

# 移動通信における音声／音響・画像符号化技術 — MPEG-4 —

## Speech/Audio and Video Coding for Mobile Communications — MPEG-4 —

移動通信への適用を視野に入れた新しいマルチメディア符号化標準である MPEG-4 の標準化作業が 1999 年 1 月の国際標準化を目指し進められている。ようやく全貌が明らかに成りつつある標準化草案を基に、MPEG-4 が可能にする新たなマルチメディア符号化の世界を紹介する。

A new multimedia coding standard “MPEG-4”, which targets also for mobile communications, is now actively under discussion in the standardization committee in order to complete an international standard on January 1999. This paper describes the outlook of the future multimedia world with MPEG-4 based on the description of the standardization drafts.

大矢 智之 三木 俊雄  
Tomoyuki Ohya Toshio Miki

### まえがき

動画像・音声／音響を含む高能率なマルチメディア符号化の代表的な国際標準化活動として、ISO (International Standardization Organization: 国際標準化機構) / IEC (International Electrotechnical Commission: 国際電気標準会議) JTC1 (Joint Technical Committee 1) の MPEG (Moving Picture Experts Group) がある。MPEG では 1988 年以降、主に高ビットレート領域で高品質をねらいとした標準化 (MPEG-1/2) を実施してきた。1992 年に標準化された MPEG-1 [1] はビデオ CD (CD カラオケ) やマルチメディアパソコン、インターネットなどですでに広く用いられている。また、1994 年に標準化された MPEG-2 [2] はデジタル衛星放送や DVD などで近年急速に普及し始めている。さらに HDTV (High Definition Television) など高品質分野への適用をねらって

MPEG-3 の標準化が検討されたが、MPEG-2 の対応範囲が拡大されたため欠番となっている。

現在、MPEG-1/2 に続く新しいマルチメディア符号化標準として、1999 年 1 月の完成予定で作業が進められているのが MPEG-4 である。MPEG-4 では、符号化効率の向上に加えて、自然画と合成画の統合 (SNHC)、オブジェクト (コンテンツ) 別符号化と操作／加工機能、高度な誤り耐性、各種のスケラビリティなどの新機能を有する汎用符号化の実現を目指しており、コンピュータ／放送／通信など多くの領域で利用できる、汎用的な次世代マルチメディア符号化規格としてまとめつつある (図 1)。移動通信などの無線分野でもマルチメディアを実現することをターゲットの一つとしていることおよび、標準完成時期が 1999 年初頭であることから、次世代移動通信システムにおける有効なアプリケーションであると期待される。

この MPEG-4 標準勧告の全体像を

まとめた CD (Committee Draft: 委員会草案) が 1997 年 10 月に完成する。本稿執筆時点では最終版の発行は行われていないが、これまでに完成したドラフト [3] ~ [12] を基に MPEG-4 の適用領域、標準の規格の概要などについて述べる。

### 標準化活動

#### ■組織

MPEG-4 は現在 ISO/IEC JTC1/SC29 (Sub Committee 29) / WG11 (Working Group 11) において標準化活動が進められている。定期会合は、ほぼ 4 カ月ごとに 1 週間の予定で開催されており、定期会合後は具体的な課題別にアドホックグループを設立し、電子メールによる日常的な討論やアドホックグループ会合の開催などを展開することにより、機動的で活発な活動を行っている。会合への参加者数も WG レベルとしては最大規模の 20 カ国、300 名である。



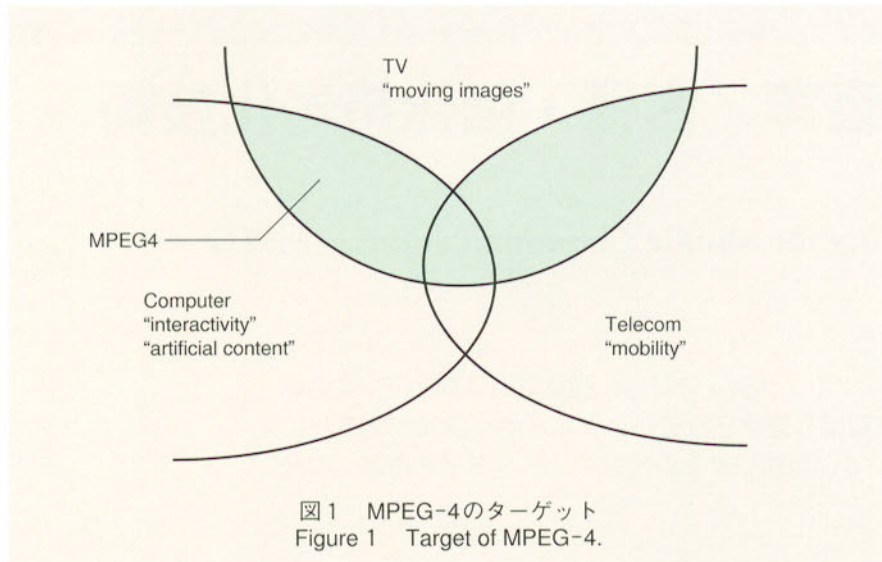


図1 MPEG-4のターゲット  
Figure 1 Target of MPEG-4.

表1 サブグループの構成  
Table 1 Sub-group List.

