

# 解説

## 次世代映像メディアの符号化・ 伝送技術の最新動向

Latest Status of Compression Coding and Transmission Technology  
for Next Generation Video Applications

内藤 整 松本修一

### A bstract

近代の映像メディアの発展は主に高解像度化に象徴され、これはBSデジタル放送や地上デジタル放送の開始により、HDTVへ急速にシフトした点からも明らかである。本稿では、これに続く次世代映像メディアとして、3D映像、自由視点映像、超高精細映像に着目し、その基本原理と信号フォーマットについて導入を行う。次に、これら新規映像メディアを、放送やインターネットのインフラを介して扱う上で重要となる圧縮符号化技術、並びに伝送技術の最新動向について解説を行い、実用化に向けた実装アプローチに関する最近のトレンドを紹介する。

キーワード：3D映像、自由視点映像、超高精細映像、圧縮符号化、デジタルシネマ

#### 1. はじめに

地上デジタル放送の開始に伴い、最も身近な映像メディアであるテレビ放送の映像フォーマットがHDTV<sup>(用語)</sup>に統一される方向にある。映像メディアの高度化はこれにとどまらず、今後も進化していくものと予想されるが、大きな流れとしては、高解像度と高臨場感の追求に二分される。本稿では、次世代の映像メディアについて、アプリケーションの位置付けを明確にした上で、これらに特化した圧縮符号化技術、並びに伝送技術の導入を行い、最後に実用化に向けた現状の取組みについて解説を行う。2.では次世代映像メディアの導入を行う。3.では2.で述べた映像アプリケーションごとに、圧縮符号化技術及び伝送技術の最新動向を説明する。4.では次世代映像メディアに対応したコーデック<sup>(用語)</sup>システムの実用化に関連し実装アプローチのトレンドについて紹介する。

#### 2. 次世代映像アプリケーション

本章では次世代映像アプリケーションの代表例として、3D映像、自由視点映像、超高精細映像についての基本原理及び信号フォーマットの解説を行う。

##### 2.1 3D映像

3D映像は、スポーツシーンをはじめとする、立体視により臨場感を味わえるコンテンツを主な対象とし、イベント会場での中継表示、更には一般家庭向け放送と、その利用形態は今後ますます広がっていくものと思われる。

人間は両眼視差により、とらえた物体ごとの奥行き位置を推定し、立体感のある像として頭の中に描いている。3D映像の典型例として立体ハイビジョンについて説明する。立体ハイビジョンは上述の立体視の仕組みに着目し、左右の眼に相当する2台のハイビジョンカメラで撮影した映像を、それぞれ左眼、右眼に対して別個に表示することにより、自然視している際と全く同様に立体感のある像を頭の中に形成する。二つの映像を左右の目に対して別個に表示する方法としてはメガネなしによる方式と、偏光メガネ式とに大別される。メガネなしによる方式としては、レンチキュラレンズ方式とイメージスプリッタ方式が有名であり、共に製品化例が存在する<sup>(1)</sup>。レンチキュラレンズ方式は、スクリーンを覆うレンチキュラ板(かまぼこ状の複眼レンズを横に連ねたもの)

内藤 整 正員 (株)KDDI 研究所映像通信グループ

E-mail sei@kddilabs.jp

松本修一 正員 (株)KDDI 研究所

E-mail matsumoto@kddilabs.jp

Sei NAITO, Member (Visual Communications Laboratory, KDDI R&D Laboratories Inc., Fujimino-shi, 356-8502 Japan) and Shuichi MATSUMOTO, Member (KDDI R&D laboratories Inc., Fujimino-shi, 356-8502 Japan).

