

モバイルマルチメディア信号処理技術特集

映像符号化技術

IMT・2000のマルチメディア通信サービスの中核の一つであるMPEG・4映像符号化方式について、その要素技術、ベースとなった従来の映像符号化方式H.261、MPEG・1、MPEG・2、H.263との関連、特徴を解説した。

やまぐち ひろゆき えとう みのもり あだち さとる
山口 博幸 栄藤 稔 安達 悟

1. まえがき

次世代移動通信（IMT・2000：International Mobile Telecommunications・2000）のマルチメディアサービスにおいては、ビデオフォン、映像配信をはじめとする色々な場面で、映像符号化方式MPEG（Moving Picture Experts Group）・4が使用される。MPEG・4は、これまでの映像符号化技術の集大成であるとも位置付けられる。

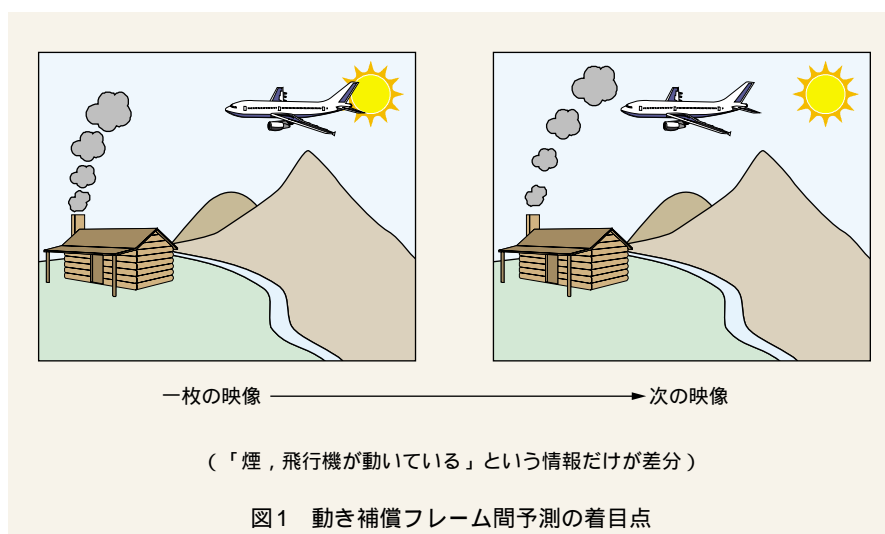
本稿では、その要素技術ならびにMPEG・4映像符号化方式に到るまでの各種映像符号化方式の特徴を解説する。

2. 映像符号化要素技術

通常の映像信号は、そのままでは約100Mbit/sもの大量の情報となる。そこで、映像の性質を巧みに利用する、高能率な映像符号化方式が各種考案されてきた。これらに共通的な要素技術として、動き補償フレーム間予測、離散コサイン変換(DCT：Discrete Cosine Transform)、可変長符号化が挙げられる[1]～[3]。

(1) 動き補償フレーム間予測

動き補償フレーム間予測とは、映像の1枚1枚を符号化するのではなく、前後の画面を参照して特定の部分が、



どの方向にどれだけ動いたかの移動量で指し示す手法である(図1)。動く方向・大きさ(動きベクトル)は、画面中の部分部分で異なる。そこで、画面を 16×16 画素程度に小さく分割し(これをマクロブロックと呼ぶ)、それぞれについて動きベクトルを求める。さらに、マクロブロックを移動させた場所での前画面との差異を予測誤差と呼び、これに対し、(2)項の離散コサイン変換を施す。

(2) 離散コサイン変換

映像の一枚一枚の画面は、単純な映像成分(低周波数成分)からより複雑な映像成分(高周波数成分)までの、重み付の足し算で表現される(図2)。

そのうち、低周波数成分に情報が集中し、視覚的にも重要な役割を果たすことが知られている。DCTは、重要な

周波数成分のみを最終的に取り出して情報圧縮することをねらいとするもので、映像の空間周波数領域への変換を効果的に行うことができるため、広く採用されている。

実際には、画面を 8×8 画素程度のブロックに分割し、それぞれのブロックに対しDCT処理を行う。図2中の a_i はDCT係数と呼ばれ、この係数をさらに量子化、量子化代表値への丸め処理をして、(3)項の可変長符号化を施す。

