

MPEG-4 動画符号化技術

MPEG-4 Video Coding Techniques

日比慶一*¹ 堅田裕之*¹ 長谷川伸也*¹
Keiichi Hibi Hiroyuki Katata Shinya Hasegawa

青野友子*¹ 伊藤典男*¹ 草尾寛*¹
Tomoko Aono Norio Ito Hiroshi Kusao

Jordi Ribas*² Regis Crinon*² Shawmin Lei*² Ibrahim Sezan*²

要 旨

現在標準化が進められているMPEG-4は、超低ビットレート向けの動画音声符号化方式である。動画では、符号化効率の向上に加えて、オブジェクト単位の符号化、誤り耐性の向上、階層性機能、自然画と合成画の統合等を有する汎用符号化方式を目標としている。本稿では、当社でMPEG-4動画に関連して開発を進めた再同期符号及びデータパーティショニングを用いた誤り耐性技術、部品画像を用いた時間階層性技術、タイリングを取り入れたウェーブレット符号化技術、2次元メッシュ技術、レート制御技術及びMPEG-4システム関連技術について述べる。これらは、MPEG-4バージョン2提案であるタイリング技術を除き、1998年12月に標準化予定のMPEG-4バージョン1の技術となっている。

MPEG-4 under the standardization work is a coding method for video and audio at very low bitrate. The target of MPEG-4 is a generic coding method and it is characterized by the object based coding, the improved error resilience, the scalability function, the synthetic natural and hybrid image coding, in addition to the improved coding efficiency. In this paper, we are going to report on the error resilience with resynchronization marker and data partitioning, the object based temporal scalability, the wavelet coding with tiling technique, the 2D animated mesh coding, the optimized rate control and some techniques in MPEG-4 systems. These techniques

except for the tiling wavelet in MPEG-4 Version 2 are included as the tools in MPEG-4 version 1 which will be standardized on December in 1998.

まえがき

ISO/IEC JTC1 MPEG は高能率なマルチメディア（音声、画像）符号化の規格化を行う国際標準化組織である。記録、通信、放送用符号化としてのMPEG-1およびMPEG-2が収束したのを受け、低ビットレートでの符号化をターゲットとしたMPEG-4が1993年に発足した。活動の初期において、64kbps以下の超低ビットレートにおける高効率符号化をスコープとしていたが、標準化作業中の情勢の変化等から、動画では、符号化効率の向上に加えて、オブジェクト（人物や物体などを表す任意形状の部品画像）単位の符号化、高度な誤り耐性技術、各種の階層性、自然画と合成画の統合、等を有する高度で汎用的な次世代マルチメディア符号化としてまとめつつある。

一部機能の標準化スケジュールの遅れにより、MPEG-4標準は2段階に分けて標準化される。1998年12月に標準化予定のバージョン1は高能率符号化を実現するための基本的な機能を含んでおり、2000年2月に標準化予定のバージョン2はバージョン1に対する付加機能を含んでいる。¹⁾²⁾

本稿ではMPEG-4標準化において当社で開発を行ってきた再同期符号及びデータパーティショニングを用いた誤り耐性技術³⁾、部品画像を用いた時間階層性技術⁴⁾、タイリングを取り入れたウェーブレット符号化技術⁵⁾、2次元メッシュ技術、レート制御技術及びMPEG-4システム関連技術について述べる。

*¹ 技術本部 マルチメディア推進本部 AVC 研究所

*² Sharp Laboratories of America, Inc.

1. 誤り耐性

モバイル環境での移動マルチメディア端末へMPEG-4を適用することを目的として、伝送誤り率の高い場合に、伝送部や多重化部で対処できない残留誤りに耐性を有する映像符号化方式の実現が検討されてきた。

1.1 MPEG-4での誤り耐性の実現

MPEG-4に採用された誤り耐性機能では、**図1**に示すように、主に符号化ビットストリームのシンタックスを工夫して、誤りの影響を時空間的に局在化し、主観的画質の著しい劣化を回避する。

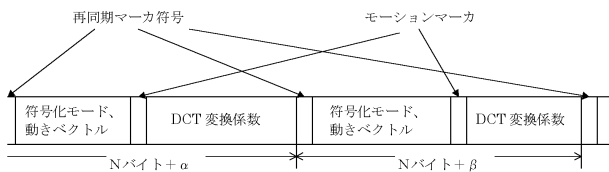


図1 誤り耐性シンタックス

Fig. 1 Bitstream syntax with error resilience.