電子情報通信学会誌 Vol.89 No.4 pp.326-332 2006年4月

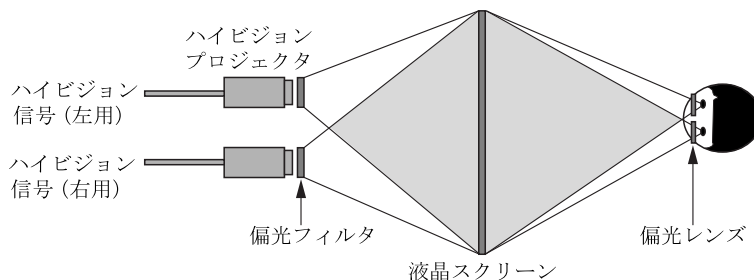


図1 偏光メガネ式の立体表示

を構成する各レンズ越しに左右の画像を縦じま状に表示するものである。イメージスプリッタ方式は液晶パネル上で水平方向に左用、右用の画素を交互に配置し、同パネルとバックライトの間に縦長のスリット（視差バリアと呼ばれる）を規則的に配置するものである。図1に、偏光メガネ式立体表示の原理を示す。偏光メガネ式は、左右の映像をそれぞれ別個のプロジェクタから偏光フィルタを通して液晶スクリーンに投射、視聴者は表示システムがはん用のものでなく、しかも高価であるため、一般に特定のイベント会場において巨大スクリーンと併用する形で使用されている<sup>2)</sup>。

## 2.2 自由視点映像

自由視点映像は、複数のカメラにより撮影された多視点映像により構成され、ユーザの選択した視点に応じて、映像出力を提供するものである。広い意味での自由視点映像は、視点を固定し、視線方向を自由に移動させる全方位映像も含まれるが、本稿においては誌面の都合上、

多視点映像の議論に限定することとする。自由視点映像の単純な例としては、スタジアムでの競技映像を相当数のカメラで取り囲み同時撮影し、受信する映像（カメラ）をユーザの意思により切替可能なインタラクティブな放送サービスが挙げられる。カメラ配置としては、図2に示すように平行に配置する方法（図2(a)）と、注目するオブジェクトにフォーカスを当てて、互いに交差させる方法（図2(b)）が考えられる。カメラを密に配置することで、ユーザに視点の選択肢を多く与えることができ、操作性は向上するといえるが、扱うデータ量の増大や運用性の悪化を招く。このため、カメラ間の実在しない視点は仮想視点として扱い、同視点に対応する映像は、近傍カメラからの情報を基に補完することで、カメラの台数が制限される場合においても、十分な視点数の映像を受信側で生成することが可能となる。このように自由視点映像の信号ソースは、基本的に複数カメラの映像出力により構成されており、カメラの配置方法によっては、2.1で述べた3D映像を包含することにより、3Dによ

## 用語解説

**HDTV (High Definition TeleVision)** 代表的なテレビ用映像フォーマットのうち、NTSCなど従来の標準解像度テレビ(SDTV: Standard Definition TeleVision)を上回る解像度を有する映像フォーマットの総称。画像サイズや走査方式の組合せによって、1080i, 720p, 1080pなど複数のフォーマットが存在するが、このうち1080iが国内及び海外で最も普及している。

**コーデック** 圧縮符号化を行うエンコーダ(encoder)と圧縮ストリームの復号を行うデコーダ(decoder)の組合せに対する総称。映像に限定されず、音声やデータの圧縮符号化に対しても同様に使用される用語。

**デジタルシネマ** 映画コンテンツを素材レベルからパッケージに至るまで、一貫してデジタル形式で扱うシステムの総称。従来の映画素材は35mmフィルムに代表されるアナログ処理を基本としていたが、今後はデジタルシネマに置換されていく流れにある。

**MPEG (Moving Picture Experts Group)** 国際標準化組織であるISO/IEC JTC1において、SC (Sub-Committee) 29 WG (Working Group) 11として主に動画像の圧縮符号化形式に関する国際標準の策定を行う団体の総称。

**MPEG-2** MPEGで策定された動画像用圧縮符号化形式の国際標準の一つ。このうちpart 2は映像符号化形式に関する規定を扱い、デジタル放送やDVDにおいて世界的に

幅広く採用されている。1995年に策定された。

**MPEG-4** MPEGで策定された動画像用圧縮符号化形式の国際標準の一つ。このうちpart 2は映像符号化形式に関する規定を扱い、固定受信やパッケージメディアを主対象とするMPEG-2 Videoとは対照的に、モバイルやインターネットを主な適用領域として、低ビットレートでの高い符号化性能、オブジェクトベース符号化への対応などを特徴とする。1999年に策定された。

