

MPEG4 活動の現状

三木 俊雄 河原 敏朗 大矢 智之 保谷 早苗

miki@mlab.nttdocomo.co.jp

NTT移動通信網株式会社 研究開発部

〒238 横須賀市武1-2356

MPEG-1/2に続く次期マルチメディア符号化標準規格候補として登場したMPEG-4は、移動通信・コンピュータアニメーション等への適用を想定し、低ビットレート・高圧縮・高機能を主なねらいとしている。さらにMPEG-4は、複数のツール・アルゴリズム・プロファイルから構成される複合体で、MSDLと呼ばれるシンタックスで結合される構成を持っており、従来規格や新技術をも取り込める柔軟な規格となっている。本稿では、MPEG-4標準化の現状と動向、規格の概要について述べるとともに、1次テストの結果を概説する。

MPEG4 Standardisation Activities

Toshio Miki, Toshiro Kawahara, Tomoyuki Ohya, and Sanae Hotani

NTT Mobile Communications Network Inc. R&D Department
1-2356, Take, Yokosuka-shi, Kanagawa, 238 Japan

The standardization of MPEG-4 was started on ISO/IEC succeeding to MPEG-1/2, and considered to be the promising candidate for next generation multimedia coding standard. This standard focuses on low bitrate coding, high compression and additional new functionalities, and its application area covers mobile communications, computer animation, and so on. Furthermore, the MPEG-4 is foreseen to be composed of four different elements: the syntax called 'MSDL', tools, algorithms, and profiles. This structure allows highly flexible systems which can involve existing standards or new technologies developed after this standardization. In this paper, we briefly introduce the MPEG-4 standard overview, and report the status of the standardization including the result of the 1st round test.

1. はじめに

高効率な動画像符号化技術の国際標準化の代表的な活動として、ISOのMPEG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狙いとしたMPEG1/2の標準化を実施し、蓄積/放送/通信など広い分野での利用が期待されている。

MPEG1/2に続く次期規格候補として登場したMPEG4は、低ビットレート・高圧縮・高機能を大きな特徴とし、移動通信などの無線通信の分野においてもマルチメディアを実現すること、各種のスケラビリティ機能やダウンロード機能などの高機能化を図ること、等を主たるターゲットとしている。時期的にみても、ITU-Rにおける次期移動通信方式の国際標準であるFPLMTSの標準化時期と整合しており、本格的な移動マルチメディア時代の到来が近いことを感じさせてくれる。

本稿では、MPEG4の標準化動向と規格の概要について述べるとともに、昨年11月に行われた1次テストの参加状況および結果を概説する。なお、オーディオの1次テストの結果については1/22-24のミュンヘン会合で報告される予定であるので、発表当日には皆様にご紹介できる見通しである。また、略語については脚注を参照されたい。

2. 標準化活動

2.1組織

MPEGはビデオ/オーディオ/テスト等のサブグループに分かれて活動している。MPEG4は当初この中でMPEG4はAOEと呼ばれるサブグループでPPDやテスト手順等の検討が進められていたが、昨年11月のダラス会合で、ビデオ、オーディオのアルゴリズム、ツールに関する検討はそれぞれのサブグループで行われることとなった。シンタクス、SNHCの検討については当面従来のAOEをMPEG4 Integrationと改名し、ここで検討が進められる。

さらに必要に応じて、各サブグループ内にアドホックグループを組織し、電子メールによる日常的な討論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。国際会合への参加者数も常に100名を超える大所帯である。

国内では、情報処理学会・情報規格調査会傘下にSC29/WG11/MPEG-4小委員会があり、34機関が参加する国内委員会としてMPEG4標準化に日本の意見を反映すべく活動を行っている。

2.2スケジュール

MPEG4は2年前のソウル会合で発足し、これまでにCFP/PPD^[1]、テスト手順文書^[2]、MSDL要求条件^[4]を出力した。今後の主な予定は、1次テストで優秀な結果を納めた提案をベースにVMを作成した後、向こう1年間協調して改良を行い、WD/CD/DISを経て、98年11月のIS化を目指している。図1に標準化スケジュール

