

# 超低ビットレート動画像圧縮技術

松本 修一 宮地 悟史

(株) KDD 研究所 画像通信グループ

## 期待されるインターネットビデオ

ISDNやPHSといった身近なデジタル回線の普及により、報道取材映像や監視カメラ映像といった動画像の伝送アプリケーションがここ数年急速に活用されつつある。また、電子メールやネットニュース、ftpなどで一般に浸透してきたインターネットは、WWW (World Wide Web) の出現により爆発的に普及し、当初は文書や静止画が中心であったコンテンツも、圧縮技術の進歩とともに現在では動画像が数多く見られ、本格的なインターネットマルチメディア時代が到来したといえる。インターネットでの動画像伝送は、ユーザが端末回線として従来の電話回線やISDN、PHS等を用いることを想定しているため、伝送ビットレートとしてアナログモデムの28.8kbpsからISDN (2B) の128kbpsまでを対象とするのが一般的であろう。

このように、インターネットや報道取材映像などでは、数十kbps～数百kbpsを想定した超低ビットレート動画像圧縮技術の発展が大きな支えとなっている。

## 動画像圧縮とは

動画像圧縮の研究は実質30年以上にもさかのぼり、本格的な実用化段階を迎えたのは、今から10年ほど前といえるだろう。そもそもの圧縮の基本は画面を構成する画素のうち、互いに隣り合っている画素の明るさや色は非常によく似ているという高い相関性を利用して、すでに送信済みの周りの画素の値で次にくる画素の値を逐次予測する予測符号化 (DPCM) である。この1つとして、画面中の個々の物体の動きを検出し、動き量で予測値を補正する通称“動き補償”(MC) 技術の開発によって、圧縮率が飛躍的に伸びた。さらに、古くからその研究が行われていた、離散的コサイン変

換 (DCT) と呼ばれるある種のフーリエ変換技術の導入によって、圧縮処理が画像信号のみならずその周波数成分にまで拡大し、さらなる圧縮率の拡大が図られた。この動き補償とDCTを組み合わせたMC-DCT技術がITUやISOのこれまでのすべての動画像圧縮標準の基本となっている。画像圧縮はもとより、画像の持つこの高い相関性と人間の視覚的な特性の2つの性質に基づくが、インターネットやISDNなどの低速ビットレート用の圧縮では、1/1000～1/10000もの高圧縮が必要となることから、解像度の低下や圧縮特有の雑音はもとより、画面サイズの縮小やフレーム間引きなど、相当の画質劣化を前提としている。この点で、数Mbpsから数十Mbpsに及ぶ放送テレビ用の圧縮技術とは要求品質を大きく異にしているといえる。

## これまでの標準化動向

動画像圧縮では、圧縮処理と復号処理が対をなすいわゆる相互接続性が要求されるため、圧縮方式の統一化が必須であり、ITUやISOは国際標準方式の確立にこれまで精力的に取り組んできた。図-1に国際標準化の流れを示す。超低レート動画像圧縮の国際標準として、最初に定められたのが1990年のITU勧告のH.261<sup>1)</sup>であり、64kbpsの整数倍でのビットレートで主に画面の動きが少ないテレビ電話、会議を対象としていた。その後、1992年にISOは、CD-ROMなどの蓄積媒体用の圧縮標準としてH.261の技術に基づいたMPEG-1<sup>2)</sup>を標準化した。ランダムアクセス、高速サーチ、逆再生などに対応するもので、ビットレートの上限は1.5Mbpsと定められた。このMPEG-1に続いて、ISOでは1995年に現行テレビからHDTVまで汎用的に圧縮が行えるMPEG-2<sup>3)</sup> (ITUの勧告番号はH.262) を標準化した。対象ビットレートは画像品質を重視し、数Mbps～数百Mbpsでかなり高く設定されている。1996