**JPEG2000** ISO/IEC JTC1において、SC29/WG1で扱う静止画用圧縮符号化形式の国際標準の一つ。このうちpart 1は映像符号化形式のうち基本部分に関する規定を扱う。2000年に策定された。用途は静止画像に限定されず動画像にも適用可能であり、動画像用のファイルフォーマットをpart 3として標準化している。DCIではデジタルシネマの標準圧縮符号化形式としてJPEG2000 part 1を採用している。

**H.264/MPEG-4AVC (Advanced Video Coding)** MPEGとITU-Tが共同で策定した動画像用圧縮符号化形式の国際標準である。2003年に策定された。MPEG-2, MPEG-4など既存の圧縮符号化方式に対する大幅な性能改善を目的として標準化された。標準化機関ごとに呼称が異なり、またMPEGではMPEG-4の一部として位置付けており、正式名称はMPEG-4 Visual part 10である。

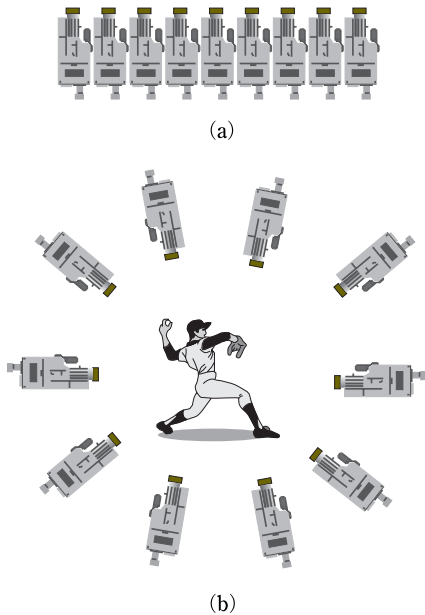


図2 自由視点映像におけるカメラ配置

る自由視点映像を実現できる。

### 2.3 超高精細映像

解像度でHDTVを上回る映像アプリケーションとして、デジタルシネマ<sup>(用語)</sup>が挙げられる。デジタルシネマとは一般に、コンテンツの記録、管理、上映といった一連のプロセスをデジタル処理で実現する映画のことを指す。従来の映画は、素材を35mmフィルムで扱う

表1 DCDMのマスタフォーマット

画像サイズ/ フレームレート	レベル1 4,096×2,160/24Hz レベル2 2,048×1,080/48Hz レベル3 2,048×1,080/24Hz (ただし、水平・垂直のうち一方のみを満たしていればよい)
アスペクト比率	1.896
画素アスペクト比	アスペクトレシオ1:1

ことを基本として、撮影から上映までの一連のプロセスはアナログ処理をベースとするものであった。これに対し、デジタルシネマでは、素材のレベルからデジタルの信号フォーマットで扱う点に特徴があり、従来のフィルムベースでの運用に対する改善点として、記録媒体の低コスト化・高耐久性、素材の加工・編集処理との高い親和性、素材交換や配給などコンテンツ流通の促進などが期待されている。EDCF (European Digital Cinema Forum) ではデジタルシネマを要求品質に応じて階層的に定義しており、35mmフィルムの素材と同等の画質を維持できるレベルを、最高レベルとして位置付けている<sup>(3)</sup>。一方で、デジタルシネマの信号フォーマットに関する標準化がDCI (Digital Cinema Initiative) により進められている。DCIは、米国の映画業界が中心となり結成した、デジタルシネマを専門に扱う標準化団体であり、2005年7月にデジタルシネマ標準規格 version 1.0 を発行した状況にある<sup>(4)</sup>。同規格では図3に示すとおり、映画素材はいったん、DCDM (Digital Cinema Distribution Master) という素材交換用の共通フォーマットを経由して、編集、配給、パッケージ化などの各種フローに至ることを前提としており、DCDMに関する規定が同規格の大半を占めている。DCDMのマスタフォーマット種別を表1に示す。空間解像度としては2種類存在しており、水平画素数に着目すると4,096画素、2,048画素のものがあり、それぞれ4Kマ

