

MPEG-4 Visual 符号化の最新動向

鈴木 芳典[†]

あらまし 昨年12月に規格化された MPEG-4 Part 10 Advanced Video Coding (AVC) | ITU-T 勧告 H.264 の基本技術を中心に、最近の MPEG ビデオ符号化の技術動向ならびに標準化活動状況について報告する。AVC は、ISO MPEG と ITU-T VCEG (Video Coding Experts Group) の共同プロジェクト (JVT; Joint Video Team) により開発された最新のビデオ符号化規格であり、既存の MPEG 規格と比較して 1.5 倍～2 倍の高画質映像を提供することが報告されている。また、現在 MPEG ビデオグループでは、スケーラブルビデオや 3 次元ビデオなど新しい検討項目について審議を行っている。

Recent Trend of MPEG-4 Visual

Yoshinori Suzuki[†]

Abstract This paper reports on the features of MPEG-4 part 10 Advanced Video Coding (AVC) | ITU-T H.264 and the current activities of MPEG Video. The AVC is the latest international standard for video coding developed by the Joint Video Team which is organized by ISO/MPEG and ITU-T/VCEG (Video Coding Experts Group), and it is reported that the visual quality of AVC is 1.5-2.0 times as good as that of the existing MPEG standards. The MPEG Video group has some new activities such as the scalable video and the 3D video.

1. はじめに

1995 年に標準化活動を開始した MPEG-4 ビデオ規格 (MPEG-4 Part 2 Video) [1] は、1998 年 10 月に基本規格、2000 年 1 月と 2001 年 1 月に拡張規格の仕様を凍結し、技術的な審議をほぼ終了している。その後、MPEG ビデオグループは、MPEG-4 Part 10 Advanced Video Coding (AVC) [2] の規格化など、MPEG-4 Part 2 の枠を越えた標準化活動を展開している。本稿では、2001 年以降の MPEG ビデオグループの活動について AVC の基本技術を中心に報告する。

2. AVC の特徴

2.1. 基本アルゴリズム

AVC は、ISO MPEG と ITU-T VCEG が 2001 年 12 月に設立した共同プロジェクト JVT (Joint Video Team) が策定した新しい動画符号化方式であり、その復号手順は、ISO/MPEG-4 Part 10 ならびに ITU-T 勧告 H.264 として両機関から公開されている。AVC の基本アルゴリズムは既存 MPEG 方式と同じくマクロブロック (MB) 単位の予測-直交変換である (図 1)。しかしながら、その構成要素には新しい機能が導入

されており、主観画質は既存 MPEG 方式の 1.5～2 倍まで向上すると報告されている [3]。

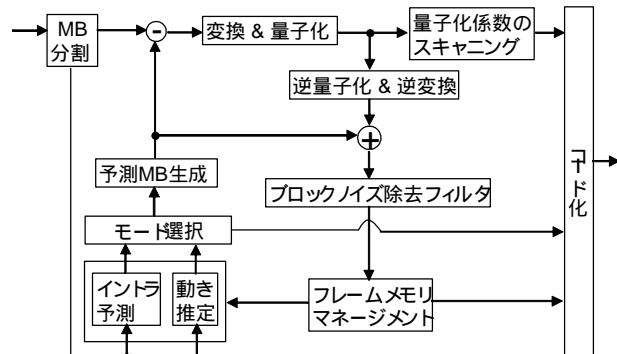


図 1. AVC 符号化のブロック図例

以下で各機能の特徴を説明する。なお、AVC の規定範囲は復号手順であり、符号化手順に関する記述は一例であることをご了解頂きたい。

2.2. 予測

AVC のフレーム内/フレーム間予測には画像の局所的な特徴の変化を考慮した数多くの方法が用意されている。これらの予測方法は予測に用いる参照画像の違いにより 3 種類に分別されており、1 種類をスライス単位で選択できる。各スライスでは、その種類の候補の中から MB 毎に予測方法を選択する。本節では、スライスの種類 (I/P/B-Slice) と各スライスに属する予測方法について説明する。