(1) 再同期マーカ符号

ユニークワードである再同期マーカ符号を適当な間隔で挿入することによって、誤りによる可変長符号の誤復号と、引き続く符号への誤りの波及を防止し、次の再同期マーカ符号以降は、正常に復号できるようにする。

このような同期回復手法に加えて、MPEG-4では以下の手法を適用することで誤り耐性を向上している。

- ・誤りによるエミュレーション等の確率を低減するため、再同期マーカ符号の挿入位置をバイト境界に合わせる（バイトアラインメント）。
- ・バイトアラインメントのために挿入されるスタッフビットが正しく検出できるかを判定することにより、復号誤り検出に利用する。
- ・再同期マーカ符号の挿入間隔を、ビットストリーム上でほぼ等間隔とする。具体的には、**図1**に示すNバイト以上となる直近の符号化データの切れ目にマーカ符号を挿入する。この時、挿入間隔を画面内の時空間的位置で規定する方式に比べて、画面上で変化が激しい領域ほど同期間隔を短くして、誤りの影響範囲を小さくする（**図2**）。

(2) データパーティショニング

データパーティショニングは、一つのビットストリーム上で、符号化属性や、動きベクトルなどの情報と、変換係数情報を分割して伝送する方式である。これらの情報の境界には、境界符号（モーションマーカ）を挿入する（**図1**）。この方式により、誤りを確率的に分散させ、比較的重要である前者の情報が誤る

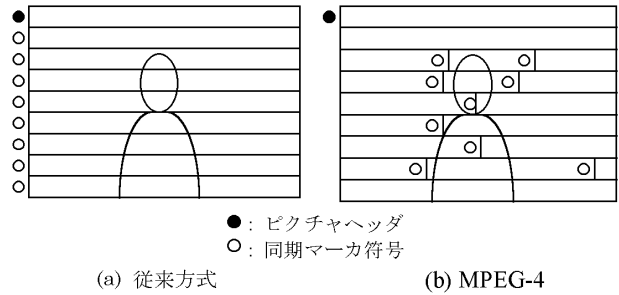


図2 同期マーカ符号の挿入

Fig. 2 Insertion of the synchronization marker code.

確率を低下させると共に、後者の変換係数に誤りが発生した場合でも、前者の情報のみを用いて誤りで劣化した画像を補間（エラーコンシールメント）することができる。また、境界符号までを正常に復号できた場合に誤り無しとみなすなど、誤り検出に用いることもできる。

MPEG-4では、これら以外にも逆方向復号可能なRVLC（Reversible Variable Length Code）、再同期単位毎の画面ヘッダ情報の繰り返し伝送、などの機能によって更に誤り耐性を向上させている。

1.2 シミュレーション実験結果

モバイル環境で典型的と考えられる誤りモデルを利用して、シミュレーション実験による評価を行った。**表1**に復号画像のPSNRを示す。これは、誤りパターンとビットストリームとの相関を除くため、50通りの誤りパターンでの結果の平均である。

誤り耐性機能を用いると、**表1**でのPSNRの向上に加えて、復号誤りによる幾何学模様や色ズレの発生などが抑制され、主観的な復号画像品質が著しく向上する。

表1 復号画像のPSNRの比較

Table 1 Comparison of PSNR of decoded picture.

シーケンス	10msec. パース誤り			
	BER : 10E-2		BER : 10E-3	
	耐性無し	耐性有り	耐性無し	耐性有り
coastguard (48kbit/s)	23.96dB	24.38dB	28.28dB	29.10dB
foreman (48kbit/s)	23.84dB	24.56dB	28.63dB	28.68dB
container (24kbit/s)	26.67dB	30.14dB	32.58dB	33.18dB
M&D (24kbit/s)	27.86dB	30.56dB	33.93dB	34.93dB

1.3 誤り耐性機能の位置付け

MPEG-4に採用された誤り耐性機能は、

- ・オーバーヘッド情報量が小さいこと、
- ・従って、符号化効率低下への影響がごくわずかで

あること、

- ・シンタックス変更のみのため、処理負担が小さいこと、

から、全てのMPEG-4デコーダで対応することが義務付けられる。

誤り発生時のデコーダでの処理方法、即ち、誤りの発見と対処、エラーコンシールメント等の標準での規定範囲外となる画質劣化低減手法などについて、今後更に検討を進める。

2. 時間階層性

MPEG-4における階層性には、画像全体の空間解像度あるいは時間解像度(コマ数)を階層的に向上させるフレーム単位の空間、時間階層性と、画像中の特定の領域(人物や物体などを表す任意形状の領域)のみのコマ数を階層的に向上させるオブジェクト単位の時間階層性(オブジェクト時間階層性)がある。本章では特に、オブジェクト時間階層性について説明する。