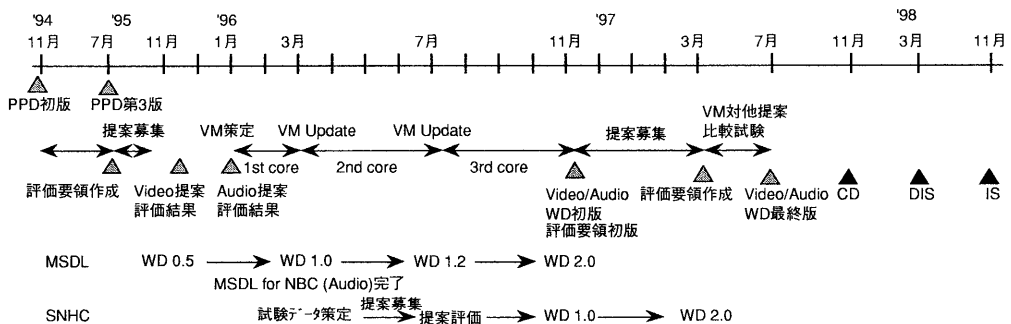


図1 MPEG4標準化スケジュール

ISO	: International Standardisation Organization, 国際標準化機構	CFP	: Call for Proposal
MPEG	: Moving Picture Expert Group, 動画像符号化専門家グループ (ISO/IEC JTC1 SC29/WG11)	CD	: Committee Draft
AOE	: Application and Operational Environment	IS	: International Standard
FPLMTS	: Future Public Land Mobile Telecommunication Systems	VM	: Verification Model
PPD	: Proposal Package description, 募集要項		
WD	: Working Draft		
DIS	: Draft International Standard		
MSDL	: MPEG-4 Syntactic Description Language		
SNHC	: Synthetic/Natural Hybrid Coding		

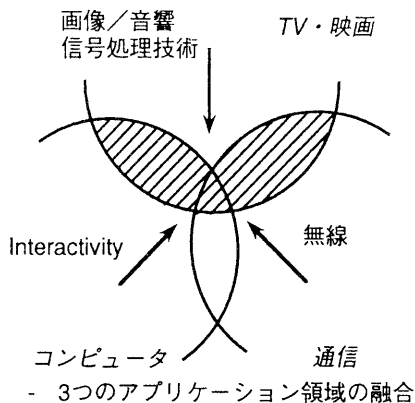


図2 MPEG4 Focus

ルを示す。当初は1次テストの結果に基づいたVM設定とコア実験に続いて2次テストを行い、再度VMを設定する予定であったが、速やかな標準化作業のためには協調フェーズをできるだけ長くとるべきとの判断から2次テストは中止され、本年1月の提案（これに対しては主観評価は行われない）以降、3度のコア実験が連続して行われることとなった。

一方、ITU-T SG15/LBC専門家会合において標準化が進められている低ビットレートビデオフォンの長期目標（H.263/L等）との協調関係も継続して維持されており、基本技術部分に関してはMPEG2とH.262の例のように共通テキスト化される可能性も高い。

2.3適用領域

代表的なMPEG4の適用領域として、PPDには以下の3種類が記載されている。

- (1)オーディオ・ビジュアル・データ・アクセス
 - ・移動網/固定網と接続可能なPDA端末
 - ・携帯形の映像再生/録画端末
 - ・対話型カクログを用いたテレビジョン
- (2)オーディオ・ビジュアル通信
 - ・TV電話
 - ・多地点間TV会議システム
 - ・ビデオ電子メール、ビデオ留守番電話
- (3)遠隔監視・制御
 - ・ビル等の遠隔保安監視
 - ・交通監視システム

後述の1次テストのカテゴリ等も考慮すると、MPEG4の主たる適用領域は、会話型の高機能マルチメディアサービスと、無線を含む高度な通信応用の2つであると言える。

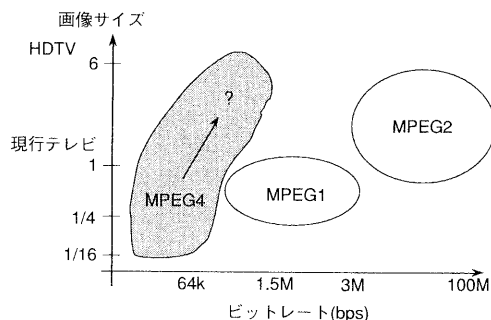
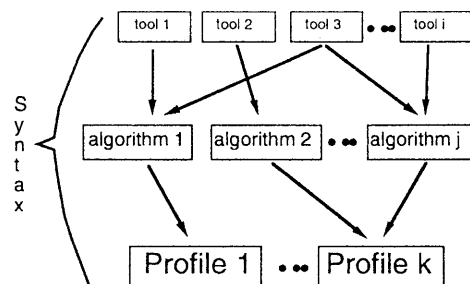


