

MPEG-4の放送への応用

MPEG-4とは

MPEG-4はMPEG (Moving Picture Experts Group: 正式には ISO-IEC/JTC1/SC29/WG11) が作成した“Coding of audio-visual objects”と題されるISO/IEC 14496規格¹⁾の通称である。

MPEG-4の活動は1993年に開始され、当初は“Very low bitrate audio-visual coding”というワークアイテムのもと、MPEG-1やMPEG-2より遙かに高効率な符号化による、超高圧縮・超低レート映像通信実現が目標とされた。その後、それに代わってオブジェクトベース処理やエラー耐性といった新たな機能の追加に重点が置かれるようになった。そして、効率改善のツールだけでなく、形状符号化ツールやエラー耐性ツールといった新たな機能実現の技術が導入された。その結果、応用分野もモバイル通信・インターネットなどの超低レート通信だけでなく、イントラネットを用いた情報提供、ビデオカメラ、放送、番組制作など広がりを見せ、ビットレートも数十kbpsからGbpsまで幅広くカバーするようになった。

MPEG-4の基本規格であるバージョン1はすでに規格書が発行され、拡張規格であるバージョン2 (正式にはAMD1: AMenDment1)²⁾も実質的な規格化作業は終了しており、最終承認を経て規格書が発行される予定である。現在はバージョン3以降^{3) ~5)}の審議が進められている。

	概要
Part 1 (System)	映像・音声ストリームの多重、シーン記述、ファイルフォーマット
Part 2 (Visual)	自然動画像、CGの符号化
Part 3 (Audio)	音響、音声符号化
Part 4 (Conformance Testing)	デコーダがMPEG-4準拠であるかどうかの判定手段
Part 5 (Reference Software)	エンコーダ、デコーダのソースプログラム
Part 6 (DMIF: Delivery Multimedia Integration Framework)	配信手段 (通信・放送・蓄積) に依存しないメディア操作のプロトコル

表-1 MPEG-4規格の構成

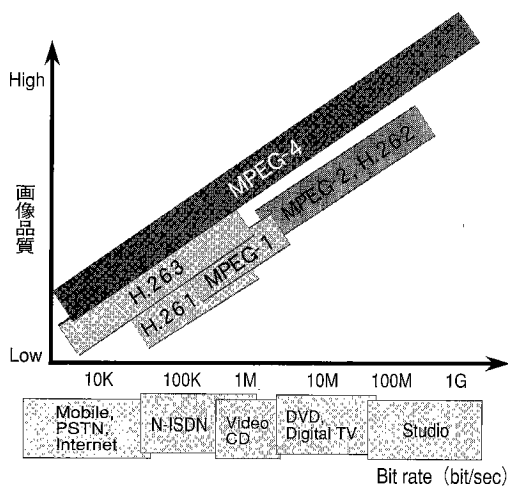
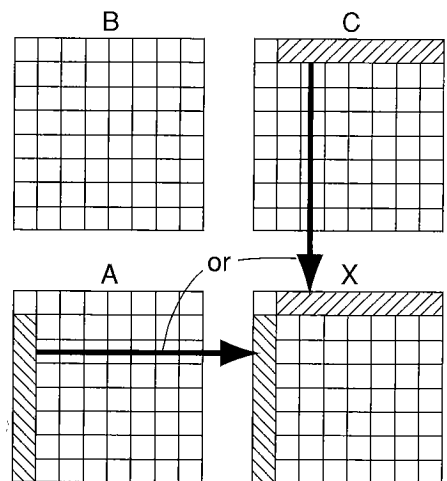


図-1 MPEG-4の適用範囲



(ブロックAあるいはCから斜線部のAC係数の予測を行う)

図-2 AC係数の予測

日本では2000年12月から衛星によるデジタル放送が開始され、本格的なデジタル放送時代を迎える。ここで用いられる符号化はMPEG-2方式であり、当面基幹放送についてはこれが変更されることはない。しかし、インターネットなどさまざまなマルチメディア環境を用いた広い意味での「放送」、あるいは番組制作、素材伝送といった分野においてはMPEG-4技術の導入が期待される。