(3) 可変長符号化

入力信号値の偏りを利用して、多く発生する信号値に短い符号を割り当て、そうでないものに長い符号を割り当てて、情報圧縮するのが可変長符号化である。

DCTの項で述べたとおり高周波数

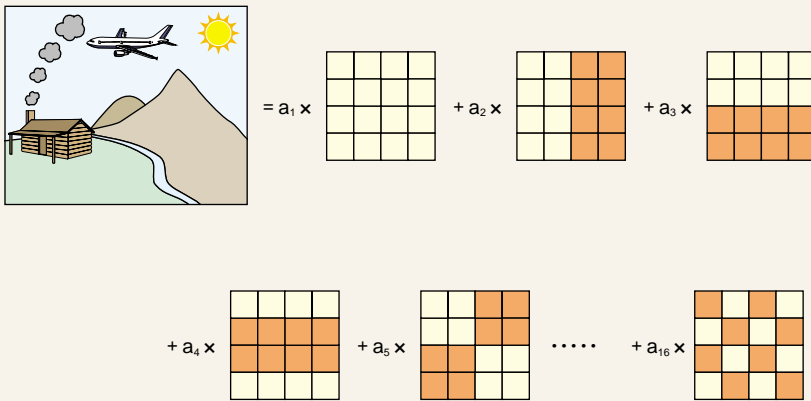


図2 画面の周波数成分への分解概念

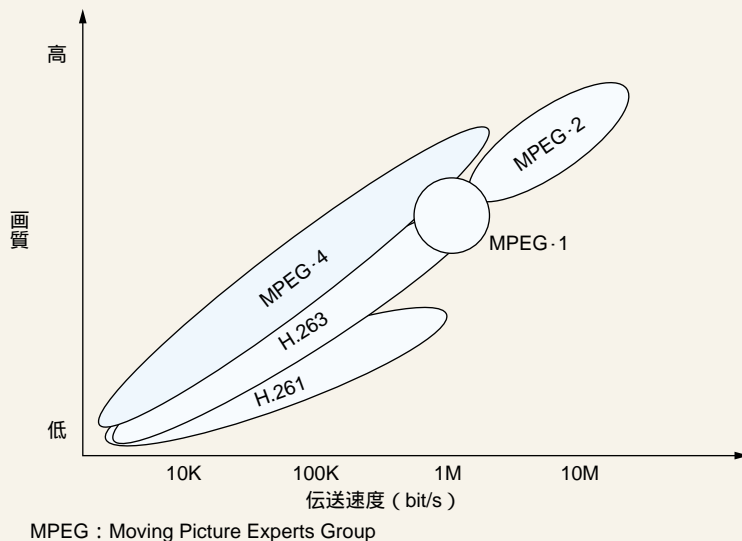


図3 MPEG・4映像符号化と他標準の関係

成分の係数は、量子化代表値への丸め処理によりゼロのものが多くなる。そこで、「これ以降の全部の値はゼロ (EOB : End of Block)」とか「値ゼロが何個続いた後に値L」という場合が多く発生する。さらに、ゼロの個数 (ゼロラン) とLの値 (レベル) の組合せについても考慮し、より多く発生する組合せに短い符号を割り付けることで情報圧縮できる。以上は、2つの組合せに符号を一つ割り当てる手法について述べた。この手法は一般に2次元可変長符号化と呼ばれている。

3. 映像符号化各方式の位置付け

国際標準化された映像符号化方式として、H.261, MPEG-1/2, H.263, MPEG-4が挙げられる。それぞれの適応領域を図3に示す。各方式において、前述の要素技術がどのように使用され、圧縮効率を向上させているか、また、それぞれの機能の差異はどこにあるかを述べる。

(1) H.261 映像符号化

ISDN (Integrated Services Digital

Network) 用テレビ電話・会議での通信を対象とした、実質的に最初の国際標準映像符号化方式であり、1990年国際電気通信連合・電気通信標準化部門 (ITU-T : International Telecommunication Union - Telecommunication Standardization Sector) にて標準化された[4]。H.261は、前節で示したすべての要素技術を使用している。すなわち次のとおりである。

- ・16×16画素のマクロブロックに対して画素単位での動きベクトルを予測し、動き補償フレーム間予測を行う
- ・前画面との予測誤差に対し、8×8画素のDCTを施す。なお、予測誤差が一定量以上となる、動きの激しい領域については、適応的に動き補償フレーム間予測せず、フレーム内での8×8画素のDCTを施すことで符号化効率を向上させる
- ・フレーム間動き補償で得られた動きベクトルと、DCT処理結果をそれぞれ可変長符号化する。なお、DCT処理結果に施す可変長符号化には、2次元のものを使用する