図3 MPEG適用領域



MPEG4 Syntactic Description Language (MSDL)で記述

図4 MPEG4構成

3. 規格概要

3.1フォーカス

MPEG4は当初、64kb/s以下の低ビットレート領域において10倍程度の情報圧縮効率の向上を図り、新しい応用範囲を広げることを狙いとしていた。しかるに、ブレークスルーとなるべき画像圧縮技術の出現が期待できないことから方針の転換を余儀なくされ、豊富な新機能とアプリケーション領域の拡大を目指すこととなった。

アプリケーションとしては、著しい成長が見込まれるコンピュータ/TV・映画/通信の融合領域を想定した(図2)。また、豊富な新機能を実現するためにビットレートの上限も撤廃した(図3)。これらの決断によって、MPEG4はより汎用的で高機能な次期マルチメディア規格候補として再スタートしたのである。

3.2構造

MPEG4は複数のツール、アルゴリズム、プロファイル等のエレメント(オブジェクト)から構成される複合体で、MSDLと呼ばれるシンタクスがこれらの関

表1 MSDLの構成

MSDL-CM	(Configuration Module)
起動時のネゴシエーション用プロトコル	
MSDL-SM	(Structure Module)
オブジェクト(Profile, Tool, algorithm)の名前/番号管理	
MSDL-OCM	(Object Construction Module)
ObjectをLinkして新たなObjectを作成	
MSDL-IDM	(Interface Description Module)
Objectとのデータ受け渡し	
MSDL-CDM	(Core Description Module)
Objectを新たに記述	

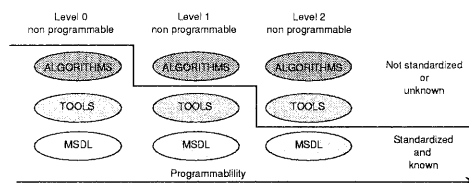


図4 MSDL Programmability Level

係等を記述している。従って、従来の標準規格のように単一のプロファイルやアルゴリズムを規定するのではなく、むしろ新規の要素の出現を予測し、これをいつでも取り込めるような柔軟な構造となっている。また、MPEG1/2やH.261/H.263等の従来規格をも包含する構造であると言える。

ユーザは標準として定められたプロファイルの使用に加え、用途に応じて各種の要素を適切に組み合わせ使用することが可能である。また、必要なツール/アルゴリズムが端末に無い場合には、MSDLが予めサーバ等からダウンロードして、プロファイルを再構築してくれる。MSDLについては3.4章で詳しく述べる。

3.3機能概要

提案されている8つの新機能を以下に示す。

- コンテンツ別の対話/操作機能
(Content-based Interactivity)
 - ・コンテンツ別データアクセス
 - ・コンテンツ別操作、ビット系列直接編集機能
 - ・コンテンツ別スケールリテ
 - ・自然画と合成画のハイリット符号化
 - ・改善されたランダムアクセス機能
- 圧縮
 - ・改善された圧縮効率
 - ・複数のコンテナデータ系列の符号化
- 汎用アクセス機能
 - ・改善された符号誤り耐性

さらに、サイバースペースもMPEG4の新しいスキームとして取り上げられることが予定されており、

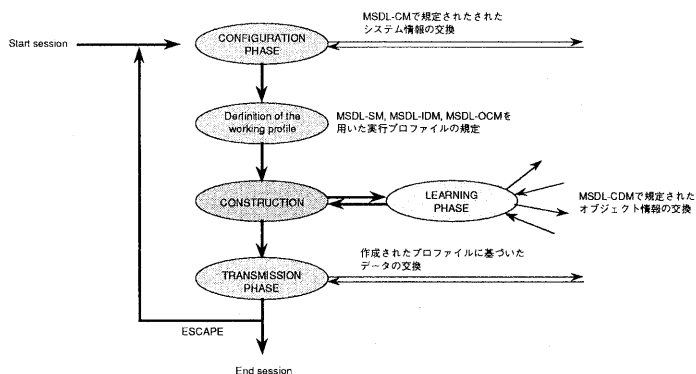


図6 MPEG4セッションの流れ

その柔軟な機能拡張性を活かして、MPEG4自身が一層進化していくものと思われる。

3.4MSDL

MSDLは、プロファイルの記述(エレメント間の関係)、ネゴシエーション、ダウンロード機能、オーディオビジュアルデータの伝送制御、等を行うための言語(シンタックス)である。言語と位置づけられているものの、その守備範囲からシステムの標準とも言えるものである¹⁴⁾。MSDLの構成モジュールを表1に示す。