本稿では、ビジュアルパートを中心にMPEG-4規格の概要を紹介した後、スタジオプロファイルを中心にMPEG-4と放送のかかわりを紹介する。

MPEG-4の概要

<構成>

MPEG-4の規格は表-1の6パートから構成される。MPEG-2との規格骨子上の顕著な違いとして以下の点があげられる。

- * Systemはエレメンタリーストリームの同期や多重化だけでなく、映像・音声の構成要素であるメディアオブジェクトを1つのシーンに合成するための記述法 (BIFS: Binary Format for Scene description) も扱う。

- * VisualだけでなくSystem, Audioにも多数のプロファイル (profile: 機能に基づく規格のサブセット) が存在する。

- * 既存のさまざまな伝送フォーマットが使えるようにシステムストリームと伝送プロトコルの間の橋渡しをするインタフェース仕様としてDMIF (Delivery Multimedia Integration Framework) を規格のパート6に定めた。

以下でパート2であるVisual規格のうち自然動画像 (Video) 符号化を中心に解説を行う^{★1}。

<MPEG-4ビデオ符号化>

MPEG-4ビデオの対応ビットレート、応用の範囲を図-1に示す。さまざまな応用をカバーすべくMPEG-1, MPEG-2の適用範囲を包括した幅広いビットレートを対象とする。また、多くの機能の実現を行うために、多数の技術ツールが導入されている。その結果、規格はかなり複雑かつ大きなものとなっている。約20のプロファイルが定義されており、各々において用途に応じた機能のサブセットを構成している。プロファイルによっては、技術内容がかなり異なるものもあり、MPEG-4はさまざまな規格の集合体とみることもできる。

MPEG-4ビデオ規格の特徴として以下のものがあげられる。

- * 高い符号化効率
- * オブジェクトベース符号化
- * エラー耐性

これらの特徴の実現のために、多くの新技術が導入されている。誌面の都合でこれらすべてを詳細に説明することはできないのでトピックのみを紹介する (これらは文献6), 7) などで解説されている)。

「符号化効率」

より高い符号化効率の実現は、あらゆる応用への普遍的な価値である。MPEG-4も基本的にブロック単位の動き補償, DCT, 可変長符号化といったMPEG-2以前の符号化と同じフレームワークを用いており、これに基づいた改善の工夫が行われた。

動き補償予測については、8×8ブロック動き補償、非

^{★1} パート2の名称もMPEG-2まではVideoと呼ばれていたが、MPEG-4では自然画の動画像の符号化だけでなく、SNHC (Synthetic Natural Hybrid Coding: CGなどの合成画像と自然画像のハイブリッド符号化) も扱うため、これらの総称としてVisualと呼んでいる。

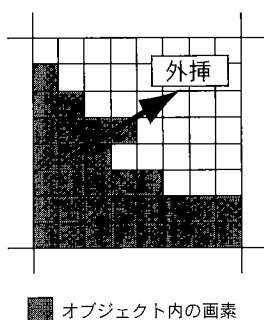


図-3 DCTブロック内のオブジェクト外画素の外挿

バージョン1	バージョン2	審議中
Simple	Advanced Real Time Simple	Simple Studio
Simple Scalable	Core Scalable	Core Studio
Core	Advanced Coding Efficiency	Simple Streaming Video
Main	Advanced Core	Streaming Video
N-bit	Advanced Scalable Texture	3D Mesh
Hybrid	Simple FBA	
Basic Animated Texture		
Scalable Texture		
Simple Facial Animation		