H.261は、現行テレビカメラ・モニタの利用を前提としている。しかしながら、テレビ信号形式 (フレーム枚数、走査線数) は、地域ごとに異なっており、国際通信に対応可能とするため、中間的な共通フォーマットに変換してこれを符号化することが考案された。

このフォーマットをCIF (Common Intermediate Format) と呼び、「352 (水平) × 288 (垂直) 画素、最大30フレーム/秒、ノンインタレース」で定義された。CIFの1/4の面積をもつQCIF (Quarter CIF) も同時に定義され、以降の映像符号化でも利用されている。

(2) MPEG-1/MPEG-2 映像符号化

MPEG-1は、CD-ROMで代表される蓄積メディアでの利用を対象として、1993年国際標準化機構 (ISO : Inter-

national Organization for Standardization) /国際電気標準会議 (IEC : International Electrotechnical Commission) にて標準化された[5]。1.5Mbit/s 程度のデータ量をターゲットとする映像符号化方式である。蓄積メディア用の符号化であるため、符号化処理の実時間性は H.261 と比べ要求条件が緩和され、新技術が採用される余地が増えるとともに、映像のランダムサーチ機能が新機能として要求される。そこで、基本的には H.261 と同じ要素技術を使用しつつ、以下の機能が追加されている。

- ・定期的に一画面全部をフレーム内符号化して、ランダムアクセス再生を可能とする
- ・ H.261 では過去の画面から、動きベクトルを予測して動き補償フレーム間予測をしているが(これを順方向予測と呼ぶ)、これに加えて、蓄積メディアの利点を利用して未来の画面からの予測(逆方向予測)もできるようにした。さらに、順方向予測、逆方向予測、順方向予測と逆方向予測の平均の3とおりを求め、最も予測誤差が小さいものを使用して、圧縮率を向上する。
- ・ H.261 では、1画素単位で動きベクトルを予測しているが、これを0.5画素単位の予測とする。そのために、隣接する画素と画素の間に、値が平均となる補完画素を加えた画面を作成し、これに対してフレーム間動き予測を行い、圧縮率を向上する

これらの機能追加により、パソコン用のビデオ符号化・プレーヤとして一般に使用されている。

MPEG-2は、通信、放送、蓄積用途の要求条件を考慮した汎用映像符号化であり、1996年にISO/IECにて標準化され、ITU-T H.262との共通テキストとなっている[6]。デジタルテレビ放送、HDTV(High Definition Television)、DVDでの3~20Mbit/sの映像符号化に使用されている。MPEG-1の技術要

素を反映し、さらに次の機能追加を行っている。

- ・ 現行のTV信号で使用されているインタレース画像の効率的な符号化の追加
- ・ 符号化データの一部を取り出すだけで必要に応じて、画面サイズや画質を調整できる機能(それぞれ、空間スケーラビリティ、信号対雑音比(SNR : Signal to Noise Ratio)スケーラビリティと呼ぶ)の追加

さまざまな用途を想定した機能追加により、特に符号化データの互換性を考慮する必要が生じる。そこでMPEG-2では、機能の違い、処理の複雑さを分類した、プロファイルとレベルという概念を新たに導入して問題を解決している。MPEG-4でも同じ概念が使用されている。

(3) H.263映像符号化

アナログ電話網でのテレビ電話を対象とした超低レート映像符号化方式であり、1996年ITU-Tにて標準化された[7]。28.8kbit/sのモデム利用を想定し、H.261に対してMPEG-1で開発された新技術の一部も盛り込んだ位置付けを有する。0.5画素単位でのフレーム間動き補償予測が基本機能(ベースライン)として必須となっている。可変長符号化を、従来の(ラン、レベル)の2次元から、EOBも含めた3次元に拡張し、符号化効率を向上させるものが考案され、同様にベースラインとなっている。さらにオプションとして、8×8のブロック単位での動き補償フレーム間予測や、画像のブロック状の歪みを緩和する処理などが新規に加わっている。

これらの機能追加により、H.263はISDNテレビ電話・会議でも一部の機器で使用されるようになってきている。

4. MPEG-4映像符号化

MPEG-4映像符号化は、ITU-T

H.263映像符号化をベースにさまざまな改良を施し、誤り耐性を向上させたものである。H.263のベースラインとは互換性を有する。

MPEG-2映像符号化は、計算機上での映像ハンドリング、デジタル放送、高速通信などを主に志向する方式であった。MPEG-4映像符号化では、これらに加え、通信分野、特に移动通信への適用も強く意識して標準化が進められた。