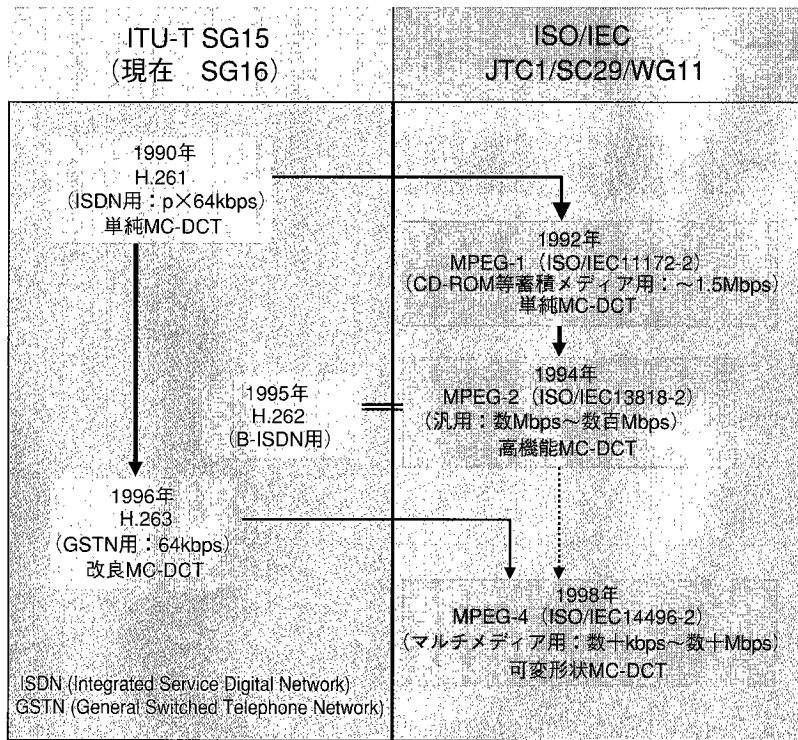


図-1 動画像圧縮の国際標準化の流れ

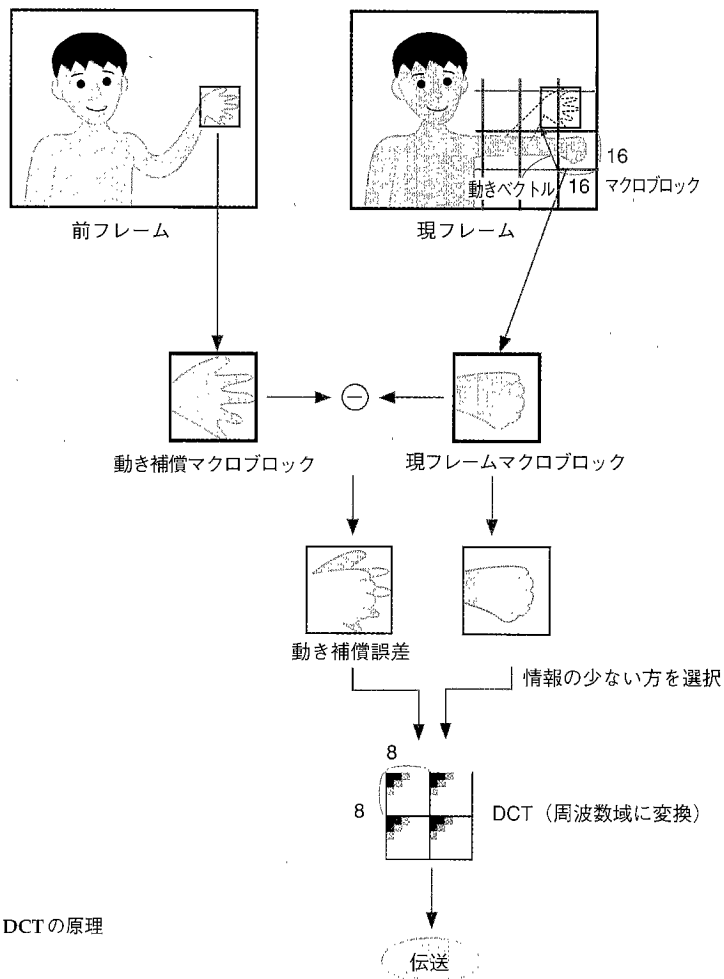


図-2 動き補償DCTの原理

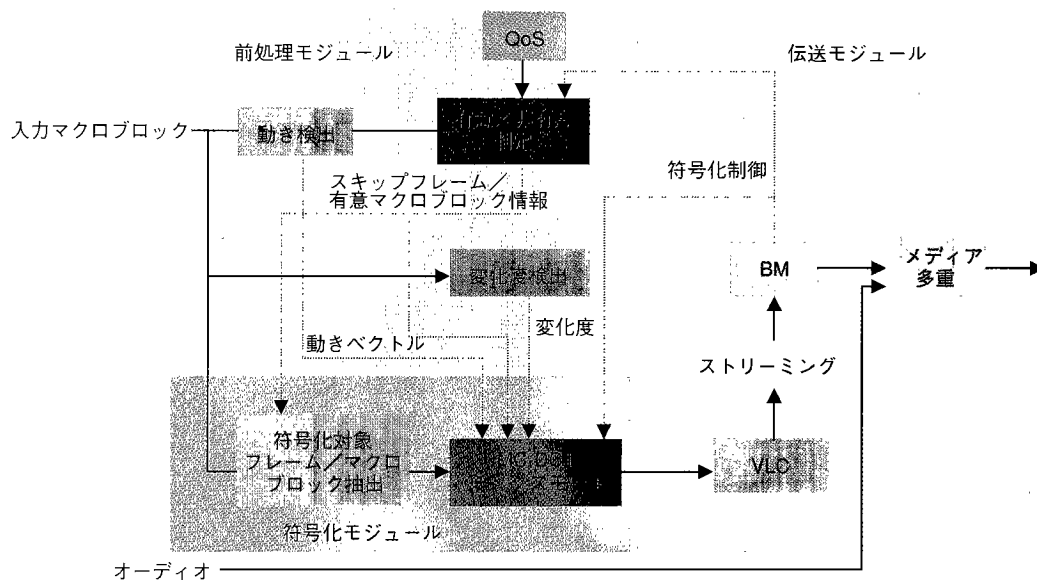


図-3 APCエンコーダのブロック構成図

年には超低ビットレート用としてITUからH.263<sup>4)</sup>が勧告化された。これは先のH.261をベースとして、動き検出精度の向上、圧縮情報に割り当てる可変長符号化の改良などにより、おおよそ2倍の圧縮率を実現させ、現在のインターネットビデオの配信アプリケーションやパソコンテレビ会議などの多くのサービスに採用されている。

一方、ISOでもMPEG-2の標準化に引き続き、超低ビットレート用の符号化標準MPEG-4<sup>5)</sup>の検討作業に入った。当初予定していた画期的な高圧縮符号化方式の実現も、圧縮限界に直面し、方針変更を余儀なくされた。その結果、符号化対象とする画像や用いる技術に関して、H.263の拡張に加えて、画面中の物体、背景ごとの符号化や操作、自然画と合成画との統合、各種階層化、回線誤りに対する高度な耐性化といった圧縮標準の実現を目指し、1998年に標準化される。圧縮技術の立場から、強いてMPEG-4の特長を1つ挙げるならば、その有意性は別として、従来矩形固定のブロック単位での動き検出やDCTを可変形状に拡張した点であろう。

### 動き補償DCTとは

圧縮の基本である動き補償DCT符号化の原理を図-2に基づき以下に説明する。

まず、入力された動画の符号化対象フレーム(画面)を $16 \times 16$ のブロック(マクロブロック)に分割する。次に各マクロブロックごとに直前フレーム中から形状や明るさの最も似たマクロブロックを探し、このマクロブロックの位置から移動量(動きベクトル)を