表-2 Visual profiles

制限動き補償，オーラップ動き補償，双方向予測における直接モード，半画素精度動き補償における内挿処理改善などが採用されている。

可変長符号化については，MPEG-2まで採用されていた2次元ハフマン符号化に代わり，3次元符号化が採用された。

この他，フレーム内符号化において，DCT係数のブロック間の相関を利用して符号化を行う手法として，MPEG-2では直流（DC）係数のみ隣接ブロックの直流係数との差分をとっていたが，MPEG-4では交流（AC）係数の一部も予測により差分をとって符号化するAC/DC予測手法が採用されている（図-2参照）。

これらの効果は主として低レートにおいて発揮され，500kbps程度のレートではMPEG-1，2に比べて画質改善が大きい。

「オブジェクトベース符号化」

MPEG-2までは長方形の画面を画像内容に関係なく適当なブロック処理により符号化を行っていたが，MPEG-4では画面を前景と背景といった画面構成要素（オブジェクト）に分割して別個に符号化するオブジェクトベース符号化が導入された。画像を部品化することにより，画面の内容を操作する映像編集が可能となる。また，受信者は復号時に自分の好みに合わせた映像再生が可能となる。

この機能の実現のために新たに形状の符号化が加えられた。形状はオブジェクトの存在領域を示す二値形状（Binary Shape）と前景と背景の混合比を示す多値形状（Gray Scale Shape）で構成される。二値形状符号化にはCAE（Context-based Arithmetic Coding）と呼ばれる算術符号化の一種が用いられている。多値形状の符号化は映像信号と同様に動き補償予測とDCTにより行われる。

オブジェクトは任意の形状をとるため，映像（テクスチャ）の符号化において8×8の正方形ブロックのみを扱

うDCTをそのまま用いることができない。このため，8×8のブロックの一部だけがオブジェクト内となるブロックにおいては，オブジェクト外にあたる画素をオブジェクト内の画素の値を用いて外挿する処理（パディング）を行い，その後DCTを行う（図-3参照：外挿手法は規格外）。

また，前景と背景に分離して符号化することにより，背景にはスプライト符号化を用いることが可能となった。スプライト符号化では，まず複数のフレームからビデオモザイクという処理によりすべての背景画面をカバーする大画面（スプライト）を得る。逆に，各フレームの背景はスプライトの一部に対し拡大・縮小・回転などの幾何変換を施すことにより得ることができる。したがって，スプライトを一度伝送しておくこと，各フレームの背景部分については幾何変換パラメータのみを伝送すればよいこと，大幅な情報圧縮が可能となる。

「エラー耐性」

情報圧縮を行うと，ビットエラーの影響は大きくなる。従来，エラー対策として情報圧縮符号化処理とは別個に，誤り訂正符号の付加や誤り検出を前提とした再送型プロトコルの採用などを行ってきた。しかし，応用によっては，これらの使用が困難なものもある。たとえば，リアルタイム通信や放送型のストリーミングにおいては再送型のプロトコルは一般的に使用できない。MPEG-4では情報圧縮符号化の中でエラー耐性を持たせるための工夫を取り入れている。

画面途中でエラーが発生しても，次の同期回復マーカを検出すれば，それ以降の情報は正しく復号できる。MPEG-4では，同期回復マーカを符号化側で任意の頻度で挿入することができる。同期回復マーカ間でのビット列では後になるほど伝送誤りにより正しく復号されなくなる確率は高くなる。そこで，たとえば動きベクトルなど重要

	VOP type				Shape		Input signal		Error		Efficiency			Scalability				
	I-VOP	P-VOP	B-VOP	Sprite	Binary Shape	Gray Scale Shape	Interface	4:2:2, 4:4:4	10 bit	Error resilience	NEWPREL	AC/DC pred.	4MV, unrestricted MV	GMC	Dynamic resolution conv.	Temporal	Spatial	Fine Granular
Simple	×	×								×		×						
Simple Scalable	×	×	×							×		×				×	×	
Core	×	×	×		×					×		×						
Main	×	×	×	×	×	×	×			×		×						
N-bit	×	×	×		×			×		×		×						
Advanced Real Time Simple	×	×								×	×	×	×	×				
Core Scalable	×	×	×		×					×		×				×	×	
Advanced Coding Efficiency	×	×	×		×	×	×			×		×		×				
Simple Studio	×				×	×	×	×	×									
Core Studio	×	×		×	×	×	×	×	×									
Simple Streaming Video	×	×	×				×			×		×	×					
Streaming Video	×	×	×				×			×		×	×					×

