

MPEG2 の標準化動向

渡辺 裕

NTT ヒューマンインタフェース研究所

email: *hiroshi@nttvdt.ntt.jp*

MPEG2 は高品質デジタルビデオ / オーディオの汎用符号化標準である。マルチメディアのアプリケーションにとって、ビデオとオーディオのデジタル圧縮フォーマットが標準化されたことは、データ流通の観点から非常に大きな意味を持つ。また MPEG2 は HDTV やデジタル衛星放送に用いられる。デジタル化によって、放送、通信、蓄積メディアの垣根を越えた新しいアプリケーションが生まれつつある。本稿では MPEG2 の標準化作業と技術内容について述べる。

MPEG2 Standardization Work

Hiroshi Watanabe

NTT Human Interface Laboratories

email: *hiroshi@nttvdt.ntt.jp*

MPEG2 is a generic coding standard for high quality video and audio. Supplying the standard for digital compression video/audio formats has a great significance for exchanging data in many multimedia applications. In addition, MPEG2 is employed in several HDTV and digital satellite broadcasting systems. Defining the digital format is going to bring new applications in a fusion area of the conventional broadcasting, communication and storage media. In this paper, the MPEG2 standardization work and its technical contents are described.

1 まえがき

MPEG(Moving Picture Experts Group, ISO/IEC JTC1/SC29/WG11の専門家グループ)の標準化活動は、MPEG1と呼ばれる蓄積メディアを対象としたビデオ符号化から始まった。現在では、汎用デジタルビデオ / オーディオの規格である MPEG2 の IS(国際標準)化に向けて作業が行われている。MPEG は、双方向 CATV によって実現されるインタラクティブ TV や、多チャンネルデジタル衛星放送、ビデオコンパクトディスクや光磁気ディスクなどのパッケージメディア、デジタル HDTV などにおいて、マルチメディアの基幹パーツとなっている。

本稿では、MPEG2 の標準化動向について述べる。また、規格に採用されている符号化技術について簡単に述べる。現在いくつかのビデオ符号化標準に採用されている技術は、基本的に特に新しいものではない。いずれも動き補償(MC)と離散コサイン変換(DCT)を組み合わせたものである。従来からその圧縮効率の程度は知られていた。VLSI 技術の進展と DCT の標準採用による価格低下が、急激に広まる発端となった。きっかけとなったのは、テレビ電話、会議用の符号化標準である CCITT 勧告 H.261 である。その後、DCT は静止画符号化標準である JPEG でも採用された。

MPEG1 は H.261 に比べて、遅延条件を緩めて双方向予測を可能にした点が大きな特徴である。ヘッドに画像サイズなどの情報を書き込む柔軟な方式になっている。MPEG2 ではさらに、インタレース画像を効率よく符号化する技術が検討された。また、段階的に異なった画像ニーズに対応するため、スケラビリティと呼ばれる階層符号化の概念が導入されている。

2 MPEG2

MPEG2 は汎用ビデオ / オーディオ符号化標準である。1993 年 11 月に CD 13818 が、また 1994 年 3 月には DIS 13818 が完成した。ビデオパート 13818-2 は ITU-T/H.262 と共通テキストである。放送、通信、パッケージメディアのあらゆる分野のアプリケーションに適用できる。オーディオ符号化の草案と、同期および多重化と扱うシステムの草案も同時に作成された。対象とする画像は主にインタレース画像である。基本となる画像フォーマットは、CCIR601 規格(4:2:2 フォーマット)および HDTV (1920 画素 × 1080 ラインまで)である。

2.1 MPEG2 の標準化作業

MPEG2 の検討は 1990 年半ばから開始された。当初は 10Mb/s でのビデオ符号化を目標とした。1991 年 11 月の久里浜会合では 32 のビデオ符号化アルゴリズムの提案がなされた。その後、ビットストリームを規定するシンタクスを作成する作業に入った。まずテストモデルと呼ばれるシミュレーションアルゴリズムの叩き台が作成された。