検出する。現フレームのマクロブロックと前フレームの移動量に対応した類似マクロブロックとの差分は動き補償誤差と呼ばれる。マクロブロックの動き補償誤差がマクロブロック自身よりも情報量が少ないと判定されたとき、動きベクトルと動き補償誤差が符号化される。それ以外の場合にはマクロブロック自身の画素が符号化の対象となる。動き補償誤差あるいはマクロブロック自身の画素は、 $8 \times 8$ のブロックでDCT(離散コサイン変換)により周波数領域に変換される。それまでマクロブロック内にランダムに分布していた誤差あるいは画素の値は、DCTにより低周波数成分に大きな値が集中する。この偏りを利用してDCTの施されたブロックは、低周波数成分から順に短い符号で動きベクトルとともに符号化され、情報圧縮される。圧縮後のビットストリームが蓄積や伝送の対象となる。品質的には、たとえばH.263方式では、画面サイズが縦・横各 $1/4$ の $1/16$ で、毎秒フレーム数が2、3枚程度、遅延時間が数秒といった性能がインターネット上で確認されている。

### 克服すべき3つの技術課題

特にインターネットを意識した超低ビットレート動画符号化技術の究極の目標は、高画質、オールソフト、双方向通信、リアルタイム伝送であろう。H.263を始めとして、さまざまなソフトウェアがインターネット上で動画配信に用いられているものの、これら課題をすべて満たすものはいまのところ存在しない。すなわち、画像品質の面では、画質とフレーム数がトレードオフの関係にあり、両立はできない。たとえば、画



図-4 APCクライアントGUI

質を優先させた場合、毎秒送るフレーム数が少なくなるいわゆるスキップフレームが発生する。さらに、これは伝送遅延時間を増大させる要因となり、双方向通信に支障をきたす。一方、画質を犠牲にして、フレーム数を増やすとパソコンに対する許容演算時間が短くなり、結果的に符号化処理が間に合わず、リアルタイム伝送が不可能となる。このように、品質、遅延、演算処理の3つの課題を従来の圧縮技術は抱えているといえる。

## APC符号化方式の開発

前述した3つの課題に可能な限り対処すべく、KDD研究所が開発したAPC (Advanced Pre-Coding) 符号化方式を紹介する<sup>6)</sup>。APCでは、高画質化を目的とした高度前処理技術、演算量の軽減を目的とした動き補償簡易化技術、遅延時間の低減を目的とした映像と音声の可変表示技術が導入されている。

図-3にAPC符号化アルゴリズムの全体構成を示す。システムは前処理、符号化、伝送の3つのモジュールから構成されている。

### (1) 前処理モジュール

前処理モジュールでは入力画像の各マクロブロックに対して前フレーム中で類似のマクロブロックを探索する動き検出を行い、各マクロブロックごとにそれらが静止領域(動きベクトルが零のとき)に属するものと移動領域(動きベクトルが非零のとき)のものに分類する。静止領域のマクロブロックは前フレームからの同一位置のマクロブロックの画像がそのまま利用できるため、符号化伝送の必要がなく非有意マクロブロックと判定される。一方、移動領域のマクロブロックについても入力マクロブロックと前フレームの動きベクトルに対応したマクロブロックとの差分値の大きさに基づき、非有意と有意の各マクロブロックに分けられる。以上が有意/非有意判定の動作となる。そして有意マクロブロックのみ符号化モジュールに入力され集中的に情報ビットが割り当てられる。同時にフレーム内で発生する有意マクロブロック数と符号化モジュールで実際に発生する情報量(ビット数)には高い相関があることから符号化モジュールでの発生情報量を事前に予測でき、この予測結果と要求されるQoS(サ

ービス品質)に基づき次の符号化対象とするフレームまでスキップさせるフレーム間引きの数を決定する<sup>6)</sup>。すなわち、符号化前処理の段階で各画像フレームのQoSを保証した状態での最適ビット数の割り当てが行われ、その情報が符号化部へと引き継がれる。同時に、後述する符号化の際に見た目の劣化を最小限なものとするため、圧縮時に視覚特性を反映する目的で、各フレーム中の信号の変化度(平坦, エッジ, 精細など)の検出も前処理モジュールで行う。