表-3 プロファイルと技術仕様 (抜粋)

な情報を同期回復マーカの直後にまとめて送るというデータパーティショニングを行う。

さらに、リバーシブルVLCの採用により、次の同期回復マーカの直前から逆向きに復号することが可能となっている。これにより伝送エラーにより復号不可能な情報を極力減らすことができる。

「プロファイル」

先に述べたようにMPEG-4の規格は全体としてかなり肥大化している。そのサブセットとなるプロファイルはSNHCや現在審議中のものも含めると表-2に示すように20になる(MPEG-2では7)。当然ながら、先に述べたMPEG-4の特徴的な技術もすべてのプロファイルに含まれているものではない。たとえばシンプルプロファイルではオブジェクトベース符号化は行わない。表-3にプロファイルとそれに含まれている技術仕様の関係の概要を示す。

「実用化の現状」

MPEG-4対応のLSIは数社から開発されており、ソフトウェア製品を含めシステム化されたものの例を表-4に示す。

MPEG-4と放送

元来MPEG-4は超低レートでの伝送を目指していたと

いう事情にも由来して、いわゆる「放送」との関連は薄かった。2000年12月から衛星を使ったハイビジョンを中心とした本格的なデジタル放送が開始されるが、そこでテレビジョン放送に用いられるのはMPEG-2であり、それがMPEG-4に取って代わられることは当分の間ないと予想される。

すでに実質的に規格化が完了しているバージョン2までの15のプロファイルにおいて、インタレース画像を扱えるのはMain profileとAdvanced Coding Efficiency profileの2つだけであるが、まだこれらを実装した製品は開発されていない。

当面はMPEG-2とMPEG-4は共存し、家庭へ高画質番組を届ける手法としてはMPEG-2が用いられ、その他の分野において、MPEG-4の特徴を生かした応用が考えられる。大きく分けて、放送画質は満足しないが、伝送帯域、伝送遅延・エラーの点で悪条件であるメディアを利用した低レート応用、逆に好条件であるメディアを利用した高機能・高画質応用が考えられる。後者については次章のスタジオプロファイルの紹介で述べる。

<低レート応用>

デジタル放送の開始にともない、多彩な情報提供への期待が高まっている。データ放送により、電子番組表やいつでもニュース・天気予報などが提供されるが、さらに番組に関連する動画の提供も考えられる。付加サービスと

MEDIAHALL	日立	社内教育用のノンリニア編集・インターネット配信	最大2Mbps	ソフト
Mobile Motion	東芝	社内教育・訓練, インターネット広告	最大1.5Mbps	ソフト
Windows Media Player	マイクロソフト	インターネットビデオ	28.8~300kbps	ソフト
インターネットビューカム	シャープ	スマートメディアへの記録	最大384kbps	ハード
RadiVision構想	東芝, 富士通, トヨタ, NTV他	モバイル衛星放送サービス	最大256kbps	2001年開始予定

表-4 MPEG-4 応用例

して送られたオブジェクトと番組映像との合成などの、番組をより高度化するサービスも可能である。また、インターネットと連動した放送サービスとして、インターネットを介した番組補足映像、あるいはオンデマンドでの番組ダイジェスト映像などの配信が考えられる。これまでも、RealNetwork社のRealVideo、Apple社のQuickTimeなどインターネットでの映像ストリーミング方式としてはデファクト標準があったが、性能あるいはインターオペラビリティの観点からMPEG-4への統一化が期待できる。