テストモデルを用いて様々な技術要素を比較することにより、多くの提案技術が収束の方向に向けて動き出した。一方、オーディオ符号化は、マルチプルステレオ符号化が課題であった。しかし、バックワード互換が必要条件として挙げられていたため MPEG1 オーディオの拡張で検討が進んだ。久里浜会合以降の最近の MPEG 会合における主なトピックを表 1, 2 に示す。MPEG2 は 1994 年 11 月には IS 化の見通しである。

標準化作業に対して最も影響が大きかったのは、米国のデジタル HDTV である ATV の提案メンバーの MPEG2 への参加である。その結果 Grand Alliance が起こった。さらに、MPEG2 は ATV に採用されることになった。また同時期にデジタル衛星 TV に MPEG2 を採用することが宣言された。このような状況の中で、ケーブルラボを中心とする CATV 業界も MPEG2 を採用せざる得なくなった。欧州もアナログの HD-MAC 方式をあきらめ、MPEG2 を採用する方向に転換した。

MPEG2 は 1991 年 5 月のパリ会合以来、ITU-T SG15/WP1 ATM 符号化専門家グループと合同で会合を開催している。ATM を用いる通信用ビデオ符号化標準は H.262 と呼ばれるが、中身は MPEG2 と同一である。この時期から通信とパッケージメディアの融合が始まったといえる。高品質なディジタ

ルビデオは、本来放送がアプリケーションとして適する分野である。したがって、MPEG2 への必要機能のリストには、蓄積メディアだけでなく放送と通信のアプリケーションに関する項目も追加された。

ATV の仕様凍結スケジュールとの関連もあり、本来年 3 回のペースで開催されるはずの会合はほとんど年 4 回以上のペースである。常時 300 人近くが参加している。ビデオ、オーディオ、システムをはじめ、幾つかのサブグループに別れてバラレルに会合が進めらる。毎回議論は深夜までにおよび、徹夜で編集作業を行うことも日常茶飯事となった。多くのメンバーが Notebook PC を持参し、文書配布はディスクあるいはあるいはフロッピーベースになっている。シンタクスの正しさの確認には email で符号化ビットストリームを交換する手法がとられた。現在では ftp サイトにプログラムとビットストリームが置かれている。

表 1 最近の MPEG 会合における主なトピック (1/2)

会合	トピック
第 16 回 WG11 会合 久里浜 91 年 11 月 18 日～26 日	MPEG2 符号化の 32 提案の画品質評価 4Mb/s, 9Mb/s, CCIR601, 2 重刺激法 主観評価, 実現性評価, アルゴリズムの検査 DCT(フレーム/フィールド)/サブバンド, ウェーブレット MPEG1 バックワード互換 / 非互換
第 17 回 WG11 会合 シンガポール 92 年 1 月 7 日～1 月 11 日	TM0 の作成, MC+DCT, 4:2:0 B-picture の有効性の確認 フレーム/フィールド適応予測, フレーム/フィールド適応 DCT 視覚特性に基づく量子化とレート制御
第 18 回 WG11 会合 ハイファ 92 年 3 月 23 日～27 日	TM1 の作成 フレームベース (フレーム構造 / フィールド構造) フィールドベース (Pure Field) FAMC/Dual USNB10Mb/s 上限の撤廃要求 (リオで MPEG3 消滅) 10kb/s Video Coding の提案 (MPEG4 へ) ロスレスビデオ符号化の提案 (NP 投票へ)
第 19 回 WG11 会合 リオデジャネイロ 92 年 7 月 6 日～10 日	フレーム構造, フィールド構造 Fr/Fi/FAMC Field16x16/Field16x8 スケラビリティ MPEG2 が MPEG3 を吸収
AdhocWG11 会合 タリタウン 92 年 9 月 27 日～10 月 1 日	ATV 技術の導入, Cable Lab の要求 Fr/Fi/FAMC/SVMC/Dual-prime 8x8MV(GI) 空間 / 周波数 / クロマスケラビリティ 低遅延 Leaky 予測 (ATT) VQ による有効 DCT 係数の選択 (ATT)
第 20 回 WG11 会合 ロンドン 92 年 11 月 2 日～1 月 6 日	Fr/Fi/FAMC は Dual-prime へ (メモリバンド幅の問題) 8x8Intra/Inter/MV の提案 プロファイルとレベルによる整理 Leaky から AC-Lealy へ イントラスライス, イントラカムによる低遅延 スキャン, DC 係数精度, 非線形 Mquant, Alternate VLC 8x1 DCT, NTC 空間スケラビリティのためのフィルタ フィルムのための 10bit 入力 表示のためのパラメータ