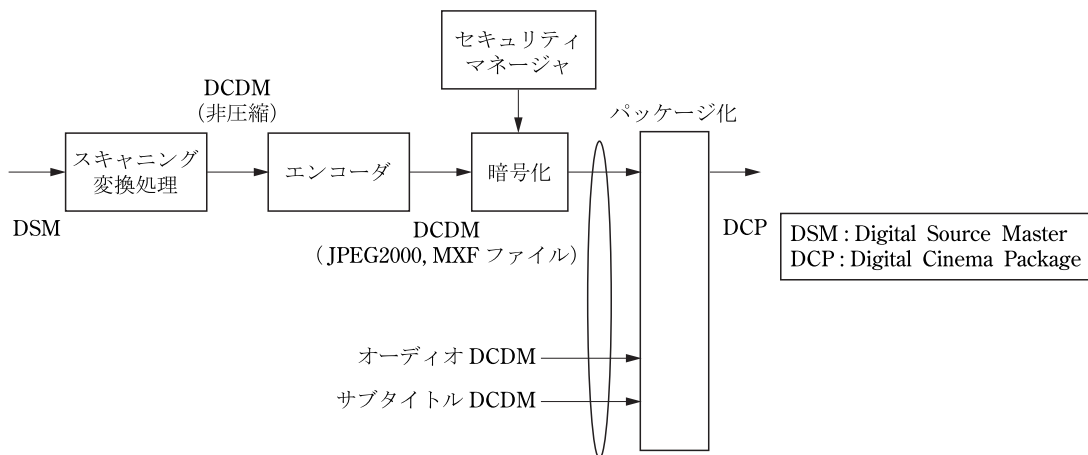


図3 デジタルシネマに関連する処理フロー

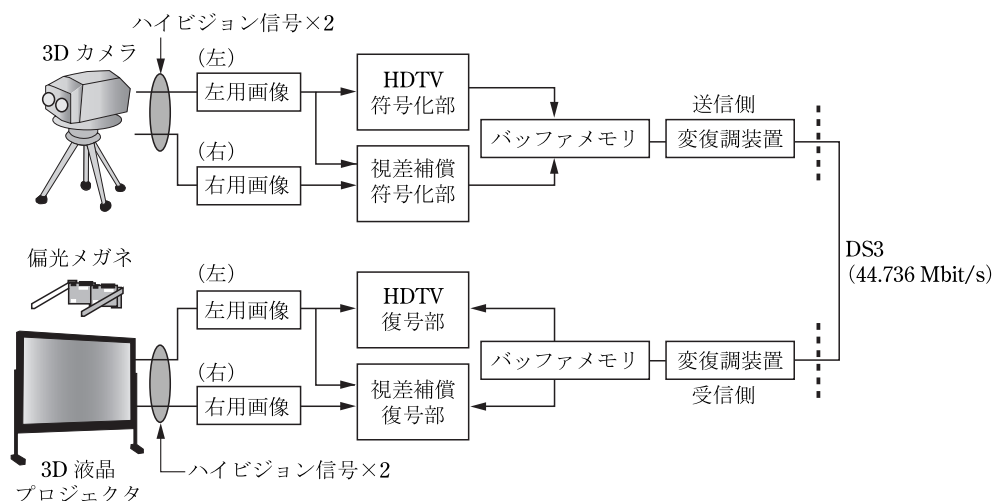


図4 3Dハイビジョン用符号化方式

スタ、2K マスタと呼ばれる。レベル3はレベル1を空間的に1/2 サブサンプリングすることで容易に取得でき、シアターで使用するプロジェクタ性能の違いを想定した配慮がなされている。各レベルに共通して、1画素はX'Y'Z空間の3コンポーネントにより構成され、コンポーネントごとのビット深度は12bitである。