名 称	議 長	担当業務
Requirements	Rob Koenen	規格への要求条件などの作成
DSM	小暮拓世	デジタル蓄積メディア（サーバ）とクライアントの関係を規定
Delivery	Vahe Balabanian	MPEG-4アプリケーションと蓄積/伝達メディア間の関係を規定
Systems	Olivier Avaro	多重化/出力の構成などの規格の作成
Video	Thomas Sikora	自然動画像符号化規格の作成
Audio	Peter G. Schreiner	自然音響/音声符号化規格の作成
SNHC	Peter K. Doenges	合成画像/音響信号の符号化の符号化規格の作成
Test	Laura Contin	主観品質試験方法の確立、実施
Implementation	Paul Fellows	複雑度（演算量/所要メモリ量など）の評価
Liaison	Barry G. Haskell	外部標準機関などとの連携調整
HoD	L. Chiariglione	MPEG運営に関する一般的事項の調整

WG11が扱う問題は広範囲であるため、表1に示すように11のサブグループが設けられ、規格の各パートの作成および参加国や外部機関との調整を担当している。

国内では、情報処理学会・情報規格調査会傘下にSC29/WG11/MPEG-4小委員会、オーディオ小委員会があり、34社が参加する国内委員会として、MPEG-4標準化に対し日本の意見を反映すべく活動を行っている。

#### ■スケジュール

MPEG-4ではこれまでに、CFP (Call For Proposal) / PPD (Proposal Package Description: 募集要項) / テスト手順文書の発行 (1995.7)、提案募

集とその評価 (1995.11)、アルゴリズムとして必要最小限の機能を集めたVM (Verification Model: 検証モデル) の作成 (1996.1)、WD (Working Draft: 作業草案) の作成 (1996.11)、1年以上にわたる改良技術提案の有効性検証作業であるコア実験とVM/WDの改訂を行ってきた。また、1997年7月には対案として提案された2つのアルゴリズムとの最終的な比較評価試験を行い、現状のVMが最も優れた性能を有することを確認した。

MPEG-4標準化の今後のスケジュールを表2に示す。1997年10月にはCDが発行され、内容がほぼ凍結される。その1年後にはDIS (Draft International Standard: 国際標準草案)

として規格がリリースされ、1999年1月にはIS (International Standard: 国際標準) となる予定である。

## 適用領域

1997年7月のストックホルム会合において、MPEG-4アプリケーションに関する文書[4]が出力された。ここで述べられているMPEG-4の代表的な適用領域と要求される主な機能を表3に示す。なお、この表はまだ暫定的なものであり、今後順次追加されるとともに、含まれるべきツールの規定などの拡張が行われる予定である。また、適用領域の例示がプロファイルの分類や要求条件の設定に影響を与えていくことが予想される。以下、主要な例について概説する。

#### ■リアルタイム（双方向）通信

これは、MPEG-4の当初から想定されてきたTV電話的な適用領域であり、PSTN (Public Services Telephone Network) / ISDN (Integrated Service Digital Network) での利用のみならず、将来の移動通信への適用をも考慮して、低ビットレート/低遅延 (400ms以下) かつ強力な誤り耐性を提供する。

#### ■インターネット/イントラネット

マルチメディアビットストリームをサーバからインターネットあるいはイントラネットを介して検索する利用形態は、WWWへのアクセス状況を見ていると、今後最も急速に伸びていく領域であると期待されている。

しかしながら、現実のインターネット利用者の回線ビットレートは64kbit/sから28.8kbit/s以下がほとんどである。イントラネットの場合でも多くの利用者が同時に利用することもあるために、実際のスループットは時間的に大きく変化する。このような状況で有効なのが画像全体を荒く符号化した階層と、画質を向上させる差分階層



表2 今後の標準化スケジュール  
Table 2 Standardization Schedule.

時 期	会場場所	内 容
1997年10月	Fribourg (スイス)	・ Systems/Visual/Audio/DMIF/Reference_Software CD ・ Conformance Testing WD ・ バージョン2WD
1998年2月	San Jose (米国)	
3月	東京 (日本)	・ 性能検証試験 (誤り耐性)
7月	Dublin (アイルランド)	・ FCD
10月	Princeton (米国)	
12月	Tel Aviv (イスラエル)	・ DIS(実質的なバージョン1標準化完了)→1999.1 IS ・ Conformance Testing CD ・ PDAM (バージョン2のCD化)
1999年3月	未定 (韓国)	
7月	未定	・ Conformance Testing FCD ・ FPDAM (バージョン2のFCD化)
12月	未定 (米国)	・ Conformance Testing DIS ・ DAM (バージョン2のDIS化)

に分けて符号化することにより状況に応じて必要な情報のみを使用する技術であるビットレートスケラビリティ機能である。

加えて、パケットロス耐性や受信側での操作性 (VTRに似た再生/停止/早送り/巻き戻し制御など) も重要な課題として検討されている。

#### ■放送

放送分野ではすでにMPEG-2を用

いた高品質デジタル化が進んでいる。MPEG-4ではさらに新しい機能として、多元・多チャンネルや、オブジェクト/コンテンツの編集機能 (CMの地域入替やマルチウインド表示など) を提供する。

#### ■その他

SNHC (Synthetic/Natural Hybrid Coding : 合成画/自然画統合符号化) 機能 (音声などに同期した顔や体のア

ニメーション、テキスト音声合成など) を利用した仮想会議/仮想プレゼンテーション、ステレオ映像を用いた臨場感のあるTV会議/遠隔協調作業、仮想3D空間を体験し設計作業などを支援する機能、放送/蓄積素材の柔軟な編集/特殊効果機能などへの適用も検討されている。

## 規格の概要

#### ■規格の構成

MPEG-4は、システム (パート1)、ビジュアル (パート2)、オーディオ (パート3)、適合試験 (パート4)、参照ソフトウェア (パート5)、DMIF (Delivery Multimedia Integration Framework) (パート6) から構成される。