[†] (株)日立製作所 中央研究所
〒185-8601 東京都国分寺市東恋ヶ窪 1-280
[†] Central Research Laboratory, Hitachi, Ltd.
1-280 Higashi-koigakubo, Kokubunji-shi, Tokyo,
185-8601 Japan

2.2.1. I-Slice

イントラ予測のみを選択候補とするスライスタイプ。MPEG-4 イントラが DCT 係数領域の予測であるのに対して、AVC イントラは画素領域の予測を採用している。符号化済み隣接ブロックの境界画素（隣接画素）を用いた外挿予測であり、予測モデルが異なる複数種類の予測モードを輝度信号と色差信号とで個別に準備している。その選択は個別に実施する。

・輝度信号

4 種類の Intra 16×16 予測モードと Intra 4×4 予測モードが用意されており、MB 単位で 1 種類を選択する。Intra 16×16 は 16×16 ブロック単位の予測であり、予測方法には、2 種類の指向性モデルモード（水平、垂直）と隣接画素の平均値を予測値とする DC モード、隣接画素に合致するように予測値をなだらかに変化させる平面予測モードがある（図 2）。一方、Intra 4×4 は 4×4 ブロック単位の予測であり、DC モードと 8 種類の指向性モデルモード（図 3）の中から、 4×4 ブロック単位で 1 種類を選択する。

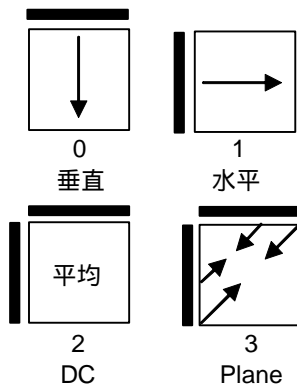


図 2. 隣接画素ブロックの境界画素

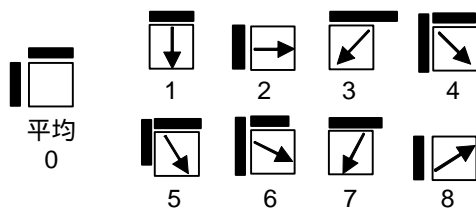


図 3. イントラ 4×4 予測の指向性モデル

・色差信号

輝度信号の Intra 16×16 と同じ 4 種類のモードが用意されており、MB 単位で 1 種類を選択する。但し、2 種類の指向性モデルモードと平面予測モードは 8×8 ブロック単位、DC モードは 4×4 単位で実施する。

2.2.2. P-Slice

イントラ予測に加えて 1 方向のフレーム間予測を選択候補に含むスライスタイプ。利用できる予測方法は 1 種類であるが、動き補償に多くのブロック分割モードが用意されており、高精度の動き検出が実

現できる。また、複数の参照画像を動き補償の候補とするマルチフレーム予測を採用しており、アンカバード部分の予測効率向上も期待できる。さらに、 $1/4$ 画素精度の動き内挿処理が導入されるなど、画質改善に向けた高精細化が図られている。

・動き補償

16×16 、 16×8 、 8×16 、 8×8 の 4 種類の基本分割モードがあり、マクロブロック毎に 1 種類を選択できる。更に 8×8 分割モードを選択したマクロブロックについては、 8×8 、 8×4 、 4×8 、 4×4 の 4 種類のサブ分割モードから 8×8 ブロック毎に 1 種類を選択できる。従って、符号化する動きベクトルの数は分割モードの選択結果により、1 ~ 16 本まで変化する。