低レートでの放送サービスの伝送路としてSバンド(2.6GHz帯)を用いるモバイル衛星放送サービスも検討されている。RadiVision構想では日本全国で自動車内での受信端末で数十チャンネルの番組を受信できるシステムを2001年を目標に作成しようとしている。これの符号化方式には圧縮効率、エラー耐性に優れるMPEG-4が使用される予定である。

一方、上記のような配信への利用の他、放送局への素材伝送への応用も考えられる。通常の映像素材としては高レートが必要であるが、緊急報道や屋外の固定カメラを用いた定点観測など伝送路への条件が十分整わない場合には低レートでのMPEG-4の利用が有効である。定点観測においてはオブジェクトベース符号化やスプライトの利用による高効率化も期待できる。将来には、家庭からの番組参加など双方向的な放送サービスも想定されるが、このような応用にはオブジェクトベースなどMPEG-4の特徴を生かせると思われる。

MPEG-4のスタジオ応用

<スタジオ応用の動向>

前述のように、MPEG-4は低レート応用が当初中心と

されていたが、オブジェクトベース符号化やスプライトなど放送番組制作に関連の深い技術が含まれている。一方、放送局用の符号化規格としてはMPEG-2 4:2:2プロファイル^{※2}などがあるが、扱える信号の範囲など必ずしも満足なものではなかった。

このような背景のもと、MPEG-4スタジオプロファイルの規格化が開始された。

スタジオプロファイルはVisualパートの規格であるが、Systemパートにおいてもスタジオ応用のための技術提案が導入あるいは検討されている。たとえば、Systemの中でオブジェクトに関連する情報をOCI(Object Content Information)として伝送する機構を持つが、これにより被写体の属性情報や撮影時のカメラパラメータ情報を伝送・保持することができる。また、シーン記述(BIFS)についても編集用途に有用な機能の追加が検討されている。シーン記述のオーサリングのためのテキスト表記法についての検討も開始されたが、スタジオにおけるコンテンツ制作に関連の深い技術である。

<MPEG-4スタジオプロファイル>

スタジオプロファイルは基本的な技術ツールはバージョン2までで提案されていたものを用いているが、全体設計上、以下の点が異なる。

*高品質画像の処理

バージョン2までは色信号フォーマットは4:2:0のものしか扱えなかったが、番組制作には不十分である。そこで、4:2:2や4:4:4の信号も扱えるようにしている。また、1画素当たりの諧調は8ビット表現であったが、10ビット表現まで扱える。さらにスタジオ機器との

^{※2} 符号化においてカラー信号は通常輝度信号と色差信号によって表現される。情報量削減のため色差信号のサンプル数が減らされることがあるが、水平、垂直とも輝度信号の半分のサンプル数となるフォーマットを4:2:0、水平のみ半分となるフォーマットを4:2:2、輝度信号と同じサンプル数で扱われるフォーマットを4:4:4と呼ぶ。

接続を考慮して、信号のダイナミックレンジ（白、黒のレベル）をスタジオ規格に合わせている。画像の解像度は2048×2048まで扱える。

*ハードウェアによる実装への考慮

MPEG-4は全体としてソフトウェアによる実装が主流であったが、上記のような高品質画像のデータ量を処理するにはハードウェアによる実現が必須となる。ハードウェアによる実現には、処理量やデータ量の最大値が問題となる。このため、二値形状符号化における算術符号化は不向きとされはざされた。また、量子化が細かい場合、DCT符号化では原画像データより大きな情報量を発生することもあるので、その場合は原画像データを伝送することとしている。

*伝送路・レートを考慮したツール選択

スタジオ内・間においてはエラーレートはかなり低いので、誤り耐性向上のための技術ツールは含まれていない。また、高ビットレートでは提案されていた符号化効率改善ツールは必ずしもその効果を発揮しない。ハードウェアへの考慮もあり、AC/DC予測をはじめとしたいくつかのツールは含まれていない。