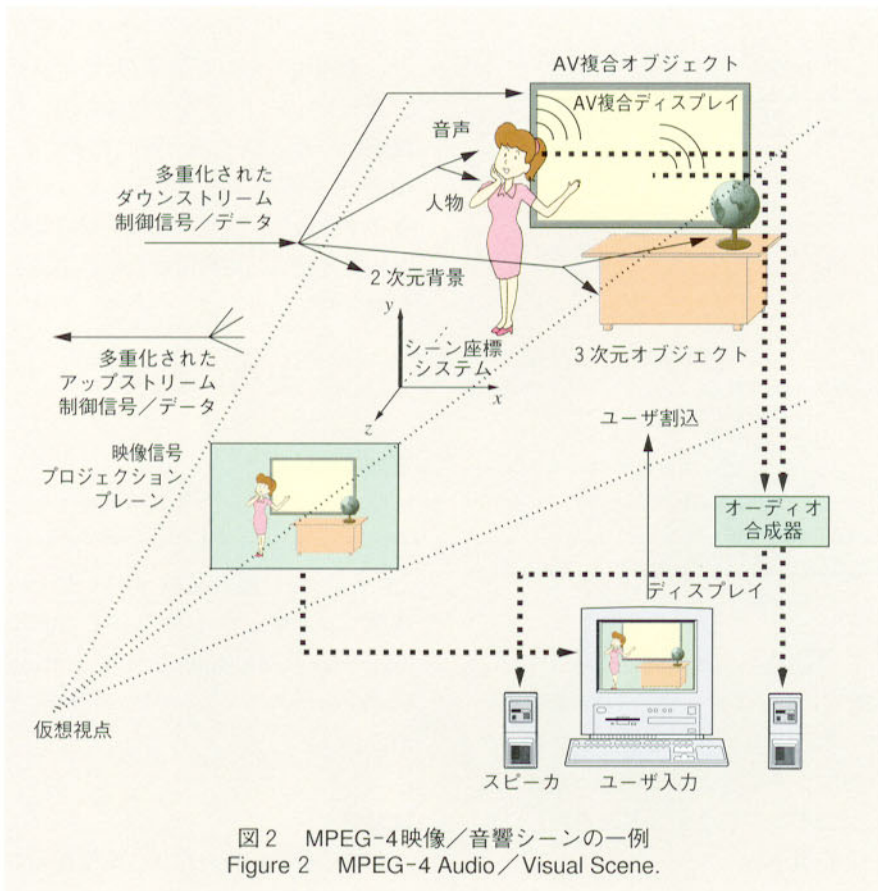
#### ■特徴

従来の動画像符号化は、画像を一つのシーンとして単純に符号化するのを基本としていた。MPEG-4の最大の特徴は、複数のAVOs (Audio/Visual Objects : 人物や背景に代表される構成要素) から成る階層化された映像/

表3 適用領域  
Table 3 Application Area.

適用領域	具体例	要求される機能							形状符号化	その他
		符号化効率	誤り耐性	低遅延	低ビットレートモード	簡易デコーダ	スケラビリティ	スケーラビリティ		
リアルタイム (双方向) 通信	TV電話	○	○	○	○	○	(コンテンツ別)	—	伝達網/I/F 多地点通信	
遠隔監視制御	監視システム	○	○	—	○	○	コンテンツ別 (SNR)	—	伝達網/I/F 12bit解像度	
移動マルチメディア	モバイルマルチメディア 移動データ端末	○	○	○	○	○	(コンテンツ別)	○	伝達網/I/F ランダムアクセス SNHC	
コンテンツ別データ蓄積・検索	電子図書館	—	—	—	—	—	(コンテンツ別)	○	ランダムアクセス 編集ツール	
インター/イントラネット	映像ストリーム提供	○	○	—	○	—	ビットレート	—	伝達網/I/F インタラクティブティ	
遠隔協調作業	CSCW	○	—	○	—	—	—	○	3D/ステレオ画像 SNHC	
放送	衛星/ケーブル放送 多様なモード	○	—	—	—	○	コンテンツ別 時間・空間	○	ランダムアクセス インタレース画像	
DVD	インタラクティブ映画 ゲーム/カラオケなど	○	○	○	○	○	コンテンツ別	2値グレイ	ランダムアクセス 多ストリーム オブジェクト同期	
編集作業	素材編集 特殊効果 (DVE)	—	○	—	○	—	○	—	ロスレス符号化 画像	





音響信号を合成することによって、シーンを表現することにある。各階層構成要素は、自然画像/音響でも合成画像/音響でも構わない。また、シーン全体を示すものでも良いし、一部分であっても構わない。各構成要素の重なり具合や時間的/空間的關係にも大き

な自由度がある。

MPEG-4が表現できる映像/音響シーンの一例[3]を図2に示す。この例では、2次元背景の前に、話者や3次元オブジェクト(動画像、音響を含めた広い意味での符号化対象物)、ディスプレイなどが置かれている。ま