### 3. 圧縮符号化・伝送技術の動向

#### 3.1 3D動画の符号化

図4に、先に筆者らが試作開発を行った立体ハイビジョン伝送方法を示す。本伝送方法を構成する3D-HDTV用符号化装置の基本原則を以下に示す。ある時刻に左眼、右眼でとらえた映像は互いにかなり似通っていることから、左用に映像は単独で符号化するが、右用の映像は左用の映像と補助的な情報(視差の大きさなど)により表現される。ここで、受信側において左用の映像は復号により得られることから、右用に対してはこの補助的な情報のみ送ることで、結果的に右用の映像は復号可能となる。これにより右用の符号化情報は左用に比べて格段に少なくすることができる。以上の考えに基づく立体映像の圧縮符号化方式は視差補助符号化と呼ばれる。

立体映像の性能指標としては、左用映像、右用映像それぞれの画質はもちろんのこと、左右映像の画質のバランスが極めて重要である。一方の画質が他方に比べてはるかに劣るような場合には、立体映像として見た場合には、その損失は明らかに検知され、部分的に立体感が失われてしまうためである。視差補助符号化に基づく立体映像を固定ビットレートで符号化する場合、右用の符号化効率が左用の符号化効率よりもはるかに高いため、受信側で得られる左右の画質のバランスを維持するためには、左用映像のビットレートの配分を右用映像に比べて

高く設定する必要がある。ただし、左用映像と右用映像の符号化効率の差は、時間的に、及び画像の種類により大きく変動することが確認されており、この配分を固定的に定めたものでは再生画像のバランスを保つのは困難である。この点を考慮して、本符号化装置においては左右のビットレートでの符号化効率の大小に応じて適切に決定することにより、受信側で得られる左右の画質のバランスを保持している。

#### 3.2 多視点映像符号化

多視点映像の符号化は、MPEG<sup>(用語)</sup>での標準化においても既に議論されており、MVC (Multi-View Coding) としてCfP (Call for Proposal: 方式提案の募集) が2005年7月のMPEG 会合にて発行された状況にある<sup>(5)</sup>。MVCは、MPEG-2<sup>(用語)</sup>やMPEG-4<sup>(用語)</sup>などの規格と同様に、符号化ストリームの記述方法を定義し、これに対するデコーダの動作を規定するものである。MVCが想定するデコーダの位置付けを図5に示す<sup>(6)</sup>。図中、エレメンタリーストリームが実質的な圧縮符号化データであるが、これに付随して、カメラパラメータなど扱う映像ソースの撮影条件に関連する詳細情報が参照可能である。3.1で述べたとおり、ある被写体に対して、同一時刻に複数視点から撮影された映像は互いに高い相関を有しており、多視点映像符号化において高い圧縮率を得るためには視点間の予測が有効である。多視点映像符号化にお