2.2 プロファイルとレベル

MPEG2 では、幅広いアプリケーションに対して適用できるように、プロファイルとレベルという概念が用いられる。“プロファイル”は符号化ツールを適応な機能単位で区切ったものである。シンタクスが段階的に異なる。“レベル”は画像サイズなどのパラメータの大きさで区切ったものである。1993 年 4 月にはメインプロファイル、メインレベルと呼ばれる部分の仕様が凍結された。

表 2 最近の MPEG 会合における主なトピック (2/2)

会合	トピック
第 21 回 WG11 会合 ローマ 93 年 1 月 25 日～29 日	データ分割 プロファイルとレベルによるサブセットの構成 周波数スケーラビリティ (最小ドリフト, シングルモードデコーダ, サブバンド/ピラミッド構造)
第 22 回 WG11 会合 シドニー 93 年 3 月 29 日～4 月 2 日	メインプロファイル@メインレベルの仕様凍結 新しい IDCT ミスマッチ対策 拡張性のあるシンタクス 3/2 ブルダウン (テレシネ変換) への対応
第 23 回 WG11 会合 ニューヨーク 93 年 7 月 12 日～7 月 16 日	AC-Leaky, 8x1DCT などのプロファイルにも属さないツールの消滅 周波数スケーラビリティの消滅 Temporal スケーラビリティの提案 SNR スケーラビリティによる 4:2:2 Chroma Simulcast メインプロファイル@メインレベルのビットストリーム交換 米国 ATV の Grand Alliance が MPEG2 に歩み寄り 非コンパチブル (NBC) オーディオの候補に Dolby AC-3
第 24 回 WG11 会合 ブリュッセル 93 年 9 月 6 日～9 月 10 日	Main+ プロファイルの作成 使われないレベル, ツールの消滅 スケーラビリティのビットストリーム交換の準備 スケーラビリティのシンタクス整備
第 25 回 WG11 会合 ソウル 93 年 11 月 1 日～11 月 5 日	ビデオ, オーディオ, システム CD の発行 Main+ は SNR スケーラビリティと空間スケーラビリティへ分裂 スケーラビリティのビットストリーム交換 プロファイル, レベル毎にビットレート上限, パックサイズ, 階層数など仕様を規定 4:2:2 は High プロファイルのみで使用
第 26 回 WG11 会合 パリ 94 年 3 月 21 日～3 月 25 日	ビデオ, オーディオ DIS の発行 システム CD は記述に問題があり, IntrimCD として発行 オーディオは追加実験の必要性あり NBC モードのオーディオは New Work Item MPEG4 (Very Low Bitrate Audio/Visual Coding) セミナー活発
第 27 回 WG11 会合 アトランタ 94 年 6 月 8 日～6 月 10 日	システム DIS の発行

これは, CCIR601 画像を両方向予測を用いて符号化するモードである。ビットストリーム交換の基礎となる部分である。その後, SNR スケーラビリティ, 空間スケーラビリティのプロファイルと, 順方向予測を用いるシンプルプロファイルが定義された。また High プロファイルは全てのツールを含むプロファイルである。プロファイルとレベルを図 1 に示す。

メインプロファイル, メインレベルの仕様を表 3 に示す。デコーダは MPEG1 のビットストリームを復号できる能力を備える必要がある。MPEG2 のシンタクスは MPEG1 の拡張になっている。異なるアプリケーション間のビットストリームの互換性については, プロファイルの包含関係のより広い範囲をカバーする方およびレベルのより高い方が, 狭いプロファイルおよび低いレベルのビットストリームを解読できる構造になっている。