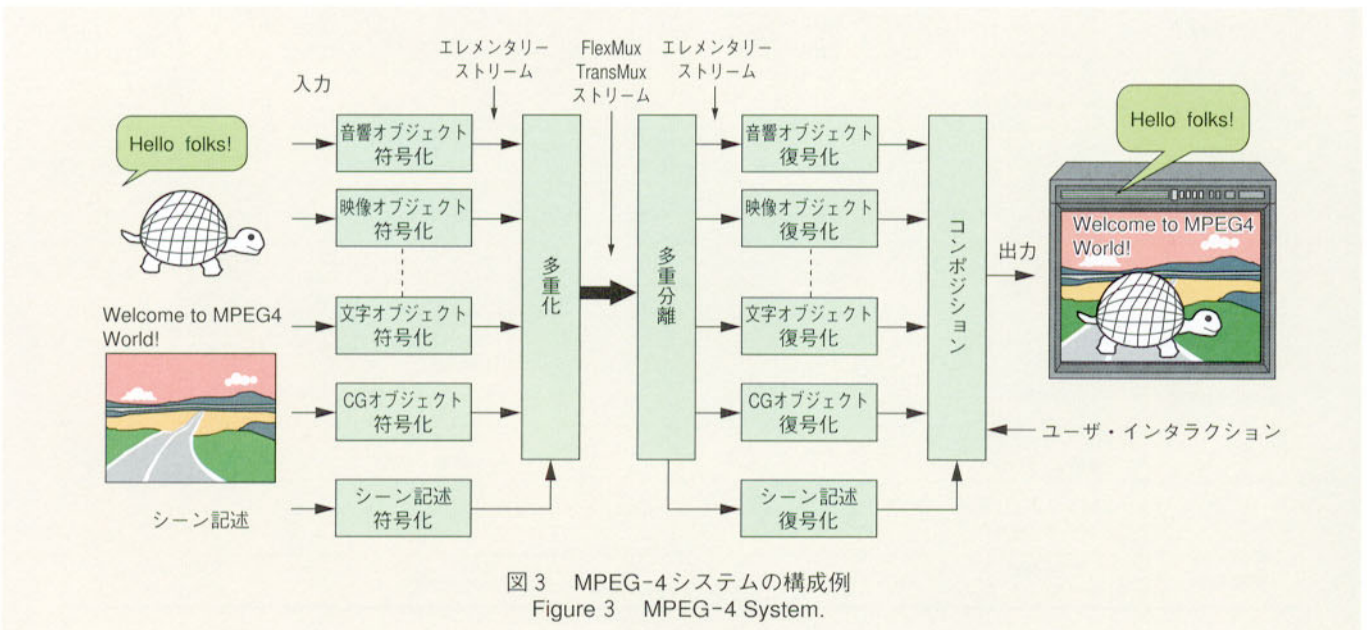
た、これらの階層構成要素がストリームとして送られる様子も示してある。

話者は実際の人物に限らず、顔のアニメーションとテキスト合成音声であっても構わない。背景などもグラフィックスであっても構わない。要は、各階層構成要素ごとに最適な符号化を行った上で多重化されたビットストリームが構築されているのである。受信側では、これらの要素の大きさ/位置/前後関係などを任意に変更し、再構成したり、視点を変更することも可能である。

MPEG-4システムの構成例[13]を図3に示す。シーン内の各オブジェクト入力が最適なアルゴリズムでエレメンタリストリームへ符号化され、多重化されて伝送される様子が見える。また、復号された各オブジェクトはシーン記述に従って、出力へと合成される。この合成作業をコンポジションと呼び、シーン記述とともにMPEG-4の重要な構成要素である。

#### ■MPEG-4システム(パート1)<sup>[9], [16]</sup>

音声と画像符号化の情報の多重化やシーンの合成を行うのがMPEG-4システムの主な機能である。図4はMPEG-4のシステムレイヤモデルを示す。符号化されたデータはアクセス





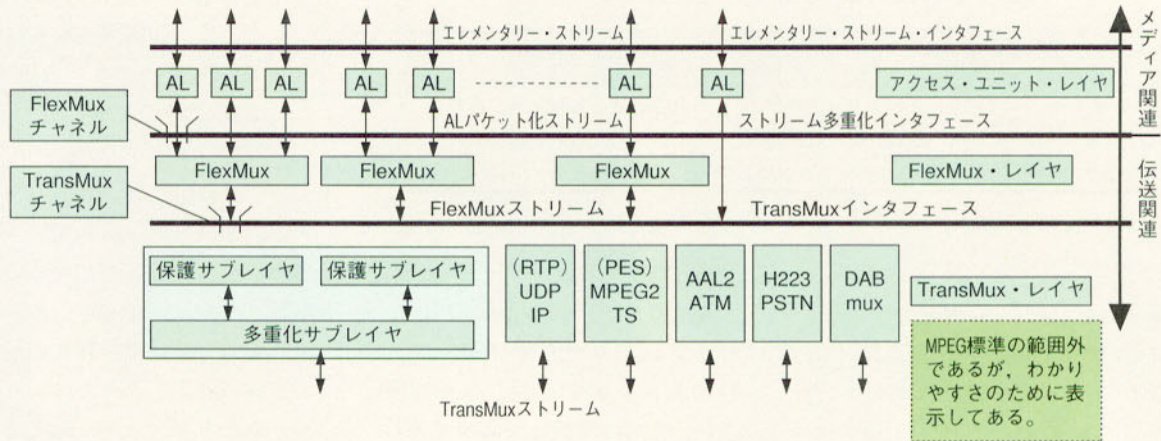


図4 MPEG-4システムレイヤモデル  
Figure 4 MPEG-4 System Layer Model.

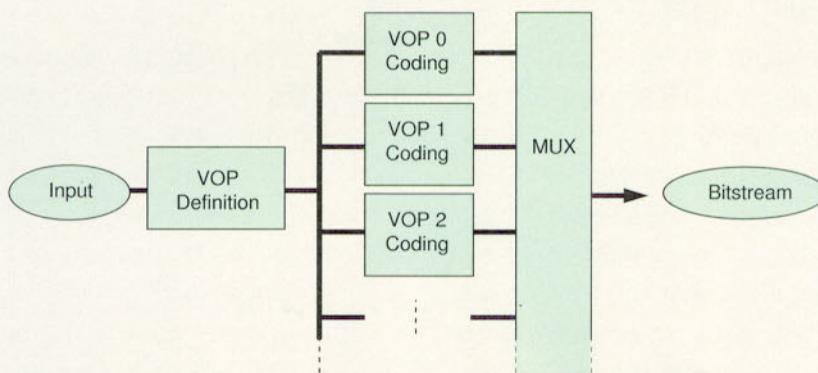


図5 VOP構造  
Figure 5 VOP Structure.

ユニット (AU) と呼ばれる単位 (パケット) ごとにアクセスユニットレイヤに渡される。ここでは、(1)映像/音響フレームやシーン記述命令などのアクセスユニットの判別、(2)タイムスタンプなどによる各オブジェクトの時刻の再生とメディア間同期を行う。

FlexMuxレイヤはMPEGで定義され、同じQoS (Quality of Service: サービス品質) を有するエレメンタリストリームを、比較的低いオーバーヘッドで多重化するツールを持っている。ただし、FlexMuxレイヤツールはオプションであり、図4に示すように、さらに下位のTransMuxレイヤが同様の機能を有する場合にはスキップすること

ができる。また、FlexMuxレイヤは符号誤り耐性が考慮されていないため、無線通信などに適用される場合にもスキップされる。この場合には、MPEG-4システム (多重化) はアクセスユニットレイヤのみとなる。

TransMuxレイヤは、所要QoSを実現するためのツールを提供するが、MPEG-4ではインタフェースの定義のみを行っている。したがって、移動通信などにおける実質的なメディア多重化は、ITU-TのH.223 Annex.A-CなどのようなMPEG外規格を用いて実現されることになる。

シーン記述すなわち各AVOの3次元空間内での配置を記述するためには

BIFS (Binary Format for Scenes) が用いられる。BIFSはVRML (Virtual Reality Modeling Language) を基にしている。これにより、3次元上に配置されたMPEG-4動画などのオブジェクトを任意の視点から観察することが可能となる。VRMLのノードにプッシュ型の実時間ストリームを導入することが可能になることによって仮想現実空間にさらに新たな応用が期待できる。