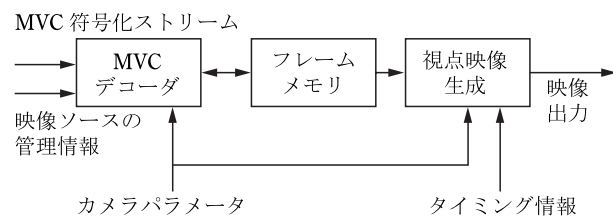


図5 MVC デコーダの位置づけ

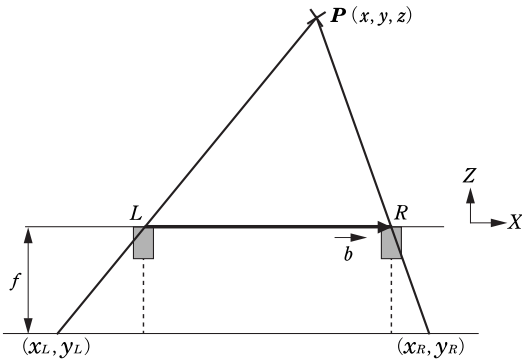


図6 視差ベクトルと奥行き情報の関係

いて、ある視点の映像を基準とした場合、基準映像以外のすべての視点において、個別に基準映像への視差ベクトルを符号化する必要がある、視差ベクトルに費やす符号量は視点数に比例して増大する。ここでカメラパラメータが既知である場合、各視点映像から基準映像に対する視差ベクトルは、基準映像の奥行き情報を基に算出可能である。L、Rの2視点の場合を例にとり以下に説明する。まず視差ベクトルと奥行き情報の関係は図6に示すとおりとなる。図中、Pは実空間座標(X, Y, Z)に位置する被写体を示す。f、bはそれぞれ、焦点距離、視点Lを基準とする視点Rの移動ベクトルを示す。ここで被写体Pに対する両視点間の視差ベクトルdは次式により算出可能である。ただし、dの起点は視点Rの画像とする。

$$d = (x_L - x_R, y_L - y_R) = \frac{f}{z} b \quad (1)$$

### 3.3 超高精細映像

#### (1) マルチ画面の並列伝送方式

解像度でHDTVを上回る映像の入出力は、複数系統のHDTVインタフェースにより扱われるのが一般的で

ある。この構成を用いるメリットとして、伝送系においてもHDTV対応の伝送機器を並列に割り当てることにより高精細映像の伝送システムが構築可能である。マルチ画面の並列伝送方式の一例として、2002年のワールドカップサッカーで実際に使用されたメガビジョン<sup>(7)</sup>の伝送システムについて説明を行う。メガビジョンとは、スポーツ・コンサート会場で行われるイベントを、高臨場感で再現する上で最低限必要とされる解像度として、6,000画素×1,000ライン相当の画面(9:48のアスペクト比)により構成される超高精細映像フォーマットとして位置付けられる。受信側では、Left用、Center用、Right用の信号がMPEG-2デコーダによって個別に復号され、3本のHDTV信号としてプロジェクタに入力される。ここで問題になるのは分割画面間の時間的なずれである。はん用のコーデック系では装置の個体差に起因する遅延差が大きく、受信側で復号画像をそのまま再生した場合の時間のずれを無視できない。同問題への配慮から、MPEG-2コーデックにはKDDI提供のKH-300Nを使用した<sup>(8)</sup>。同コーデックの特長として、大画面映像を複数系統により分割伝送した際の同期再生をサポートしており、同機能はコーデック単体の同期管理を高精度に行うことで、系統間の特別な同期制御を一切介することなく実現される。KH-300Nコーデックで採用している同期メカニズムを図7に示す。エンコーダ側では、映像入力から抽出した59.94Hzから27MHzのシステムクロック(STC)の生成、伝送を行い、これを受けてデコーダ側では、回線入力から抽出した27MHzのSTCから59.94Hzの映像同期信号を生成し、映像タイムスタンプを基準に位相調整を行った上で映像出力をこれにロックさせる。メガビジョン伝送では、エンコーダの映像入力複数系統間で完全に同期しているため、結果としてデコーダの映像出力は同一クロックを基準に処理され、複数系統間で再生時刻が一致する仕組みである。

