MPEG 1 用の復号 LSI の規模を表 2 に示す。MPEG 2 対応 LSI と外付けメモリーとの間のデータ転送バンド幅は 100 MB/s を超えることが予測される。データ転送バンド幅を小さくする工夫が報告されている⁽²⁰⁾。

DSP を用いれば, プログラムの変更だけで JPEG, H. 261, MPEG などの様々な方式に対応でき, 制御アルゴリズムの更新や追加も行える。また, 縮小, 拡大, 回転などの画像加工にも対応できるため, 専用 LSI に比べ有利である。

画像圧縮 LSI が処理すべきデータ量と所要演算性能の関係を図 4 に示す。CIF/SIF 画像では GOPS オーダー, CCIR 601 画像では数 GOPS 以上, HDTV では数十 GOPS 以上の性能が必要である。

現在, 汎用 MPU や汎用 DSP で実現されている性能は 100 MIPS, 100 MOPS 程度である。これは実時間処理に必要な GOPS オーダーの性能に比べると 1 桁以上低い, これに対して, 並列演算器やデータメモリーの 2 次元アドレッシング機能などを持つ画像用 DSP は, 汎用 DSP の約 10 倍相当の処理能力を実現

表 2 MPEG 2 メインレベル・メインプロファイル用と MPEG 1 用の復号 LSI の規模⁽¹⁹⁾

項 目	MPEG 2	MPEG 1
名付け DRAM の容量	8~32 Mb	4~16 Mb
製造技術	0.5~0.8 μm の CMOS	0.7~1.0 μm の CMOS
トランジスタ数	80 万~250 万	30 万~120 万
消費電力	1~3 W	1~2.5 W

図 4 処理すべきデータ量と所要演算性能⁽¹⁶⁾

できる。ピーク値では1 GOPS以上の画像用DSPも開発されている。

半導体プロセスの進歩と性能向上予測によれば、5年で1桁程度の性能向上が見込めるので、数年後には1チップの汎用MPUまたは汎用DSPによるSIF画像の符号化が可能になると予測される。

しかし、画像の高精細化に伴う画素数の急激な増加と処理の高度化に、DSP単体の高性能化だけで対処するのは困難である。複数のDSPと大容量メモリを1チップに集積化したSIMD (Single Instruction Multiple Data stream)型の並列DSPや、超並列コア(例えば、動きベクトル検出専用)と汎用MPU/DSPコアの混載形式などが、今後の画像用DSPとして有望である。

2.4 む す び

30年以上にわたる画像情報圧縮アルゴリズムに関する知識の集積は、国際標準化の成功とLSI技術の進歩により実を結ぼうとしている。画像情報圧縮の基本は画像表現法にある。画像情報の本質の探究と、それを踏まえた画期的な画像表現法の出現を期待したい。

(谷本)

(1994年5月16日受付)

〔参 考 文 献〕

1) 安田浩編著：“マルチメディア符号化の国際標準”，丸善(June 1991)

- 2) H. S. Malvar: “The LOT: Transform Coding Without Blocking Effects”, IEEE Trans. Acoust., Speech, Signal Process., **ASSP-37**, 4, pp. 553-559 (Apr. 1989)
- 3) O. Rioul and M. Vetterli: “Wavelets and Signal Processing”, IEEE SP Magazine, 8, 4, pp. 14-38 (Oct. 1991)
- 4) P. J. Burt and E. H. Adelson: “The Laplacian Pyramid as a Compact Image Code”, IEEE Trans. on Commun., **COM-31**, 4, pp. 532-540 (Apr. 1983)
- 5) M. Tanimoto and A. Yamada: “A New Complimentary Pyramid Scheme of Images”, Picture Coding Symp., 6, 1, pp. 6. 1. 1-6. 1. 2 (Sep. 1988)
- 6) J. Katto and Y. Yasuda: “Performance Evaluation of Subband Coding and Optimization of Its Filter Coefficients”, SPIE Visual Commun. and Image Processing '91, **1605**, pp. 95-106 (Nov. 1991)
- 7) M. Tanimoto and N. Wakatsuki: “A New Subband Coding Scheme Employing Emphasis”, International Workshop on HDTV (Nov. 1993)
- 8) 斎藤 鄭: “新しい画像符号化技術—フラクタル理論を中心として—”, 信学誌, **75**, 12, pp. 1343-1355 (Dec. 1992)
- 9) M. Kaneko, F. Kishino, K. Shimamura and H. Harashima: “Toward the New Era of Visual Communication”, IEICE Trans. Commun., **E76-B**, 6, pp. 577-591 (June 1993)
- 10) 小野: “超高精細画像”, テレビ誌, **47**, 9, pp. 1204-1215 (Sep. 1993)
- 11) 大越: “三次元画像工学”, 朝倉書店 (July 1991)
- 12) 坂内: “三次元画像技術の発展”, 画像電子学会アドバンストイメーシング '92 [先端の3次元画像技術], pp. 1-10 (Apr. 1992)
- 13) 吉川: “ホログラフィック3次元画像の圧縮技術”, テレビ誌, **47**, 2, pp. 177-182 (Feb. 1993)
- 14) 渡辺: “MPEG 2/H. 262”, テレビ誌, **48**, 1, pp. 44-49 (Jan. 1994)
- 15) 山内: “DSPの現状と将来展望”, 電子材料, pp. 38-43 (June 1992)
- 16) 笠井, 小倉: “画像用LSI技術の現状と展望”, テレビ誌, **48**, 1, pp. 14-19 (Jan. 1994)
- 17) 松田, 浅見, 川井: “画像圧縮用LSI”, テレビ誌, **48**, 1, pp. 31-37 (Jan. 1994)
- 18) 山品: “画像処理DSP”, テレビ誌, **48**, 1, pp. 38-43 (Jan. 1994)
- 19) “画像圧縮方式MPEG 2次世代インフラの要に”, 日経エレクトロニクス, 803, pp. 77-116 (Mar. 14, 1994)
- 20) ワテル, 相澤, 羽鳥: “低メモリアクセスの半ペル精度動き推定回路”, マルチメディアと映像処理シンポジウム, pp. 79-83 (Apr. 1994)