#### ■ MPEG-4ビジュアル (パート2)<sup>[7] [17]</sup>

MPEG-4動画像符号化は、ITU-T勧告H.263[15]をベースとして様々な改良を加えたものとなっている。従来の符号化と比較して以下のような特徴を持っている。

- (1) オブジェクト単位の符号化
- (2) オブジェクト単位のスケーラビリティ
- (3) 符号化効率の向上
- (4) 誤り耐性の向上
- (5) 合成画/自然画統合符号化
  - (1) オブジェクト単位の符号化

オブジェクト単位のアクセスや対話操作を可能とするために、図5に示すようなVOP (Video Object Plane) という新しい構造が導入されている。符号化するシーンが複数のオブジェクトから構成されている場合には、例えば人



表4 誤り耐性要求条件  
Table 4 Requirements for Error Resilience.

	ランダム 誤り耐性	バースト誤り耐性 (平均バースト長)	誤り回復特性
システムレイヤ	1.0E-2	T.B.D. (T.B.D.)	1 往復遅延時間以内に回復
ビデオレイヤ	1.0E-4	1.0E-3 (10msec)	1 フレーム時間後に回復
オーディオレイヤ	1.0E-4	1.0E-3 (10msec)	80msec以内に回復

物と背景のように別々のVOPとして符号化することができる。それぞれのオブジェクトの形状は別に符号化される。オブジェクトの形状は矩形でも任意形状でもよく、境界の表現も二値(Binary Shape Coding: 二値形状符号化)だけでなく半透明の重なり(Gray-scale Shape Coding: 多値形状符号化)を表現する符号化も用意されている。

#### (2) スケーラビリティ

スケーラビリティはオブジェクトデータへの柔軟なアクセスを提供する機能である。空間スケーラビリティは、画像全体を荒く符号化した階層と画質を向上させる差分階層に分けて符号化する機能で、ユーザの状況に応じて同一のコンテンツから適切な情報のみを利用することが可能となる。また、時間スケーラビリティはコマ落としや特殊速度再生を可能にする機能であり、同じくデータへの柔軟なアクセスを提供する。

#### (3) 符号化効率の向上

ITU-T H.261[19]勧告にいくつかの機能追加を行うことにより符号化効率を2~3倍に向上させたものがITU-T H.263勧告である。MPEG-4ではH.263に取り入れられた機能追加に加え、さらに圧縮効率を高める工夫がなされている。

##### ① ハーフペル動き補償の改良

オブジェクトの移動を符号化する動き補償を、1/2画素単位で行う機能がH.263にもあるが、MPEG-4ではその計算に誤差が集積しないようさらに工夫が加えられた。

##### ② 両方向予測VOP

前後のフレームの情報を用いてあるフレームでのオブジェクトの動きを予測する両方向予測機能が、さらに柔軟な予測ができるよう拡張された。

##### ③ フレーム/フィールド適応予測

現在の放送用素材の大部分が、奇数番目の走査線と偶数番面目の走査線を1フレームおきに交互に記録するインタレース方式を採用している。インタレース画像にも適用可能な動き補償が採用されている。

##### ④ Static Sprite符号化

ある限られた範囲の背景が繰り返しあらわれる画像では、あらかじめ広い範囲の背景のみを符号化しておき、これに射影変換やアフィン変換などの幾何学変換を適用して繰り返し使用すると効率が良い。特にCG (Computer Graphics) 画像などでは劇的な圧縮効率改善を行うことが可能となる。

##### ⑤ DC/AC予測

H.263の符号化は、画像を細かいメッシュ状に区切った一つ一つ(マクロブロック)に動き補償を施した後、DCT (Discrete Cosine Transform) を施して量子化することを基本としている。DCTのDC成分およびAC成分のブロック間相関を取り除くことにより符号化効率が向上する。

##### ⑥ ウェーブレット符号化

H.263をベースとしたDCTによる符号化のほかに、静止画像用の符号化モード(Texture Coding

Mode)として、ウェーブレット変換を用いた符号化も採用されている。低解像度から高解像度までの幅広いスケーラビリティを柔軟に実現できるという特徴を持つ。

#### (4) 誤り耐性の向上

動画像符号化を移動通信へ適用する場合にもっとも問題となるのが、伝送路誤りへの対応である。各レイヤで想定している誤り率を表4に示す。システム側で訂正しきれない誤りが符号化レイヤに上がってきた場合にも、画像の劣化を最小限に押さえるための数々のツールが提供されている。

##### ① 可変長間隔同期方式

VLC (Variable Length Code: 可変長符号) を使用しているビットストリームに誤りが混入すると、それ以降の符号の同期がとれなくなり、バースト的な復号誤りになってしまう。そのため、いくつかのマクロブロックをグループ化しそれぞれの先頭に同期回復マーカを挿入することにより誤りの伝搬を防ぐことができる。

##### ② バイトアラインメント

システムレイヤとの整合上、ビットストリームは8bitの正数倍の基本構造を持つ。8bit境界に満たない場合は最後にスタッフビットを挿入するが、復号時にこのスタッフビットの内容を検証することにより、より確実に復号誤りを検出することが可能となる。

##### ③ 重要情報の多数回送信

フレームの先頭に含まれる重要情報が失われてしまうと、同期回復マーカによってもフレームの途中から正しい復号を行うことはできなくなってしまう。これを防ぐために、重要情報についてはフレームの中で複数回挿入することを可能にする。

##### ④ データパーティショニング

動きベクトル情報とDCT係数をそれぞれまとめて、同期語で区



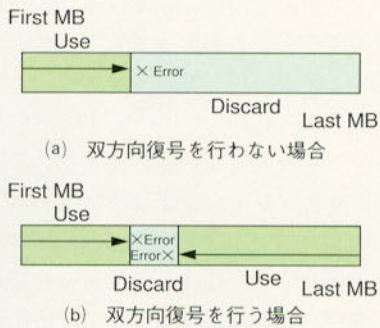


図6 可変長符号の双方向復号  
Figure 6 Reversible Decoding of Variable Length Code.