MSDLは端末のプログラム能力レベルに応じてこれらのモジュールのうち必要なものを用意する。プログラム能力レベルは

- ・レベル0 Non-Programmable
- ・レベル1 Flexible
- ・レベル2 Flexible and Extensible

の3レベルが規定されており、レベル0では標準プロファイルのみ、レベル1では標準ツールを用いた非標準アルゴリズム/プロファイルのサポート、レベル2ではダウンロードによる非標準ツールまでがサポートされる(図6)。

前述の通り、MSDLはシステムの標準に近い性質を持っており、セッションの開始から終了までの動作を

規定している。図6に一般的なセッションにおける処理の流れと、そこで用いられるMSDLモジュールを示す。

具体的な使用言語としては現在C++が用いられているが、これは暫定的なものであり、他にもJAVA等が有力候補として挙げられている。

現時点ではMPEG4の中で唯一WDがバージョン0.1の不完全なものながら設定されており¹⁹⁾、今後の標準化作業はNBC Audioを対象としたものを本年3月までに策定、同時にWD1.0を完成することが決定されている。ただしVMとの関係などは未だ不明確である。

4. 1次テスト概要

4.1ビデオ

コンテンツ別スケーラビリティ、圧縮、符号誤り耐性の3カテゴリについて評価を行う。ビットレートは10~1024kb/sと幅広い。原画像としては、解像度/動き速度別の3種類にステレオ画像とハイブリッド画像の計5クラスが用意された。

符号誤り耐性試験に用いる誤りパターンは、ランダム誤り(BER=10⁻³)、バースト誤り(16~24ms長、3回/10秒間)、ランダム/バースト混合誤りの3種類である。また、誤り回復試験も行われ、1~2秒間の長いバースト誤りからの復旧の程度が評価された。

評価方法は、アルゴリズム提案に対しては主観評価、ツールに対してはエキスパートによる評価が実施された。主観評価法はSS/DSIS/

DSCQSがカテゴリ/クラスに応じて使い分けられる。比較対象のアンカーにはMPEG1とH.263が使用された。

提案数は、コンテンツ別スケーラビリティが15件、圧縮が29件、符号誤り耐性が11件、ツールが67件と非常に多く、各社の本標準への意気込みが伺われるものとなった。表3に提案組織とカテゴリを示す。また提案機関と提案カテゴリを表3に示す

4.2オーディオ

圧縮、ビットレートスケーラビリティ、スピード制

表3 第1次評価提案組織 (ビデオ)

提案組織	提案カテゴリ		
	コンテンツ別スケーラビリティ	圧縮	符号誤り耐性
Columbia U		A	
AT&T	A, B/C	A, B	A, B
Microsoft	A, B/C, E	B, C, E	
UC Berkeley		A	
Ti	A	A, B	A, B
日立		A, B	
JVC		A	
松下	E, B/C/E	B, E	A, B
三菱	A	A, B	
NEC	A	A, B	
NTT		B, C	
NTT DoCoMo			A, B
沖		A, B, C, E	A, B
シャープ	A, B/C/E	A	A
ソニー		C	
東芝	A		A, B
Daewoo		A	
Hyundai		A	
Samsung		A, B	
TU Berlin		C	
HHI	A, B/C/E		
U Hannover	A	A	
Bosch			A, B
Thomson		C	
Philips		B, E	
EPFL	A	A, B	
U Stratchclyde		A	A

表2 第1次評価要領 (ビデオ)

機能	テスト	シーケンスクラス*
Content-based Scalability	Object Scalability	A, B/C, E
	Spatial Scalability	A, B/C/E
	Temporal Scalability	A, B/C/E
Compression	Subjective Quality	A, B, C, E
Universal Access-Robustness in Error Prone Environment	Error Resilience	A, B, C
	Error Recovery	A, B, C