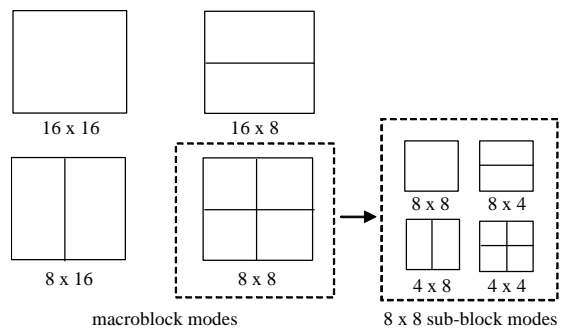


図 4. ブロック分割

・画素内挿処理

輝度動きベクトルの精度は $1/4$ 画素であり、実数画素精度の予測画素は 6 タップ $[1, -5, 20, 20, -5, 1]$ と 2 タップ $[1, 1]$ の内挿フィルタを用いて算出する。図 5 に $1/2$ 画素精度の予測画素の算出方法を示す。図中の \square は整数画素であり、 \circ の $1/2$ 画素は、水平と垂直方向の隣接 6 整数画素に対して、それぞれ 6 タップフィルタ処理を施して生成する。の $1/2$ 画素は、垂直あるいは水平方向の \square の隣接 6 画素に対して 6 タップフィルタ処理を施して生成する。

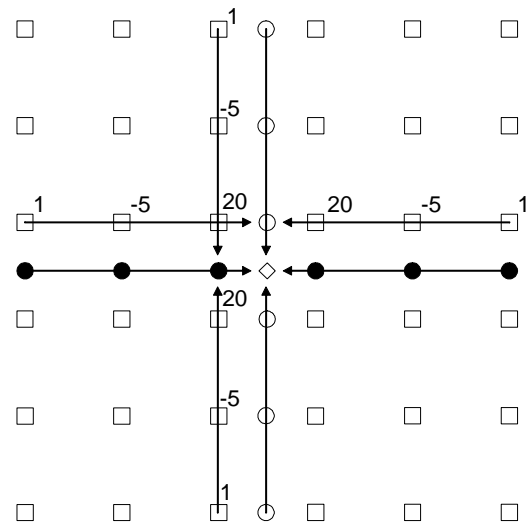


図 5. $1/2$ 画素内挿処理

図 6 に 1/4 画素精度の予測画素の算出方法を示す。図中の □ は整数画素、△ は 1/2 画素を示している。○ と △ の 1/4 画素は、水平と垂直方向の隣接する 2 つの整数あるいは 1/2 画素に対して、2 タップフィルタの線形内挿処理を施して生成する。△ で示す 1/4 画素は、斜め方向の隣接する 2 つの 1/2 画素に対して、2 タップフィルタの線形内挿処理を施して生成する。

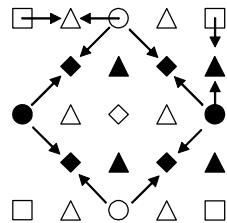


図 6. 1/4 画素内挿処理

色差成分の動き補償では、輝度成分のブロックサイズと動きベクトルの値をそれぞれ縦横半分にした値が用いられる。動きベクトルの精度は 1/8 画素であり、整数精度以下の予測画素は、隣接する 4 個の整数画素に対する単純線形内挿により生成する。

・マルチフレーム

輝度成分の基本分割モードの分割ブロック毎に、複数の候補から予測に用いる参照画像が選択できる。AVC では、全ての再生画像を対象に、参照画像用に保存するか否かを指定する。保存した再生画像の中から参照画像の候補をスライス毎に優先番号つきで選択する。この参照画像の候補を参照ピクチャリストと呼ぶ。図 7 は現フレーム（黒）内のスライスについて、保存した 4 枚の再生画像（白抜き）から 3 枚の参照画像を優先番号つきで選択した例である。

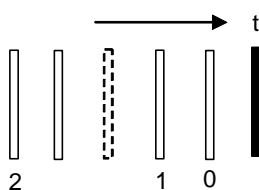


図 7. 参照画像番号割当て例

・動きベクトルの符号化

MPEG-4 と同様、隣接ブロックに属する動きベクトルの中間値予測により算出した予測ベクトルと符号化対象動きベクトルとの差分値を水平・垂直個別に符号化する。なお、AVC では、差分符号化を伴わない Skip モード (Not Coded MB) の再生 MB は、予測動きベクトルを用いた 16×16 モードの 1 方向予測により算出する。