切ることにより、それぞれのグループで生じた復号誤りが他方に伝搬しなくなり、画像歪みの隠蔽（コンシールメント）を有効に行うことができるようになる。

#### ⑤ 双方向復号可能VLC (Variable Length Code)

逆方向からも瞬時復号可能となるように特別に設計された可変長符号化テーブルを用いる。これにより、図6に示すように符号誤りにより復号不能となった部分についても、次の同期語位置から逆方向に復号することにより多くの情報を復号することが可能となる。

#### (4) 合成画／自然画統合符号化

MPEG-4で符号化可能なオブジェクトは、①自然画ビデオ、②静止画像、③視点変化に対応した画像、④メッシュ、⑤顔アニメーションに分類される。これらの符号化オブジェクトを統合して一つのビデオシーケンスを構成できる点が大きな特徴である。特に③～⑤のように主に人工画像を対象とした符号化を用いることにより大幅なビットレート削減を達成することが可能となる。

##### ・視点変化に対応した画像

ウェーブレット変換を利用してあらかじめ符号化してあるオブジェクトに対し、視点や着目点の変化情報のみを伝送することにより受信側で適切な合成を行う

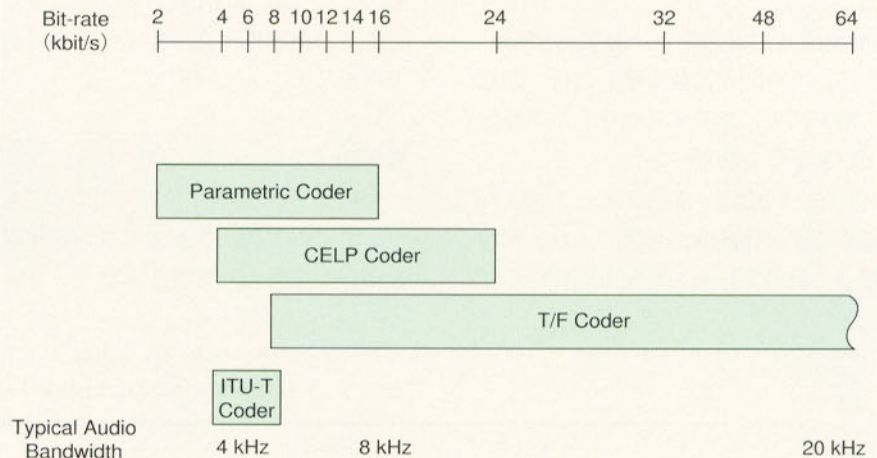


図7 MPEG-4オーディオのターゲット  
Figure 7 Target of MPEG-4 Audio.

機能。

##### ・メッシュ

オブジェクトを2次元または3次元のメッシュで表現し、各メッシュノードの動きのみを伝送する機能。自然画や合成画によるテクスチャマッピングも可能。

##### ・顔アニメーション

顔画像の形状とテクスチャを符号化する機能と、顔の動きや表情を表現する機能から成る。仮想ミーティングなどへの応用も検討されている。

#### ■MPEG-4オーディオ (パート3)<sup>[9]、[10]</sup>

MPEG-4オーディオでは低ビットレート音声符号化から高品質音響まで幅広い符号化対象を扱うと共に、音声合成やMIDIに対応するような柔軟で豊富な機能が盛り込まれている。図7に示すように、ビットレートや帯域に応じて①パラメトリックコーダ、②CELPコーダ、③T/Fコーダの3つの符号化を使用することができる。

##### (1) パラメトリックコーダ

2～4kbit/sの音声符号化を目的としたHVXC (Harmonic Vector Excitation Coding) と6～16kbit/sの楽音符号化を目的としたIL (Individual Line

Coding) の2つの方式から構成される。ともに振幅やピッチ周期、スペクトル形状などのパラメータを抽出して伝送する。HVXCは有声音部分を周波数軸上でハーモニクス表現し、位相情報を削除した後量子化するところに特徴を持つ。また、ILはFFTスペクトルから聴覚マスキングを考慮して重要スペクトルのみを抽出し符号化する方式である。

##### (2) CELP (Code Excited Linear Prediction) コーダ

4～12kbit/sで4kHz帯域の符号化を行うNB-CELP (Narrow Band CELP) と、14～24kbit/sで8kHz帯域の符号化を行うWB-CELP (Wide Band CELP) とから成る。CELP方式は中程度のビットレートで効率の良い符号化を行う方式として日本のPDC方式を含め移動通信用音声符号化としても広く用いられている方式であり、入力音声をスペクトル包絡と音源情報に分離して符号化する特徴を持つ。

##### (3) T/F (Time/Frequency mapping) コーダ

8～40kbit/sで符号化を行うTwin-VQ (Transform-domain Weighted Interleave Vector Quantization) と、24～64kbit/sで符号化を行うMPEG2-



AAC (Advanced Audio Coding) に準拠した方式からなる、ともにMDCT (Modified Discrete Cosine Transform) を用いて時間周波数変換した後、聴覚心理モデルに基づきスペクトル処理および量子化を行う。

Twin-VQは8~40kbit/sまで1kbit/s刻みで符号化速度が可変であり、固定ビット数のフレームを使用しているこ

とから符号誤り耐力も高い。一方MPEG2-AACはMPEG-2の一部としてすでに標準化が完了している方式をほとんどそのまま使用している。

■参照ソフトウェア (パート5)

MPEG-4の標準化の中ではソフトウェアでの実現がアルゴリズムの検討と並行して行われてきた。参照ソフト

ウェアそのものが、MPEG-4ではCPU速度の向上とも相まってリアルタイム動作できる環境が整いつつあり、すでにMPEG-4プレイヤーやブラウザを含むソフトウェア群が多数開発されている。参照ソフトウェアは標準化必須部分のデコーダ、非必須部分のエンコーダともに、研究用途・商業用途を問わず無償での利用が合意され、規格化と

表5 要求条件 (自然映像)  
Table 5 Requirements for Natural Video.