*シーケンスクラス
A:低解像度、動き小
B:中解像度、動き小/低解像度、動き中
C:高解像度、動き中/中解像度、動き中
D:ステレオ画像
E:自然画と人工画の合成

Bitrate: 10, 24, 48, 112, 320, 512 and 1024 kbps

表4 第1次評価要領 (オーディオ)

機能	テスト
Compression	Subjective Quality
Content-based Scalability	Bitrate Scalability
Speed Control	
Universal Access-Robustness in Error Prone Environment	Error Resilience

- Bitrate

2, 6, 16, 24, 49, 64 kbit/s (fixed bitrate)

- テスト素材

ハープシコード、ボーカル、オーケストラ、カスタネット
スピーチ (背景雑音あり/なし、背景音楽あり/なし)

SS : Single Stimulus Method
DSCQS : Double Stimulus Continuous Quality Scale Method

DSIS : Double Stimulus Impairment Scale Method
MOS : Mean Opinion Score

表5 ツール提案組織（オーディオ）

提案組織	内容
ソニーIPC	Pitch Change
ソニーIPC	Large speed change decoding
Samsung	Low complexity encoding/decoding
CCETT	Low delay coding
Alcatel	Coding of periodic waveforms
Univ. Hannover	Pitch change
NTT DoCoMo	PROTECTS error resilient audio coding tool
MAVT	Small step scalability/adaptation to channel conditions
ソニー	64kbps coder with improved frequency/hardware scalability
NIT	Multiple streams as described in N0999
MAVT	Multiple streams as described in N0999

御、符号誤り耐性の4カテゴリについて評価を行う。ビットレートは2~64kb/sである。音源としては、楽音と音声を用いる（表4）。

主観評価法は64kb/s(符号誤り耐性以外)ではITU-R BS-1116法が、64kb/s(誤り耐性)および2~40kb/sではMOSが用いられる。比較対象のアンカーにはMPEG1、MPEG2レイヤ3、G.722(56kb/s)、G.729(8kb/s)、FS1016(4.8kb/s)が使用される。

11月のダラス会合では、1次テストへの提案が締め切られた。ツールについては11件の提案があり、主観評価は行わずデモと技術文書により審議が行われた。表5にツールの提案組織とその内容を示す。スピード制御と符号誤り耐性のカテゴリではPre-screeningテストが行われ、主観評価試験をする提案が絞られた。その結果、提案数は圧縮が36件、ビットレートスケラビリティが10件、スピード制御が5件、符号誤り耐性が7件となった。現在主観評価がすすめられており、1月のミュンヘン会合において結果が報告される予定である。

4.3 1次テスト結果と今後の標準化作業

前述のとおりオーディオの主観評価結果については1月のミュンヘン会合で報告されるため、ここではビデオの結果を概説する。ビデオのアルゴリズム提案に対する主観評価結果では、3カテゴリのすべてにおいて日本の企業からの提案が上位を占めた⁴⁾。

特筆すべきはこれらの符号化アルゴリズムがH.263に準拠したものであること、およびアンカーとして用いられたH.263自身が主観評価で上位となった点である。これにより、1月会合で有望なアルゴリズムが新規提案に提案される可能性はあるものの、今後の標準化作業においてH.263が核として用いられる見通しとなっている。

今後のスケジュールは当初の予定から変更され、2次テストの中止が決定された。今後はすべて協調フェーズの作業となり、1次テストの結果に基づいたVMの設定およびVMへの各種ツールの導入によるコア実験が主要項目となる。MPEG4の適応領域に広さから、現在30項目以上のコア実験が提案されているが、対応するVMの個数は最小限であることが望ましく、項目の統合が期待されている。

5. むすび

以上簡単に述べたように、MPEG4は従来にない高機能かつ高効率なマルチメディア規格であり、移動マルチメディアなど画期的な新世界を開く旗手として注目されている。

<参考文献>

- [1] ISO/IEC JTC1/SC29/WG11 N0997, "MPEG-4 Call for Proposal," July 1995.
- [2] N0998, "MPEG-4 Proposal Package Description(PPD)"
- [3] N0999, "MPEG-4 Testing and Evaluation Procedure Document"
- [4] N1022, "Requirements for the MPEG-4 Syntactic Description Language"
- [5] N1111, "MSDL specification Version 0.1"
- [6] N1056, "Report of the Ad Hoc Group on MPEG4 Video Testing Logistics"