2.2.3. B-Slice

イントラ予測に加えて 2 方向のフレーム予測を選択候補に含むスライスタイプ。このスライスでは、2

種類の参照ピクチャリスト(リスト 0 とリスト 1)を予測に利用する。図 8 は、現フレーム(黒)内のスライスについて、2 種類のリストを生成した例である。このように、リスト 0 では現フレームに対して時間的に過去の再生画像しか選択できないが、リスト 1 では時間的に将来の再生画像も選択できる。

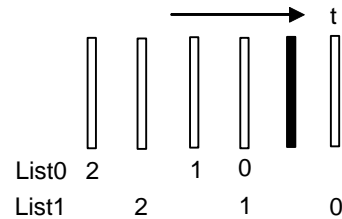


図 8. 2 リストへの参照画像番号割当て例

B-Slice では、この 2 種類のリストを用いた 4 種類の予測方法が定義されている。

- ・ 1 方向予測(List 0) ; List0 を用いた 1 方向予測
- ・ 1 方向予測(List 1) ; List1 を用いた 1 方向予測
- ・ 2 方向予測 ; 両リストを用いて個別に 1 方向予測を実施。2 個の予測ブロックの各画素を内挿処理して予測値を生成する。内挿処理には通常は平均内挿を適用するが、ディゾルブのような特殊効果を伴う画像の予測効率を高める重み付け内挿処理 (Weighted prediction) も利用できる。
- ・ 直接予測 ; 動きベクトル、参照フレーム、予測方向を周囲の情報から自動選定する予測方法。Temporal 予測と Spatial 予測の 2 種類があり、スライス毎に選択できる。Temporal 予測では、自動的に指定される後方フレームにおいて、現 MB と空間的に同一位置の MB の符号化情報(予測モード、動きベクトル、参照フレーム)から選定する。一方、Spatial 予測では、同一フレーム内の隣接する 3 個の MB の符号化情報から選定する。

予測方法と参照画像の選択は、基本分割モードの分割ブロック毎に実施する。但し、直接予測については、16×16 モードの MB と 8×8 サブモードを選択した 8×8 ブロックにのみ適用可能な予測方法であり、他の分割モードでは選択できない。なお、B-Slice では、Skip モードの再生 MB は、16×16 モードの直接予測により算出する。

2.3. 周波数変換と量子化

AVC では、4×4 DCT を整数近似した 4×4 整数変換を予測誤差信号の周波数変換に採用している。そのため、変換ミスマッチは発生しない。量子化については、量子化精度が約 12.5% ずつ変化する 52 種類の量子化器を用意している。そのため、エンコーダの画質・符号量制御を実施する際の自由度が高い。本節では、変換・量子化の方法について説明する。

・変換、逆変換

4×4DCTの変換行列 H とそれを整数近似した 4×4 整数変換行列 H_F を以下に示す。

$$H = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ c & s & -s & -c \\ 1 & -1 & -1 & 1 \\ s & -c & c & -s \end{bmatrix}$$

但し、 $c = \sqrt{2} \cos(p/8)$, $s = \sqrt{2} \sin(p/8)$

$$H_F = \text{round}(2.5 \cdot H) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$$

上記のように、4×4DCTの変換行列 H の各係数を 2.5 倍して整数値に丸め込んだ行列を 4×4 整数変換として採用している。この整数行列は、予測誤差信号の 9 ビット入力を乗算なしで 16 ビットの係数に変換できるという利点があるが、一方で、丸め込み誤差の発生や各行の非直交性、1, 3 列と 2, 4 列のノルムが異なるという変換効率を劣化する要因も含んでいる。そのため、後述する量子化処理にて、DCT の整数化に伴う丸め込み誤差の調整を行っている。入力ブロックを X 、DCT 変換行列を Y 、整数変換行列を W 、整数化に伴う丸め込み誤差を補正する調整行列を E とすると、整数変換は下記の式にて示される。