2.1 オブジェクト時間階層性

図3はオブジェクト時間階層性を用いた画像符号化の様子を示すものである。図3(a)で、下位レイヤのフレーム1、4、7は、実線の矢印で示すように動き補償予測を用いて符号化されている。図3(b)の上位レイヤでは、下位レイヤで符号化されていないフレーム2、3、5、6の物体領域のみが符号化される。この時、動き補償予測の参照フレームとして、下位レイヤあるいは上位レイヤの既符号化フレームが用いられる。例えば図3で参照フレームは、フレーム2に対しては下位レイヤのフレーム1であり、フレーム3については上位レイヤのフレーム2となる。

復号の際に下位レイヤのデータのみを用いれば、コマ数の低い動画が再生され、上位レイヤまでを復号すれば、物体領域についてのみコマ数の高い動画が得られる。図3の(c)は上位レイヤまでの復号で再生される表示画像が、下位レイヤの復号画像と上位レイヤの復号画像から作成される様子を示している。

2.2 背景合成

上位レイヤの背景に相当する画像を得る処理を、背景合成と呼ぶ。図4は、図3のフレーム2における背景合成の様子を示すものである。図4で左側の領域はフレーム1の物体領域、右側の領域はフレーム4の物体領域である。この時、合成されるフレームの画素値は以下のようにして求められる。

- ・フレーム1の物体領域内かつフレーム4の物体領域外(図4の斜線部分)に対しては、フレーム4

の画素値をコピー。

- ・フレーム4の物体領域内かつフレーム1の物体領域外(図4の網点部分)に対してはフレーム1の画素値をコピー。
- ・フレーム1の物体領域外及び4の物体領域外については以前の低位レイヤフレーム(フレーム1)の画素値をコピー。
- ・上記以外の領域の画素値は、既にコピーされた画素値を用いて補間する。

図5に背景合成を行なう場合(a)と行わない場合(b)の実験結果を示す。(b)では背景として時間的に前の低位レイヤフレームを用いている。上位レイヤの人物が大きく動いているため、(b)では人物が2重となる歪みが生じるが、(a)ではこの歪みが発生しないことがわかる。

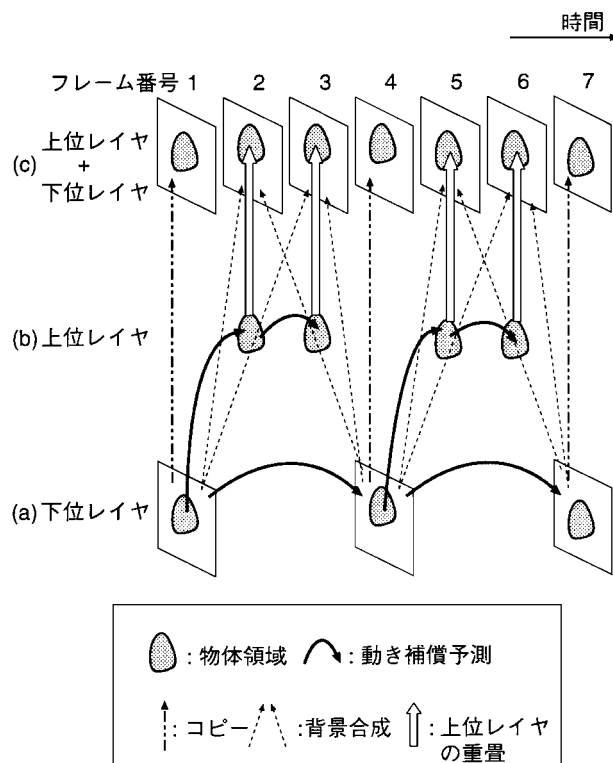


図3 オブジェクト時間階層性による画像符号化
Fig. 3 Video coding with object based temporal scalability.

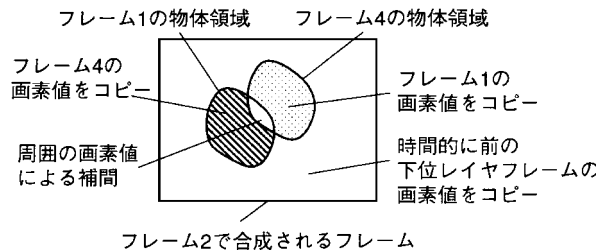


図4 背景合成
Fig. 4 Background composition.