その結果、極めて汎用的な映像符号化方式として1999年ISO/IEC標準として成立した[8]。それゆえ、IMT-2000におけるテレビ電話での利用のみならず、ビデオメール・配信も含めた映像利用マルチメディアサービスのキー技術と位置付けられる(図4)。

(1) プロファイルとレベル

相互通信性も含めた符号化データの相互利用を確保するため、MPEG-2と同じ思想で、プロファイルで機能を、レベルで処理量を類別する。プロファイルとしては、Simple, Core, Main, Simple Scalableなどが定義されており、そのうちSimpleプロファイルが共通機能である。H.263でオプションとされた8×8画素の動き補償フレーム間予測はSimpleプロファイルとして位置付けられている。

また、Simpleプロファイルでは、レベル1はQCIF画像、レベル2ではCIF画像を、それぞれ対象としている。

さらに、Core, Mainプロファイルでは、映像中の任意の領域を“オブジェクト”と定義し、その画質を良くしたり、他の符号化データに組み込むことも可能としている。CG画像との合成などにより、高度なプロファイルもMPEG-4では提供されている。

(2) IMT-2000での規格

本誌別稿「オーディオビジュアル端末技術」で詳しく述べるIMT-2000のビジュアルフォン規格3G-324Mでは、映像符号化方式としてH.263ベースラインを必須とし、MPEG-4のSimpleプロファイル、レベル1を推奨してい

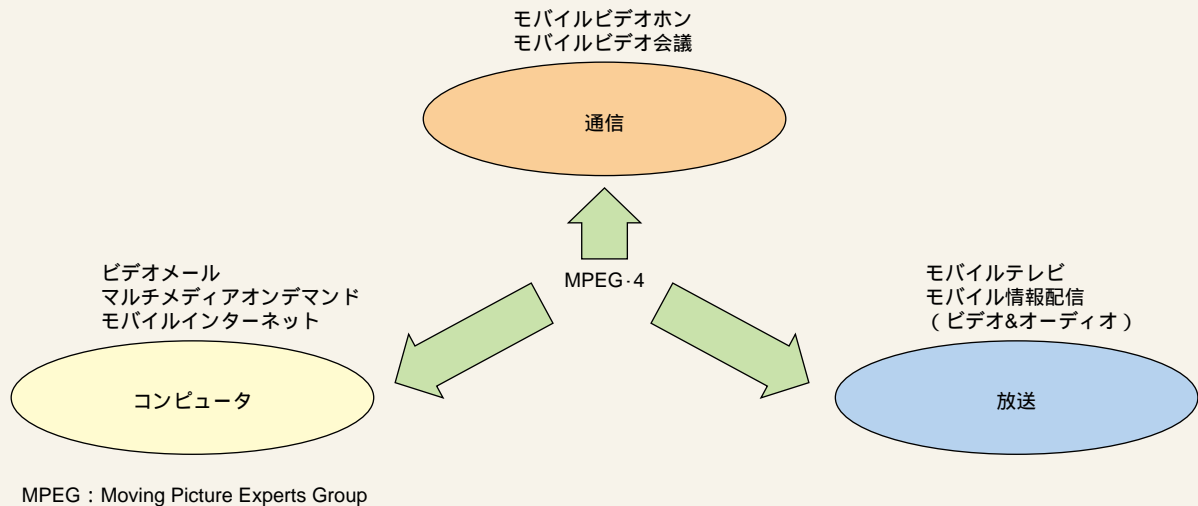
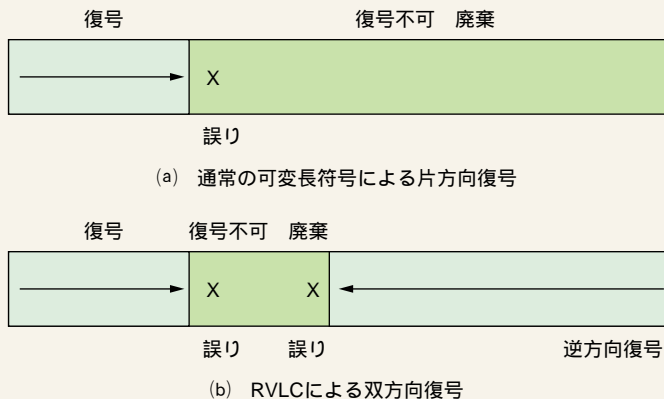