$$Y = (H_F \cdot X \cdot H_F^T) \otimes E = W \otimes E$$

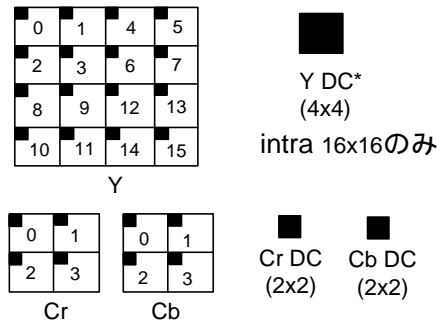
ここで、 \otimes は行列内の要素毎の乗算を示す。

逆変換行列には、 H_F の転置行列ではなく、 H_F の転置行列の第 2 行と第 4 行の各係数を 1/2 倍した H_I を適用する。これは、16 ビット演算に対応するための処置であり、乗算分は後述する逆量子化処理にて調整する。なお、1/2 倍演算は 1 ビット右シフトにより対応することが規格書で規定されているため、変換ミスマッチは発生しない。

$$H_I = \begin{bmatrix} 1 & 1 & 1 & 1/2 \\ 1 & 1/2 & -1 & -1 \\ 1 & -1/2 & -1 & 1 \\ 1 & -1 & 1 & -1/2 \end{bmatrix}$$

画素解像度が低いために 4×4 整数変換後も冗長性が残る色差と Intra16×16 輝度の DC 成分については、更に 2×2 または 4×4 アグマール変換を実施する(図 9)。変換後の DC 係数も 16 ビットで表現できる。なお、これら DC 成分の逆変換については、逆変換時のダイナミックレンジ使用効率の観点から逆量子化

の前に実施する。



近似整数変換 Hadamard 変換

図 9. 変換ブロックの種類

・量子化

スカラー量子化を適用しており、52 種類の量子化器を用意している。量子化精度は量子化パラメータ (QP) と量子化ステップサイズ (Qstep) にて識別される。QP と Qstep は表 1 に示す関係があり、QP の値が 6 大きいと Qstep が 2 倍となるように設計されている。

表 1. 量子化パラメータとステップサイズの関係

QP	0	1	2	3	4	5
Qstep	0.625	0.6875	0.8125	0.875	1	1.125
QP	6	7	8	9	10	11
Qstep	1.25	1.375	1.625	1.75	2	2.25
QP	12	18	24
Qstep	2.5	5	10
QP	30	36	42
Qstep	20	40	80
QP	48	51			
Qstep	160	224			

量子化変換行列を Y_Q 、DCT 変換行列を Y 、整数変換行列を W とすると量子化処理は下記にて定義される。

$$Y_Q = \text{round}(Y / Q_{step}) = \text{round}\left(\frac{W \otimes E}{Q_{step}}\right) = \text{round}\left(W \otimes \frac{L(QP\%6)}{2^{15+(QP/6)}}\right)$$

ここで、 L は整数変換係数をスケールングするための整数行列であり、 $(QP\%6)$ の値に対して 6 種類が用意されている。下記の計算により算出できる。

$$L(QP\%6) = \frac{E \cdot 2^{15}}{Q_{step}}$$

この量子化処理には、中間値表現に 32 ビットを必要とするが、量子化変換係数は 16 ビットで表現できる。

逆量子化された変換行列を Y' とすると、逆量子化処理は以下の式にて定義される。

$$Y' = Y_Q \otimes E_D \cdot Q_{step} \cdot 64 = Y_Q \otimes R(QP\%6) \cdot 2^{(QP/6)}$$