2.3 フレーム構造とフィールド構造

フレーム内符号化フレーム (I-picture), フレーム間予測符号化フレーム (P-picture), フレーム内挿予測符号化フレーム (B-picture) を用いた GOP 構造は基本的に MPEG1 と同様である。低遅延を目的として P-picture に周期的な Intra-slice を用いることもできる。MPEG2 では画像の符号化の単位をフレームだけでなく, フィールド単位にも設定することができる。フレームの状態のままマ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では, フレームピクチャを指定する必要がある。参照画像はインタレースされた前フレームの画像である。

それに対し, フィールドの画素からなるマクロブロックを用いるフィールド構造の符号化では, フィールドピクチャが指定される。参照画像は異なった 2 つのバリエーションのフィールドである。符号化の対象と

High 1920x1080x30 1920x1152x25		MP@HL 80 Mb/s			HP@HL 100, 80, 25Mb/s
High-1440 1440x1080x30 1440x1152x25		MP@H1440 60Mb/s		SSP@H1440 60, 40, 15Mb/s	HP@H1440 80, 60, 20Mb/s
Main 720x480x29.97 720x576x25 (CCIR601)	SP@ML	MP@ML 15Mb/s	SNP@ML 15, 10Mb/s		HP@ML 25, 15, 4Mb/s
Low 352x288x29.97 (CIF)		MP@LL 4Mb/s	SNP@LL 4, 3Mb/s		
Level* Profile	Simple 4:2:0 IP	Main 4:2:0 IPB	SNR Scalable 4:2:0 IPB	Spatial Scalable 4:2:0 IPB	High 4:2:0 / 4:2:2 IPB

* Typical Image Size, not the upper bound


 Currently No User

図 1 MPEG2のプロファイルとレベル

表 3 メインプロファイル, メインレベルの仕様

Item	Content
1. Image Format	≤ CCIR601 (720×480×29.97Hz, 720×576×25Hz)
2. Bitrate	≤ 15Mbits/sec
3. Chroma format	4:2:0
4. Picture type	I, B, P-pictures
5. Coding Unit	Frame Structure + Field Structure
6. MC Prediction	
Frame Structure	Frame 16×16 / Field 16×8 / Dual-Prime (only for I+P scheme)
Field Structure	Field 16×16 / Field 16×8 / Dual-Prime (only for I+P scheme)
7. MC Range	[-127.5, +128.0], 0.5 pel unit
8. Buffer size	≤ 1835008 bit
9. Compatibility	MPEG1 Forward Compatible
10. Intra DC Precision	≤ 10 bit
11. Intra VLC	MPEG1/New Table (at Picture Layer)
12. Coefficient Scan	MPEG1/New Scan (at Picture Layer)
13. Error Resilience	Intra MB Motion Vector
14. VBR Operation	Included

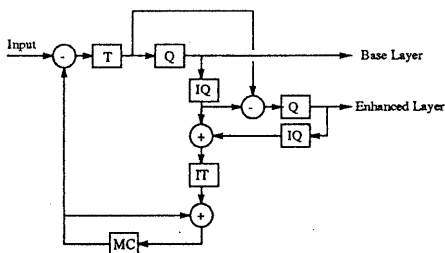


図2 SNR スケーラビリティの構成

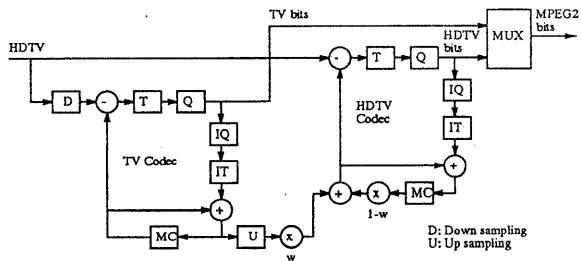


図3 HDTV-TV 空間スケーラビリティの構成

なっているフィールドが奇フィールドであるか、偶フィールドであるかにより予測フィールドが異なる。直前に符号化された2フィールドが参照フィールドとみなされる。

2.4 フレーム / フィールド適応予測, 適応 DCT

フレーム構造の符号化では、フィールド単位の動き補償とフレーム単位の動き補償をマクロブロック単位に切り替えることができる。フレーム構造の符号化では、 16×16 画素のマクロブロックに対して、 16×16 画素を単位として1ベクトル、あるいは各フィールドの 16×8 画素の領域を単位として2ベクトルを用いて動き補償される。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用いることもできる。