図4 MPEG・4のスコープ



RVLC : Reversible Variable Length Code (双方向復号可能可変長符号)

図5 双方向復号可能可変長符号の復号例

するものである。

以上述べたように、MPEG・4 Simple プロファイルレベル1を使用することで、移動通信に適した非常に簡単な復号化器が構成できる。

5. あとがき

「百聞は一見にしかず」のことわざは、映像利用マルチメディアサービスの広がる可能性を一言でいい表し、その反面、映像の情報量の膨大さを示している。映像符号化技術の発展がMPEG・4映像符号化として結実し、広帯域符号分割多元接続方式(W・CDMA : Wideband Code Division Multiple Access)無線アクセス技術開発もあいまって、2001年には、無線での映像利用マルチメディア通信サービスが開始されるに至った。今後さらに符号化器ハードの高性能化、低消費電力化が推進されていくことだろう。

本稿では、映像符号化の3つの要素技術を解説し、各映像符号化方式においてこれらがいかに利用され、どのように機能拡張し、情報圧縮率の向上を果たし、MPEG・4標準に反映されているかを述べた。翻ると、先に述べた3要素技術に集約できているともいえるが、学会レベルでは注目に値する新要素技術へのチャレンジも継続的に行わ

る。Simple プロファイルは、以下の誤り耐性ツールを含んでいる。

- ① 再同期...可変長符号化される符号化データに再同期コードを挿入して、画面上の適切な位置で区切り、伝送誤りを局在化する。挿入する再同期コードに続いて、符号化パラメータを指定するヘッダ情報が記入されるため、速やかに符号化誤りの状態から回復できる。再同期コードの挿入間隔は、ヘッダ情報のオーバーヘッドと、撮影対象物および伝送路特性とを勘案して最適化ができる。
- ② データ分割...符号化データ中のデータ種別が異なる境界に同期コ

ードを挿入し、誤り隠蔽を可能とするものである。例えば、動きベクトル、DCT係数の間に同期コードを挿入することで、DTC係数にビット誤りが混入しても、動きベクトルは正しく伝達されて、より自然な誤り隠蔽が可能となる。

- ③ 双方向復号可能可変長符号...図5に示すように、逆方向からも復号可能な可変長符号である。DCT係数に適用する。これにより、ビット誤りが混入したマクロブロック以外はすべて復号可能となる。
- ④ 適応イントラリフレッシュ...動き領域に対し頻繁にフレーム内符号化を実施して、誤り伝播を防止

れており、さらなる発展が期待される。

文 献

- [1] 安田浩：「MPEG/マルチメディア符号化の国際標準」丸善（1994）。
- [2] 藤原洋：「最新MPEG教科書」アスキー出版（1994）。
- [3] G.Sullivan and T.Wiegand: “ Rate - Distortion Optimization for Video Compression. ” IEEE Signal Processing Magazine, 15, 6, pp.74 - 90 (Nov.1998) .
- [4] ITU - T Rec. H.261 (1993), Video codec for audiovisual services at p × 64 kbit/s.
- [5] ISO/IEC 11172 - 2: 1993, Information technology - Coding for moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s - Part 2: Video.
- [6] ITU - T Rec. H.262 (1995) || ISO/IEC 13818 - 2: 1996, Information technology - Generic coding of moving pictures and associated audio information: Video.
- [7] ISO/IEC 14496 - 2: 1999, Information technology - Generic coding of audio-visual objects - Part 2: Visual.

用語一覧

CIF : Common Intermediate Format

DCT : Discrete Cosine Transform (離散コサイン変換)

EOB : End of Block

HDTV : High Definition Television

IEC : International Electrotechnical Commission (国際電気標準会議)

IMT - 2000 : International Mobile Telecommunications - 2000 (次世代移動通信)

ISO : International Organization for Standardization (国際標準化機構)

ITU - T : International Telecommunication Union - Telecommunication Standardization Sector (国際電気通信連合・電気通信標準化部門)

MPEG : Moving Picture Experts Group

QCIF : Quarter CIF

RVLC : Reversible Variable Length Code (双方向復号可能可変長符号)

SNR : Signal to Noise Ratio (信号対雑音比)

W - CDMA : Wideband Code Division Multiple Access (広帯域符号分割多元接続方式)