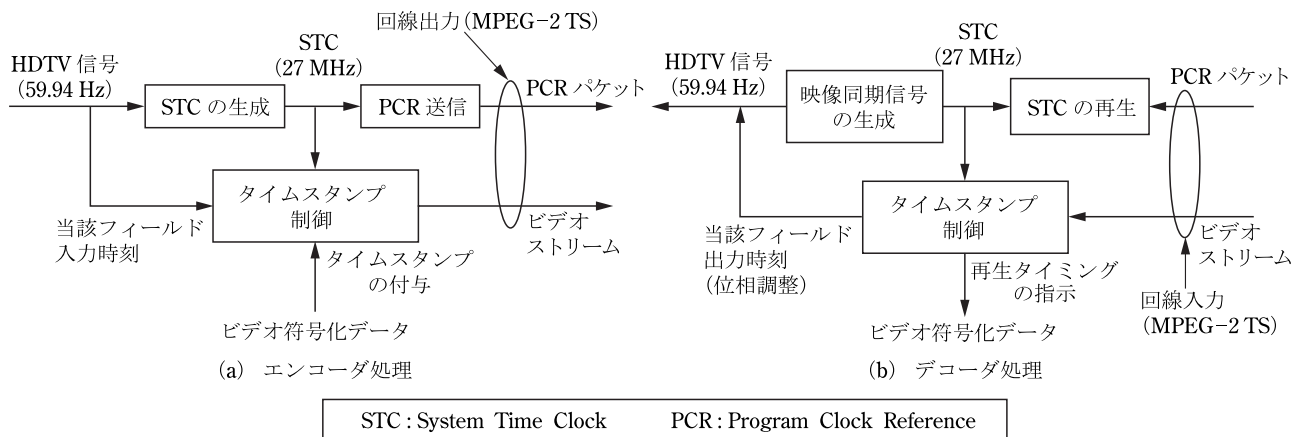


図7 画面間同期メカニズム



表2 DCDMにおけるJPEG2000符号化パラメータ

符号化方式	JPEG2000
デコーダ要求	マスターデータの各レベルがデコードできること
最大ウェーブレット変換レベル	2K:6, 4K:5
コードブロックサイズ	32×32
Precinct サイズ	256×256 (LLサブバンドのみ128×128)
コンポーネント内のタイル分割	非適用
符号化レート	最大250Mbit/s

表3 デジタルシネマ用符号化方式の比較一覧

符号化方式	MPEG-4 Simple Studio プロファイル	JPEG2000 Part 1	H.264 High 4:4:4 プロファイル
DCDM マスタレベルとの対応	レベル 2,3	レベル 1,2,3	レベル 1,2,3
符号化レートの最大値	1,800Mbit/s	制限なし	960Mbit/s
ビット深度の最大値	12bit	38bit	12bit
予測タイプ	Iのみ	Iのみ	I,P,B

(2) デジタルシネマ用符号化

DCIのDCDMでは圧縮方式及びファイルフォーマットの規定を含んでおり、映像符号化にはJPEG2000<sup>(明)</sup> part 1<sup>(9)</sup>を使用し、ファイルフォーマットはMXF形式としている。JPEG2000 part 1で使用する符号化パラメータについては表2に従う制限が定義されている。一方で、DCI規格では採用されていないものの、デジタルシネマへの適用を意識した圧縮符号化標準として、MPEG-4スタジオプロファイルとH.264<sup>(明)</sup>ハイプロファイルが挙げられる。両符号化方式の比較一覧を表3に示す。参考のため、JPEG2000 part 1に関する情報を併せて示す。

ここで、MPEG-4やH.264は動き補償予測とDCTの併用を基本とする符号化アルゴリズムであるが、元々想定している画像サイズは高々HDTVクラスであるため、4Kマスタなど高精細画像に適用した際には、更なる改良の余地が見込まれる。例えばH.264ハイプロファイルでは動き予測の単位となるマクロブロックサイズの上限は16画素×16ラインであるが、同制限を拡張し、32画素×32ラインや64画素×64ラインといった、より大きなサイズの予測単位を選択可能とすることで、4Kマスタへの適用時における低ビットレートの効率改善が確認されている<sup>(10)</sup>。