フィールド予測では 16×16 画素のブロックサイズと、それを上下に分割した 16×8 画素のサブブロック単位に細かな予測を行う動き補償とを用いることができる。またフレーム構造と同様に、Dual-prime も用いることができる。

フィールド / フレーム適応 DCT はマクロブロック中のデータに対してフィールド別に DCT を行うモードと、交互に混ぜてフレーム構成として DCT を行うモードとを切り替える技術である。

2.5 スケーラビリティ

SNR スケーラビリティは、ビットレートに応じて画品質を段階的に向上させる手法である。ただし、画像サイズは同一である。SNR スケーラビリティの構成を図2に示す。DCT 係数を粗く量子化し、逆量子化した後、係数上での差分値を求め、再び細かな量子化を行う。

空間スケーラビリティは、基本的に CCIR601 と4倍のサイズの HDTV 画像を階層的に符号化する際に用いられる。HDTV-TV 空間スケーラビリティの構成 (エンコーダ) を図3に示す。HDTV-TV 空間スケーラビリティのデコーダは、多重化されたビットストリームを HDTV のデータと TV のデータとに分解する。次に、TV のビットストリームより TV サイズの復号画像を生成する。その後、低域フィルタをかけた拡大して加重し、HDTV の復号画像とから予測を行なう。この予測画像に HDTV データから復号された差分画像を加算して表示する。

2.6 低遅延モードとセルロス対策

通信分野でのアプリケーションでは、符号化-復号遅延が短いことが望まれる。B-picture を用いないことでフレーム並べ替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ることで、バッファサイズを減少させ、遅延を減らすことができる。また、シーンチェンジが発生した場合には、一時的に次の画像を数フレーム期間伝送しないことで対処する。

セルロス対策として、短いサイズのスライスが用いられる。スライスヘッダはユニークワードであり、セル損失後にすばやく回復することができる。受信できなかったスライスのデータの代わりに、前フレー

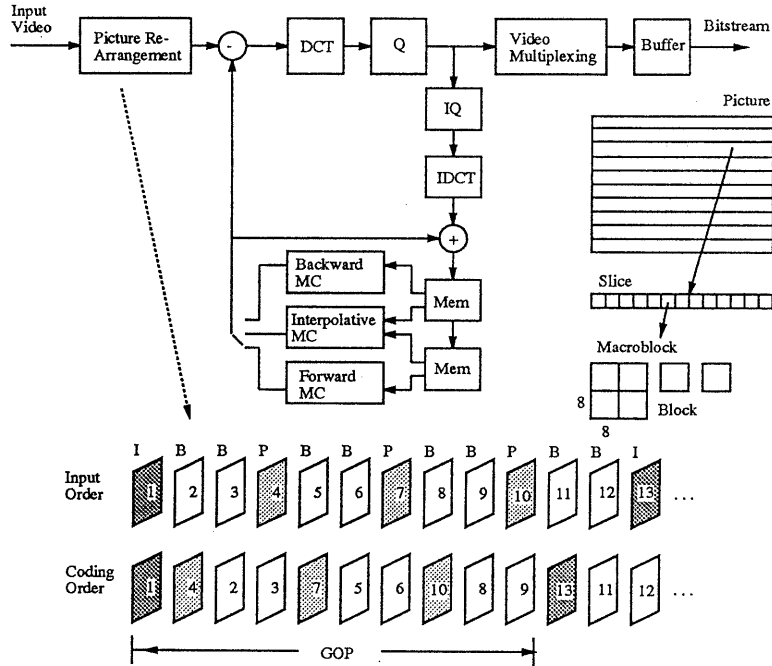


図4 MPEG エンコーダの基本構成

ムの画像を動き補償して欠損部分を埋める。欠損部分の動きベクトルの代わりに、画面上でのマクロブロックの動きベクトルが用いられる。

3 MPEG1 との相違

MPEG1 はマルチメディア PC や WS 上での娯楽用 AV, 教育, ビジネス, ビデオメールなどのアプリケーションが主体である。MPEG1 では 1.5Mbit/s を符号化レートの目標とした。この値は, CD-ROM, DAT などのデジタル蓄積メディア (Digital Storage Media, DSM) のスループットに相当する。