要求項目	要求条件	仕様
オブジェクトベース表示	<ul style="list-style-type: none"> <li>空間/時間関係記述に従って、任意形状オブジェクトを合成/表示、重なり領域/透過性もサポート</li> <li>各オブジェクトを異なるパラメータ/品質/アルゴリズムで符号化</li> </ul>	<ul style="list-style-type: none"> <li>任意形状オブジェクトには、2値形状/2値形状とテクスチャ/グレイレベル (透過性) とテクスチャを含むこと</li> </ul>
コンテンツ	<ul style="list-style-type: none"> <li>すべてのピクセルベースの映像コンテンツをサポートすること</li> </ul>	—
オブジェクトベースビットストリーム操作/編集	<ul style="list-style-type: none"> <li>映像オブジェクト編集機能を有すること</li> </ul>	<ul style="list-style-type: none"> <li>最も粗いオブジェクトアクセス単位が0.5秒</li> <li>テクスチャを復号することなく、形状を復号できること</li> </ul>
オブジェクトベースランダムアクセス	<ul style="list-style-type: none"> <li>任意時刻からの再生、早送りなどの高解像度サポート</li> </ul>	<ul style="list-style-type: none"> <li>任意時刻において、0.5秒以内で対応の品質が提供できる時間ランダムアクセス機能</li> </ul>
品質・忠実性	<ul style="list-style-type: none"> <li>他の規格よりも高い主観品質を実現すること</li> </ul>	<ul style="list-style-type: none"> <li>イントラでも高品質であること</li> </ul>
複数並列データストリームの符号化	<ul style="list-style-type: none"> <li>複数視点からの映像/独立複数映像の結合符号化</li> <li>ステレオ映像に対する冗長度圧縮</li> <li>通常復合とのコンパチビリティ維持</li> </ul>	<ul style="list-style-type: none"> <li>1シーンに対し、最大4つの映像の結合符号化</li> <li>ステレオ映像に対し、MPEG-2MVPと同様の冗長度圧縮</li> </ul>
誤り耐性	<ul style="list-style-type: none"> <li>誤り耐性ツールを提供すること (無線/有線/蓄積)</li> <li>ビットストリームに対し、不均一な誤り保護が行えること</li> <li>誤り隠蔽、耐性、段階的劣化/復旧能力、OFF機能</li> </ul>	<ul style="list-style-type: none"> <li>ランダム誤り<math>10^{-4}</math>、平均長10msのバースト誤り<math>10^{-3}</math>まで使用可能</li> <li>誤り回復時間は、システム復旧後1フレームであること</li> <li>データの優先順位、誤り検出/隠蔽機能を有すること</li> </ul>
オブジェクトベース符号化柔軟性	<ul style="list-style-type: none"> <li>符号化コンテンツ/テクスチャ品質/形状精度の変更ツール</li> </ul>	<ul style="list-style-type: none"> <li>オブジェクト別符号化機能</li> <li>異なる空間/時間品質、SNR、形状精度、複雑度</li> </ul>
オブジェクトベーススケーラビリティ	<ul style="list-style-type: none"> <li>コンテンツのテクスチャ (時間/空間/SNR)、形状 (空間/SNR) のスケーラビリティ</li> </ul>	<ul style="list-style-type: none"> <li>テクスチャ: ベースレイヤ+最大4つの拡張レイヤ (空間/時間/SNR)</li> <li>形状の拡張レイヤ数は未定</li> </ul>
遅延モード	<ul style="list-style-type: none"> <li>各種の遅延モードをサポートすること。例えば、低遅延復号モードや端末間低遅延モード</li> </ul>	<ul style="list-style-type: none"> <li>初期画面で最大0.5秒、通常時で最大150ms (24 kbit/s)</li> <li>復号遅延50msをサポートすること</li> </ul>
フォーマット	<ul style="list-style-type: none"> <li>各種の映像フォーマットをサポートすること</li> </ul>	<ul style="list-style-type: none"> <li>輝度空間分解能: QSFIF/SQCIF, QCIF/QCIF, SIF/CIF, 4*SIF/CIF, CCIR601, 8x8-2048x2048の任意サイズ</li> <li>色空間: モノクロ, Y/Cb/Cr, アルファとの組み合わせ</li> <li>色差空間分解能: 4:0:0, 4:2:0, 4:2:2</li> <li>時間分解能: 最大キャプチャレートまで (通常60fps)</li> <li>ピクセル: コンポーネント当たり最大8bit</li> <li>プログレッシブ/インタレース映像</li> </ul>
ビットレートモード	<ul style="list-style-type: none"> <li>64kbit/s以下 (low), 64~384kbit/s (intermediate), 384k~4Mbit/s (high) において最適化されていること</li> <li>CBR/VBR</li> </ul>	—
複雑度モード	<ul style="list-style-type: none"> <li>各種の複雑度モードをサポートすること</li> </ul>	<ul style="list-style-type: none"> <li>I-486/75MHz/4MBメモリで15fpsの4:2:0 QSFIF/QCIFが実時間復号できる</li> </ul>
静止画	<ul style="list-style-type: none"> <li>静止画の高効率符号化</li> </ul>	<ul style="list-style-type: none"> <li>空間輝度分解能: 8x8-4096x4096の任意サイズ</li> <li>ビットレート: 0.01bit/pixelから</li> <li>空間8-11階層符号化, SNRを含め最大32階層符号化</li> </ul>



同時に公開される予定である。

## ■DMIF (Delivery Multimedia Integration Framework) (パート6)<sup>[10]</sup>

セッション開始時において、ビットレート／符号誤り率／遅延時間などのアプリケーションからのQoS要求に応じて、符号化リソースや伝送ネットワーク／蓄積メディア割当て、レイヤ間インタフェースの定義などを行うのがDMIFである。DMIFは従来のDSM-CC (Digital Storage Media-Command and Control) を拡張定義したものである。

## 要求条件とプロファイル

自然画像に対するMPEG-4の要求条件を表5に示す[5]。「適用領域」で述べた多くの適用領域を包含するために文献[5]ではこれ以外にも非常に多くの要求条件が規定されている。

上述の要求条件をすべて満たすCODECは、用途によっては不必要な機能をも含んでいるために、複雑で大規模なものになってしまうことが心配される。そこで、適用領域をいくつかのクラスタに分類し、各クラスタ内に共通な要求条件を満たすのに必要最小限のツールセットを規定するのが賢明である。この分類をプロファイル[6]と呼ぶ。

MPEG-4ではプロファイルの整理が始まったばかりであり不確定要素も多いが、ここではビジュアルパートについてのみ文献[12]に基づきツールセットを整理し、表6に示す。

## 知的所有権の扱い

### ■特許・著作権

MPEG-4に関する必須特許の取り扱いについては、まだ予備的な検討が始まったばかりであるが、特許問題についてはMPEG-2において苦労した結果、最近になって特許プール制度が発足したことを鑑みると、同様の手法

表6 プロファイル (ビジュアルパート)  
Table 6 Profile (Visual Part).