ここで E_D は、逆変換行列 H_I 生成時における 1/2 倍

化処理の影響を E に加えた調整行列である。 R は量子化変換係数をスケールリングするための整数行列であり、下記の計算にて算出される。 L と同様に(QP%6)の値に対して6種類が用意されている。

$$R(QP\%6) = E_D \cdot Q_{step} \cdot 64$$

なお、64の乗算は、 R の要素の整数化に伴う丸め込み誤差の発生を防止するための処理である。乗算した分は、下記のように再生ブロック X_D 算出時の正規化処理により修正する。

$$\text{逆変換} ; X' = (H_I \cdot Y' \cdot H_I^T)$$

$$\text{正規化} ; X_D(i, j) = (X'(i, j) + 2^5) \gg 6$$

Intra 16×16の輝度DC成分と色差DC成分の量子化・逆量子化については、上記と同じ原理で実施されるため、詳細説明は割愛する。但し、DC成分のみを対象とした処理となるため、 L と R については、全要素を(0,0)要素の値に置き換えた行列を用いる。

2.4. 符号化

AVCのエントロピー符号化方法には、可変長符号を適用するUVLC/CAVLC (Universal Variable Length Coding/ Context-based Adaptive VLC)と算術符号を適用するCABAC (Context-based Adaptive Binary Arithmetic Coding)の2種類があり、ピクチャ単位で選択できる。以下、これらの特徴を説明する。

・UVLC

符号要素によって変化しない単一符号表を用いた可変長符号化方式。符号要素の各値に対して事前にコード番号を割り当てておけば、コード番号から符号が一意に決定する(表2)。

表 2. UVLC 符号表

コード番号	符号
0	0
1	010
2	011
3	00100
4	00101
.....	

符号要素固有の生起頻度には対応できないが、符号表を個別に用意する必要がないため、符号化表を保存するためメモリコストが低減できる。スライスヘッダやCAVLC適用時のMBヘッダ情報の符号化に用いられる。AVCでは、誤り耐性の高いGolomb符号が単一符号として採用されている。

・CAVLC

予測誤差信号の符号化に用いられる。量子化変換係数を1次元系列にスキャンし、run(連続する0の数)とcoeff(有意係数の絶対値)の組み合わせを可変長符号化するというコンセプトは既存のMPEG方式と同

じである。これに加えてCAVLCでは、予測効率が高く有意係数の発生頻度が低いというAVCの特徴を考慮した2種類の符号要素と、符号要素の各値の発生頻度が局所的に変動することを考慮した適応符号表が導入されている。表3に特徴をまとめる。

表 3. CAVLC の特徴

技術項目	新しい機能
符号要素 (coeff)	係数の特徴(ブロック内の有意係数の数と有意係数のうちスキャン順で先頭から連続する'1'の数)を2次元テーブル(coeff_token)にて符号化
符号要素 (run)	スキャン順で最後の有意係数に至るまでの'0'の数(total_zeros)を符号化
適応符号表	符号要素に対して複数の可変長符号表を用意。隣接ブロックの符号化パラメータに応じて自動的に切り替える。coeff_token用に5個、total_zeros用に18個、runの符号用に7個用意

・CABAC

MBデータの符号化に用いられる。MB内の各符号要素の符号化は以下の手順で実施される。

- (1) 符号要素の値をバイナリデータに変換、
- (2) バイナリデータの各ビット(bin)について、binの算術符号化に適用するコンテキストモデルの選択とbinの算術符号化を実施、
- (3) 発生頻度に基づいてコンテキストモデルを更新。バイナリ化の方法には、Golomb形式や固定長形式など数種類が用意されており、適用する方法は符号要素毎に規定されている。コンテキストモデルは399種類用意されており、MPS(most probable symbol)の値(0 or 1)とLPS(least probable symbol)の生起確率(64通りの値に量子化)により規定される。適用するモデルは、符号要素、binの位置ならびに隣接ブロックの符号化値の組み合わせにより自動的に選択される。