(a) 背景合成を行う場合 (b) 背景合成を行わない場合

図5 背景合成の結果

Fig. 5 Result of background composition.

3. ウェーブレット

MPEG-4には、動き補償予測とDCTを用いた動画符号化とは別に、ウェーブレット変換を用いた静止画像の符号化モード(Texture Coding Mode)を備えている。このモードは、特にCG画像と自然画像の合成を扱うSNHC(Synthetic and Natural Hybrid Coding)グループにおいて、高解像度から低解像度までの様々な空間解像度を備えた(空間解像度の階層性)高画質の符号化方式が望まれた結果、静止画像用の符号化として採用されたものである。また、ウェーブレット符号化は、空間解像度の階層性に加えて、DCTで問題となるブロック歪が生じないという特徴も有する。

3.1 ウェーブレット符号化

始めにMPEG-4で用いられているウェーブレット符号化の基本的なブロック図を図6に示す。図中のウェーブレット変換部では、まず水平方向・垂直方向のサブバンド分割によって入力画像が4つのサブバンドに分割され、このうち水平方向・垂直方向とも低域のサブバンドについてのみ同様なサブバンド分割を繰り返して、図7に示すサブバンド分割データを得る。ウェーブレット変換後は、最低域サブバンドとそれ以外の高域サブバンドで処理が異なり、図6に示すように、最低域サブバンドの係数は量子化、DC予測、適応算術符号化され、高域サブバンドの係数は量子化、ゼロツリースキャン、適応算術符号化されて符号化データが作られる。

3.2 タイリング技術

ウェーブレット符号化は、符号化効率が高い、階層

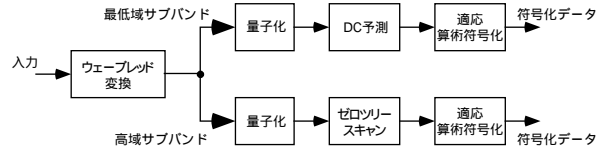


図6 ウェーブレット符号化ブロック図

Fig. 6 Block diagram of wavelet coding.

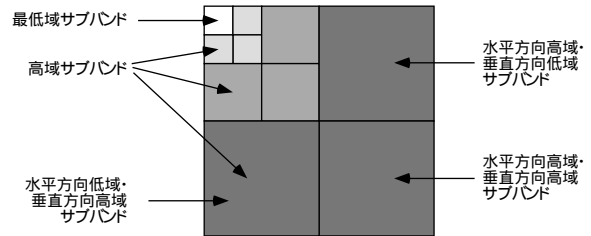


図7 サブバンド分割

Fig. 7 Sub-band decomposition.

性を有している等の優れた特徴がある一方、ブロック単位ではなくフレーム全体を単位として処理を行うことから、必要メモリ量が多い、画像内の任意の箇所へのランダムアクセス性が低い、といった問題点もあった。しかし、原画像を複数のタイルに分割してウェーブレット変換を施す手法を適用することで、ここにあげた問題点を解決することができる。

タイリングの方法には、ウェーブレット変換時の処理によって2つの形態が考えられる。

(1) 各タイルを独立の画像と見なし、タイルの周辺部では折り返し処理を施すことで、それぞれのタイルを別々にウェーブレット変換する。

(2) 図8に示すように、タイル内部の画素に加えて、サブバンド分割に必要な周囲画素も含めた広い範囲に対してウェーブレット変換を施す。ウェーブレット変換後は、タイル内部の画素に対応するウェーブレット係数のみを符号化の対象とし、量子化以降の処理を行う。

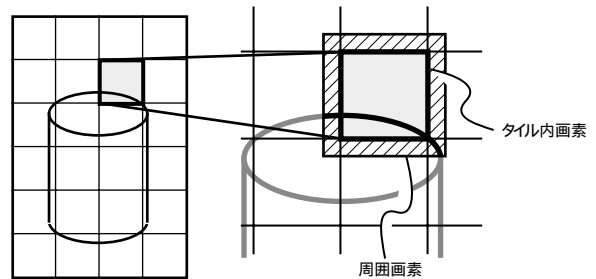


図8 方式(2)のタイリング

Fig. 8 Tiling of method (2).