4. 実用化に向けて

次世代映像メディアとして期待される、自由視点映像や超高精細映像を放送や通信のインフラで扱うためには、圧縮伝送用のコーデックが必須であるが、実用化段階にあるものはまだ存在していない状況にある。放送や配信といったサービス用途で利用されるコーデック装置は従来から専用LSIによるハードウェア実装が一般的であり、これら次世代映像メディアのコーデック実装においても例外ではない。一方で、近年のCPU速度の向上に後押しされる形で、はん用のCPUを多用し、分散並列型のソフトウェア処理により超高速の演算性能を達成するHPC(High Performance Computing)が注目されている<sup>(11)</sup>。動画像符号化はもともとアルゴリズム自体の並列度は高く、HPCの活用により、オールソフトウェアでのコーデック実装も可能になると期待されている。従来のハードウェア実装と比べ、ソフトウェアベースの実装には、LSI開発を伴わないという点での低コスト化が見込まれるのに加え、システム導入以後に生じると予想される、新たな映像フォーマットや符号化ツールの追加といった機能強化の要望にもソフト的に対応できるといった利点がある。デメリットとしては、ハード実装に比べ、長時間動作における安定性がやや劣る点が懸念されるものの、CPU性能の向上、OSの安定化、及びPCアーキテクチャの進化が今後も見込まれるため将来的には解消される問題であると期待される。

5. ま と め

次世代映像メディアとして、3D映像、自由視点映像、超高精細映像に着目し、その基本原理と信号フォーマットについて導入した上で、これらを放送やインターネットのインフラを介して扱う上で重要となる圧縮符号化技術、並びに伝送技術の最新動向を解説した。将来の映像メディアが高臨場感、インタラクティブ、高精細の方向に進んでいることは明らかであるが、実サービス展開へ導く上では、圧縮符号化のみならず、撮影や編集といったコンテンツ制作技術、及び表示デバイスやストレージに代表される端末系技術の確立が必須である。特に端末系においては、心理学や医学の観点から、ストレスを感じさせない自然な映像提示が大前提であり、これに関連する革新的技術の創出が今後の実用化スピードを大きく左右するものと予想される。

文 献

- (1) 安東孝久, 増谷 健, “メガネなし立体ディスプレイ,” 映情学誌, vol.55, no.8/9, pp.1071-1075, Aug. 2001.
- (2) 内藤 整, 松本修一, “3Dシアターとその伝送用機器,” 映情学誌, vol.52, no.6, pp.786-788, June 1988.
- (3) 青山友紀, “800万画素デジタルシネマの実現に向けて,” 映

- 情学誌, vol.57, no.2, pp.187-191, Feb. 2003.
- (4) Digital Cinema Initiatives, LLC, "Digital Cinema System Specification V1.0," July 2005.
  - (5) "Call for Proposals on Multi-view Video Coding," ISO/IEC JTC1/SC29/WG11 N7327, July 2005.
  - (6) "Requirements on Multi-view Video Coding v.4," ISO/IEC JTC1/SC29/WG11 N7282, July 2005.
  - (7) 内藤 整, 大場省介, 松信章一, "ワールドカップサッカーのメガビジョン中継," 映情学誌, vol.56, no.10, pp.1559-1561, Oct. 2002.
  - (8) 内藤 整, 松本修一, "地上波デジタル中継用 HDTV コーデックの開発," 情処学研報, vol.29, no.4, pp.19-24, June 2000.
  - (9) "JPEG2000 part I Final Draft International Standard," ISO/IEC JTC1/SC29/WG1 N1890, Sept. 2000.
  - (10) 松村篤志, 内藤 整, 川田亮一, 小池 淳, 松本修一, "動的に更新される背景バッファを用いた自由視点映像に対する補完方式," 映情学誌, vol.59, no.6, pp.901-908, June 2005.
  - (11) <http://hase.hpcc.jp>



ないとう せい (正員)

平6 早大・理工・通信卒, 平8 同大学院修士課程了。同年国際電信電話(株)入社。以来, 同社研究所にて主に動画像符号化の研究に従事。現在, (株)KDDI 研究所映像通信グループ研究主査。平11 年度本会学術奨励賞受賞。



まつもと しゅういち (正員)

昭52 北大・工・電子卒, 昭54 同大学院修士課程了。同年国際電信電話(株)入社。以来, 同社研究所にて HDTV, SDTV の高能率符号化方式の研究・開発に従事。現在, (株)KDDI 研究所取締役。工博, 昭58 年度本会論文賞, 昭61 年度本会篠原記念学術奨励賞, 平11 年度本会業績賞各受賞。