対象とする画像はノンインタレース信号である。基本となる画像フォーマットは SIF (Source Input Format) と呼ばれ, CCIR601 規格の 1/4 サイズである。ただし, 色差信号は輝度信号の水平・垂直方向解像度の 1/2 である。フレーム周期は 24 ~ 30Hz を主に扱う。映画フィルムから NTSC までのビデオが中心である。

MPEG1 勧告は基本的にはビットストリームの定義と復号論理だけを規定する。アプリケーションに応じて画像品質や特殊再生の能力を選択できる。従って基本的にはどのように符号化すればよいかは規定されていない。エンコードに関する自由度が高く, 高品質な符号化画像を得ようとすればノウハウが必要となる。復号論理は正確さのために疑似 C 言語で書かれている。これから符号化法を推測することは困難である。そこで, 典型的な符号化法が付録に付加されている。標準化作業中に用いられたシミュレーションモデルを基にした符号化器の構成を図4に示す。

MPEG1 符号化データはヘッダ情報の中に水平・垂直方向画素数, 画素アスペクト比, 1秒あたりの入力画像フレーム数, 符号化レート, 複号可能なバッファサイズ, 量子化マトリクス, タイムコード, 孤立 GOP を示すフラグ, 符号化順序ナンバー, 符号化フレームタイプ, バッファ状態(遅延量), 動き補償範

表 4 MPEG1 基本パラメータ群 (Constrained Parameters)

Horizontal picture size	≤ 768 pels
Vertical picture size	≤ 576 pels
Picture area	≤ 396 macroblocks
Pixel rate	≤ 396 × 25 macroblocks
Picture rate	≤ 30 Hz
Motion vector range	≤ ± 64 pels (using half pel vectors) (Vector scale code ≤ 4)
Input buffer size	≤ 327680 bits
Bitrate	≤ 1856000 bits/second (Constant bitrate)

囲拡大フラグ、ユーザデータ用フラグなどが含まれている。これらのパラメータは MPEG1 の応用に幅広い自由度を与えている。例えば、水平・垂直方向画素数は最大それぞれ 4096 画素である。画素アスペクト比は CCIR 601 の 525/625 ラインの両方だけでなく VGA モニターにも対応できる。画像レートは 23.976 ~ 60 フレーム / 秒までのうち 8 種類から選ぶことができる。

ところで、画素数やビットレートなどの限られた範囲が、基本パラメータ群として規定されている。これらを表 4 に示す。これは、ハードウェアコストの点から、ある合理的な範囲のビットストリームをデコードできるものを MPEG1 デコーダと呼ぼうとしたものである。

4 むすび

MPEG2 の標準化動向と符号化技術についてビデオを中心に解説した。MPEG は基本的にオーディオビジュアル符号化標準である。ビデオだけでなく、オーディオ符号化の圧縮率、品質に対しても効率の良さが認識されている。また、それらデータの同期と多重化と取り扱うシステムの標準化がなされている。ビデオの復号や表示だけでなく、マルチメディアのアプリケーションには、いずれの符号化標準も重要であり、今後急速に広まると考えられる。

参考文献

- [1] "Information Technology - Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5Mbits/s", ISO/IEC 11172 Part1 System, Part2 Video, Part3 Audio (Aug.1993)
- [2] "Information Technology - Generic Coding of Moving Pictures and Associated Audio", ITU-T Recommendation H.262, ISO/IEC DIS 13818, Part2 Video, Part3 Audio (Mar.1994)