方式(1)では、各タイルが完全に独立であるため、必要なメモリ量が大幅に削減され、処理も単純である

反面、タイル境界にブロックノイズが発生して画質を劣化させる問題点がある。方式(2)は方式(1)よりは必要メモリ量が増加するが、ブロックノイズの発生はない。

一つのタイルに相当する部分画像の復号にかかる処理時間は、例えば方式(1)のタイリングを使用した場合とタイリング無しの場合を比較すると、方式(1)では約65分の1に復号時間が削減されることがシミュレーション結果から得られている。

タイリング技術を用いることで、全体の符号化データから、任意の部分領域かつ任意の解像度に対応した符号化データを直接取り出し、復号あるいは伝送することが可能となる。

4. 2次元アニメティッドメッシュ符号化

MPEG-4のSNHCグループでは、2次元三角形メッシュを用いたアニメーション機能も検討された。2次元三角形メッシュは、2次元画像フレームを三角形パッチの集まりで表したもので、これらのパッチには静止画のテクスチャがマッピングされる。三角形パッチの頂点をノード、ノード間を結ぶ直線をエッジと呼ぶ。

ダイナミック2次元メッシュは2次元三角形メッシュの時間的なシーケンスからなり、各メッシュは同一のトポロジー(構造)を持つが、ノードの位置はメッシュ毎に異なっている。

従って、ダイナミック2次元メッシュは、初期の2次元メッシュの配置とそれに続くメッシュにおける各ノードの動ベクトルによって規定される。ここで、動ベクトルは一つ前のメッシュのノードから現在のメッシュのノードを指すベクトルである。即ち、ダイナミック2次元メッシュは、連続する画像フレームにおける三角形パッチの形状変化として表され、この三角形パッチの変形に合わせてマッピングされるテクスチャを変形することで2次元アニメーションを生成する。これは、例えば旗の静止画から波打つ旗の動画像を作成するのに用いられる。

5. レート制御

ここでは、MPEG-4の符号化モデルVM8(Verification Model version-8)にて推奨方式として採用された符号量制御について説明する。

符号量制御の目的は、画質を維持しつつ通信路のレートに合わせて圧縮データを生成し、またエンコーダバッファの制限を満たすように符号化パラメータを選択することである。本手法は、バッファ占有量(buffer fullness)を正常に保ちながら、与えられた目

標ビットレートで最高の画質を得る最適化手法に基づいており、低遅延の符号化において、VM7で用いられていた従来法より優れている。図9に示す両者のバッファ占有量より、本手法の有効性が確認できる。図9で、実線はVM8の提案手法、破線はVM7の従来法、点線は符号化バッファの容量である。

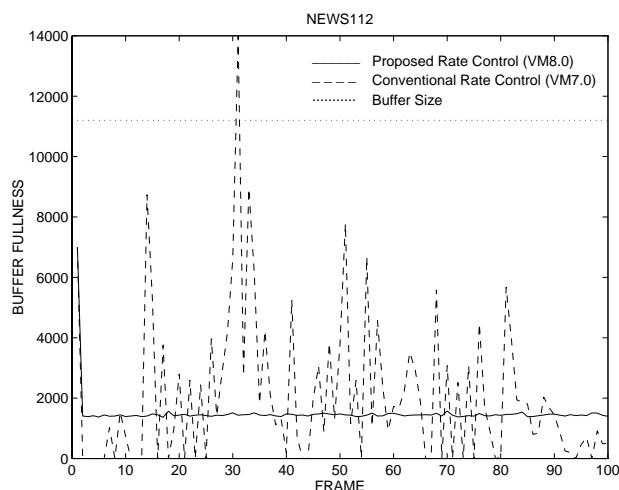


図9 バッファ占有率の比較(テスト画像: News 112kbps)
Fig. 9 Comparison of buffer fullness after encoding each frame of "News" at 112 kbps.

このように、従来法ではバッファ占有量が大きく変動するのに対し、本手法では変動が少なく、占有量がバッファ容量を越えるオーバーフローやバッファが空になるアンダーフローが発生しない。

6. MPEG-4 システム

MPEG-4システムは、映像、音声などの複数のメディアと、これらのメディアに含まれる複数のオブジェクトの同期、多重化を規定している。また、これに加え、シーンを構成するオブジェクトをどのように配置するかを記述するシーン記述方法も規定している。

6.1 MPEG-4ストリームのMPEG-2トランスポートでの伝送

MPEG-4ストリームを送信する際には、ITU-T勧告H.223 Annex A/B/Cや、UDP/IPなど、実際の環境、目的に応じた伝送プロトコルが選択される。