*番組制作環境への適合

一方で、次世代番組制作環境へ有用な技術は仕様を向上して取り入れている。スプライトで用いられる背景用大画面は10年前からシンセビジョン⁸⁾として用いられており、今後もバーチャルスタジオ応用に活用されていく。バージョン2までのスプライト符号化仕様に、処理画面の拡大、デフォーカス処理の追加を行った上で採用される見込みである。また、SMPTE規格に合致したタイムコード、透明度だけでなく奥行き情報など複数枚の画像関連情報(Multiple Auxiliary Components)の伝送も可能となる予定である。

*既存設備との整合性

MPEG-2 4:2:2プロファイル準拠の機材がすでに放送局に導入されており、このフォーマットでの素材映像の蓄積もある。これらの素材を活用し、なおかつ再符号化による画質劣化を避けるためにMPEG-2 4:2:2プロファイルのビットストリームからのロスレス変換を可能としている。具体的には、量子化手法や動き推定手法がMPEG-4とMPEG-2では異なっていたため、これらをMPEG-2に合わせている。

スタジオプロファイルは現在、シンプルスタジオプロファイルとコアスタジオプロファイル⁹⁾から構成されている。シンプルスタジオプロファイルはフレーム内符号化の

みであり、VTRなど主としてスタジオ内での使用を目的としている。コアスタジオプロファイルではフレーム間符号化やスプライトをサポートしており、スタジオ間伝送やバーチャルスタジオ応用を目的としている。

今後、放送局において、効率的な番組制作のためネットワークを利用した制作、映像部品の活用、過去のアーカイブの再利用が進むであろう。また、バーチャルスタジオ³⁾に代表される高度な映像制作も同時に期待されている。これらの環境にスタジオプロファイルは適しており、今後の利用が期待される。

まとめ

本稿では、規格化が進展し、実用化も進みつつあるMPEG-4について、その概要および放送との関連を紹介した。

マルチメディア時代を迎え、映像コンテンツの供給方法、映像素材の伝送方法の多様化、コンテンツ制作の効率化や高度化が今後進むと考えられる。このような環境における共通のプラットフォームとしてMPEG-4は期待される。

しかし、現時点ではMPEG-4の持つ潜在能力を十分発揮されたシステムが開発されているとはいえない。たとえばオブジェクトベース符号化の特徴が活用されていない。1つの理由として、符号化入力となるオブジェクトの抽出方法やオブジェクト間のレート制御、自然な合成法などが規格の外であり、まだ研究段階であることが挙げられる。MPEG-4の持つ潜在能力を十分発揮させるだけの周辺技術の開発が必要である。今後の研究開発、応用面への展開の動向が注目される。

参考文献

- 1) ISO/IEC IS 14496: Information Technology – Coding of Audio-Visual Objects (1999).
- 2) ISO/IEC JTC 1/SC 29/WG 11 N3056: Text of ISO/IEC 14496-2/PDAM1 (1999).
- 3) ISO/IEC JTC 1/SC 29/WG 11 N3146: Text of ISO/IEC 14496-2/PDAM2 (1999).
- 4) ISO/IEC JTC 1/SC 29/WG 11 N3313: Text of ISO/IEC 14496-2/PDAM3 (2000).
- 5) ISO/IEC JTC 1/SC 29/WG 11 N3315: Text of ISO/IEC 14496-2/PDAM4 (2000).
- 6) 三木編: MPEG-4のすべて、工業調査会 (1998).
- 7) 中屋: 新しい動画符号化規格MPEG-4の詳細、インターフェース1月号 (2000).
- 8) Shimoda, et al.: New Chroma-key Imaging Technique with Hi-Vision Background, IEEE Trans. Broadcasting, Vol.35, No.5 (1989).
- 9) ISO/IEC JTC 1/SC 29/WG 11 N3314: Study of Future Extensions to ISO/IEC 14496-2 Studio Profile Amendment PDAM3 (2000).

(平成12年4月27日受付)

³⁾ カメラで撮影した人物などの前景とCG等の背景を合成するシステム。カメラの動きに合わせて合成用背景が変化する。