プロファイル	ツールの名称
Simple Video Object	<ul style="list-style-type: none"> <li>• Intra Mode(I)</li> <li>• Predicted Mode(P)</li> <li>• AC/DC Prediction</li> <li>• Slice再同期</li> <li>• Reversible VLC</li> <li>• Data Partitioning</li> <li>• (Binary Shape)</li> </ul>
Main Video Object	<ul style="list-style-type: none"> <li>• 全Simple Video Object Profileツール</li> <li>• Bi-directional prediction mode(B)</li> <li>• H.263/MPEG-2 Quantization Tables</li> <li>• Advanced Prediction</li> <li>• Static Sprites</li> <li>• Alpha (Gray Scale) Shape Coding</li> </ul>
unknown	<ul style="list-style-type: none"> <li>• Temporal Scalability</li> <li>• Interlaced Content</li> </ul>

で解決される可能性が高いと思われる。

特許プール制度では、数十件の特許をまとめてサブライセンスする機関を設立し、特許保持者がこの機関にサブライセンス権を与えることによって、特許実施者は個別交渉を行うことなく、無差別にライセンスを受けることができる。ただし、MPEG-2においてもその充足率は7割程度と言われており、制度適用外の特許が規格の普及に影響を与える可能性が残っている。

### ■コンテンツの著作権

デジタル符号化規格の普及に伴い、コンテンツ業界からは著作権保護の問題が強く提起されるようになってきた。すなわち、著作権が保護されることが十分に保証されていない規格に対しては、安心して優良なコンテンツを提供することができない、という問題である。

MPEG-4では従来以上にこの問題は深刻である。なぜならばMPEG-4ではオブジェクト別に符号化され、操作／加工することが可能であるからである。したがって、利用されるすべてのオブジェクトごとにIDを付与したり、著作権保護のための管理を行う必要が生じる。これを可能とするために、多少冗長にはなるがビットストリームにIDスロットを規定する方法や、すかし (Watermark) 技術を取り入れ

ることが検討されている[11]。

## むすび

以上簡単に述べたように、MPEG-4は従来にない高機能かつ高能率なマルチメディア符号化規格であり、コンピュータグラフィックスやアニメーションと自然画像が統合された2D／3Dコンテンツが、蓄積／放送のみならずインターネットや移動通信など多くの伝送メディアでインタラクティブに利用することを可能とする画期的な新規格となる見通しである。規格の実現にはまだまだ多くの労力が必要で紆余曲折も予想されるが、マルチメディアの新時代の旗手となることだけは間違いなく、特に次世代の移動通信システムではマルチメディアアプリケーションの中核として広く用いられていくであろうことが期待される。

### 文献

- [1] ISO/IEC 11172-1/2/3, "Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5Mbit/s", International Standard, 1992.
- [2] ISO/IEC 13818-1/2/3, "Information Technology- Generic Coding of Moving Pictures and Associated Audio", International



- Standard, 1994.11.
- [3] ISO/IEC JTC1/SC29/WG11 N1730, "Overview of the MPEG-4 Standard", 1997.7.
- [4] ISO/IEC JTC1/SC29/WG11 N1729, "MPEG-4 Applications Documents", 1997.7.
- [5] ISO/IEC JTC1/SC29/WG11 N1727, "MPEG-4 Requirements Documents", 1997.7.
- [6] ISO/IEC JTC1/SC29/WG11 N1728, "MPEG-4 Profiles Documents", 1997.7.
- [7] ISO/IEC JTC1/SC29/WG11 N1797, "Text of ISO/IEC 14496-2 visual working draft V.4.0", 1997.7.
- [8] ISO/IEC JTC1/SC29/WG11 N1745, "MPEG-4 Audio WD Version 4.0", 1997.7.
- [9] ISO/IEC JTC1/SC29/WG11 N1825, "MPEG-4 systems WD v5.0", 1997.7.
- [10] ISO/IEC JTC1/SC29/WG11 N1763, "DMIF WD V1.0", 1997.7.
- [11] ISO/IEC JTC1/SC29/WG11 N1732, "Management of identification and protection", 1997.7.
- [12] ISO/IEC JTC1/SC29/WG11 N1768, "Translation from Profile Requirements to Profile Specifications", 1997.7.
- [13] 小暮, 妹尾: "国際標準規格 MPEG-4の概略決まる", 日経エレクトロニクス, No.699, 1997.9.22.
- [14] 三木: "MPEG-4の位置付けと規格への要求", 映像メディア学会講習会「見えてきたMPEG-4の全貌」予稿, 1997.10.2.
- [15] ITU-T H.263, "Video Coding for Low Bitrate Communication", 1996.3.
- [16] 栄藤: "MPEG-4システム", 映像メディア学会講習会「見えてきたMPEG-4の全貌」予稿, 1997.10.2.
- [17] 如澤, 堅田, 渡邊: "MPEG-4ビデオ", 映像メディア学会講習会「見えてきたMPEG-4の全貌」予稿, 1997.10.2.
- [18] 金子, 守谷, 中嶌, 西口, 及川, 田中: "MPEG-4オーディオ", 映像メディア学会講習会「見えてきたMPEG-4の全貌」予稿, 1997.10.2.
- [19] ITU-T H.261, "Video Codec for Audiovisual Services at px64 kbit/s", 1993.3.