それらの一つとして、MPEG-4バージョン1のシステムでは、MPEG-4ストリームのMPEG-2 DSM-CCのダウンロードプロトコルへのエンカプレーション方法が規定されている。これによりMPEG-2 DSM-CCダウンロードプロトコルを用いるATSC DTV放送やインタラクティブデータサービスとの互換性が保証され、今後のMPEG-4放送の展開において、DTVにおけ

るインタラクティブな応用との整合が容易になる。

6・2 MPEG-7ストリームを参照するメカニズム

MPEG-4でのオブジェクト単位の符号化に加えて、オブジェクトに対して意味、内容に関する属性記述を行なうことが、MPEG-7として検討される予定である。

MPEG-4のシステムでは、いくつかのMPEGストリームを参照できるような規定を含めることが予定されている。MPEG-4のオブジェクト識別子は、MPEG-4ストリームのネットワーク内の位置と、ストリームを受けべきデコーダの種類を特定する。現在、MPEG-4のオブジェクト識別子が複数のMPEG-7ストリームを参照することも可能となっている。MPEG-7ストリームとAVストリームとの関連付けは、同一のオブジェクト識別子で実現してもよく、また、オブジェクトの組み合わせ法を示すMPEG-4シーン記述に、MPEG-7オブジェクト識別子を参照する別の子ノードを追加してもよい。

これらにより、将来MPEG-7が標準化された時にも、現在のMPEG-4システム規格の枠組を変更するこ

となく、AVストリームに意味、内容を記述する情報を追加するというMPEG-7の機能を利用することができる。(図10)

むすび

MPEG-4の開発では、技術考案に加えて、約3ヶ月毎に開催される国際標準化委員会への対応にも多くの時間を割くことが必要である。その標準化活動もMPEG-4バージョン1はほぼ収束しつつあり、今後は引き続きMPEG-4バージョン2に向けての開発を進めていく。

またMPEG-4バージョン1が固まったことから、インターネットやモバイル画像通信分野などでMPEG-4応用商品の登場が期待されている。MPEG-4はマルチメディアの新たなインフラとなり得るものであり、国際標準化活動とあわせて、新規商品創出に向けての取り組みも行っていく。

謝辞

本開発を進めるにあたり、日・米研究所間のグローバルな連携に御尽力頂いた方々をはじめ、御指導ならびに御協力頂いた関係各位に深謝致します。

参考文献

- 1) “Information Technology coding of audio-visual objects: VISUAL”, committee draft, ISO/IEC JTC1/SC29/WG11 N2202, (1998).
- 2) H.Katata, N.Ito and H. Kusao, “Temporal-Scalable Coding Based on Image Content”, IEEE Trans. on CSVT, Vol.7, No.1, (1997).
- 3) “小特集 マルチメディアを目指すMPEG-4標準化動向”, 映像情報メディア学会誌, Vol.51 No.12, (1997).
- 4) 長谷川, 伊藤, 青野, “タイリングを使用したウェーブレット変換による静止画像圧縮の検討”, 信学会ソサイエティ大会, D-11-19, (1998).
- 5) 内海, 佐藤, 江間, 日比, “階層化による誤り耐性を備えた映像符号化方式”, 1996年画像符号化シンポジウム, P-4.4, (1996).

(1998年6月10日受理)

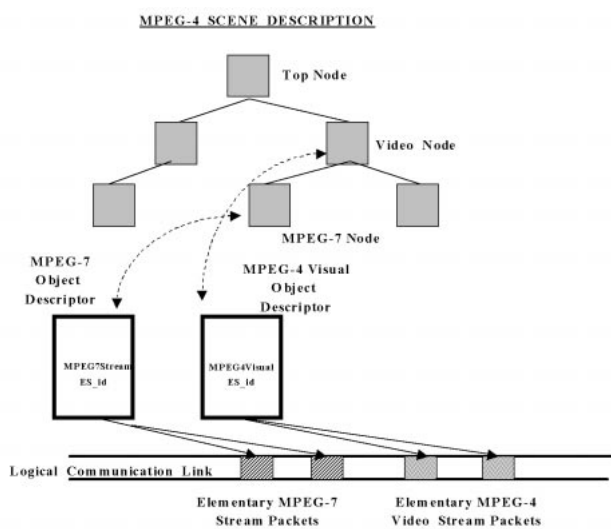


図10 画像や楽音のオブジェクトからMPEG-7ストリームを参照する様子

Fig. 10 Description of the mechanism for referencing by an audiovisual object to its associated MPEG-7 stream.