## (2) 符号化モジュール

符号化モジュールでは、前処理モジュールで決定されたフレーム間引き後の画面中の有意マクロブロックのみが抽出され、これに対して、動き補償DCT符号化が適用される。ただし、単純に動き補償モードと入力マクロブロックモードの2つを切り替えるのではなく、これら2つのモードをDCT係数上で低次係数を動き補償モード、高次係数を入力マクロブロックモードでミックスさせる独自のモードが採用されている。これは、符号化雑音の多い超低ビットレート符号化において、たとえ発生情報量が多くてもその視覚的印象が動き補償で符号化されるよりもマクロブロックを直接符号化した方がよいことに基づいている。ミックスモードにおける動き補償と入力マクロブロックの境界は、フレーム間引き数や発生情報量により適応的に制御される。また各DCT係数は与えられた数で割られる量子化と呼ばれる丸め操作が適用されるが、この際視覚的な要因を考慮して、割る数をDCTの係数位置や画像の変化度によって適応的に変え、その結果、量子化による劣化を最小限なものにしている。

## (3) 伝送モジュール

伝送モジュールでは、前処理モジュールで発生した動きベクトルや各制御情報および符号化モジュールで発生した量子化情報を可変長符号化し、ビットストリームをいったんバッファメモリに蓄積した後、一定のレートで伝送路に出力する。同時にバッファメモリ蓄積量を、前処理や符号化の両モジュールの制御のためにフィードバックさせる。

受信側では、送信側とまったく逆の操作により復号画像を得る。表示は映像と音声の相対遅延時間と映像の絶対遅延時間のトレードオフを考慮して、低遅延モードには前者を、高遅延モードには後者を犠牲にする可変表示技術を導入している。

以上のような入念な符号化制御により、APC方式はH.263と比較して、SN比で最大3~4dBの品質向上が得られており、低ビットレート画像で頻繁に見られるブロック形状のノイズや不自然な画面の動きが少なく、

視覚的な品質も大幅に向上している。

## インターネットによる動画配信システムの構築

APCを用いたインターネットによる動画配信システムは、独自開発のサーバ、エンコーダ、クライアントから構成されている。対象とする伝送ビットレートは24kbps~384kbpsと幅広く、インターネット上でAPCの高画質映像を配信する。クライアントのGUIを図4に示す。APC符号化された動画は、サーバにおいてITU-T標準方式G.723.1(5.3/6.3kbpsデュアルレート音声符号化方式)で圧縮された音声データとともにオーバーヘッドの非常に小さい独自フォーマットの packets 多重化がなされる。配信形態はリアルタイムライブ形式とオンデマンド形式に対応し、UDP(ユーザデータプロトコル)もしくはTCP(伝送制御プロトコル)上のHTTP(ハイパーテキスト伝送プロトコル)の各プロトコルによりインターネットへ配信される。HTTPを用いることにより、ファイアウォール環境下でのアクセスも可能としている。

オンデマンドアクセス時には、クライアント側からの再生、一時停止、ランダムアクセスなどの要求コマンドに対応してサーバが応答し、冗長度を極力小さくした独自のコマンド形態によりレスポンスのよいアクセスを実現している。

## ポイントとなるQoS

ユーザ要求がしっかり定義され、それに基づき規格が定められているデジタルTV放送規格とは異なり、超低レート動画圧縮の対象となるインターネットや移動体による画像通信サービスでは伝送環境が不安定であるため、サービス品質の規定ならびにその保証とといったいわゆるQoSの問題が、今後大きくクローズアップされよう。

### 参考文献

- 1) ITU-T Recommendation H.261: Video Codec for Audiovisual Services at p x 64kbit/s (1993) .
- 2) ISO/IEC 11172-2: Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at Up To About 1.5Mbit/s - Part 2: Video (1993) .
- 3) ISO/IEC 13818-2: Information Technology - Generic Coding of Moving Picture and Associated Audio Information - Part 2: Video (1995) .
- 4) ITU-T Recommendation H.263: Video Coding for Low Bit Rate Communication (Mar. 1996) .
- 5) ISO/IEC CD 14496-2: Information Technology - Coding of Audio-Visual Objects - Part 2: Visual (Mar. 1998) .
- 6) 宮地, 松本: 統合的時空間制御による超低ビットレート動画符号化方式, 映像情報メディア学会誌, Vol.51, No.10, pp.1706-1714 (Oct. 1997) .

(平成10年10月2日受付)