2.5. デブロッキングフィルタ

AVCのブロック歪除去処理は、参照画像に適用する規格必須の機能であり、表示画像の画質改善のみならず、予測誤差信号の符号量を削減する効果もある。MB単位の処理であり、図10に太線で示す輝度成分の垂直・水平4境界ならびに色差成分の垂直・水平2境界のそれぞれ隣接6画素を対象とする。修正画素値の生成には、隣接8画素を利用する。

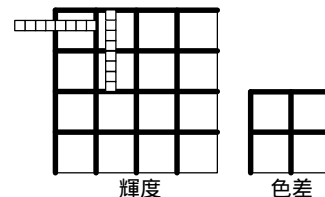


図 10. フィルタ処理対象のMB内のブロック境界
フィルタ処理実施のON/OFFは、MB境界とMB内

のブロック境界に分けて、スライス単位で個別に指定できる。フィルタの強さと勾配は、境界上下または左右のブロックの符号化パラメータ(予測タイプ、参照画像、動きベクトル、QP)とスライス単位で指定する隣接画素間差分閾値と隣接画素値の関係により決まる。なお、修正画素は次のフィルタ処理に利用されるため、1画素が複数回修正される場合もある。

2.6. インタレース信号対応

AVCは、4:2:0色信号フォーマットで8ビット/サンプルのプログレッシブ映像とインタレース映像を符号化対象としている。符号化ピクチャタイプには、フレーム、フィールド、フレーム/フィールド適応の3種類が用意されており、ピクチャ単位で切り替え可能。さらに、フレーム/フィールド適応では、MB単位でフィールドとフレームの切り替えが可能である。

2.7. Profiles@Levels

AVCでは、3種類のプロファイル(表4)と最大ビットレートを64kbps~240Mbpsの範囲で設定した15種類のレベルが用意されている。

表4. プロファイルの機能

	予測	符号化	インタレース
Baseline	I/P	CAVLC	未対応
Extended	I/P/B	CAVLC	未対応
Main	I/P/B	CAVLC/CABAC	対応

3. 審議中のビデオ技術検討項目

本章では、AVC以降のMPEGビデオグループの動向について概要を説明する。

3.1. AVCの拡張規格

対応可能なサンプル当たりのビット数と色信号フォーマットを拡張するAVC新規格。サンプル当たりのビット数を10ビットまたは12ビット、色信号フォーマットを4:2:2または4:4:4まで拡張できる。また、AVCの基本規格が扱うYCrCb色変換以外にRGBとの整合性が高い独自の色変換方法を採用する見込み。拡張規格は、Fidelity Range Extensions Amendmentの名称で3種類の新プロファイル(4:2:2/10, 4:4:4/12, 4:4:4/8)を規格化する予定。仕様凍結は2004年7月。

3.2. スケーラブルビデオ符号化

Wavelet技術を適用した多重解像度符号化方式をMPEG-21 Part 13 Scalable Video Codingとして標準化するための議論が進んでいる。既に技術提案募集[4]が始まっており、23の機関が応募したと伝えられている。2004年2月に各提案方式の画質評価実験が行われ、同年3月のMPEG会合にて実験結果の評価と今後の方針について審議する予定。スケーラブル符号化に関しては、2001年12月から技術調査を開始。Waveletを利用した3種類の技術(wavelet in loop, in-band prediction, interframe wavelet)について、実

験・評価が行われ、その有効性が報告されている[5]。

3.3. 3次元ビデオ符号化

3次元オーディオ・ビデオ符号化(3DAV)技術について、新しい標準規格策定の必要性を問うコメントが募集されている[6]。集められたコメントは2004年3月会合にて審議される予定。3Dビデオについては、インタラクティブ性の向上を目指して2001年1月に技術調査を開始。全方位ビデオ、自由視点ビデオ、立体ビデオなどの実現に必要な要素技術について議論してきたが、自由視点ビデオを除いては既存のMPEG-4技術にて基本的にかバーできることが報告されている[7]。

3.4. MPEG Intra only 符号化

MPEGのイントラ符号化のみを適用した動画像符号化方式(Intra-only MPEG)に関する検討が2003年7月から始まっている[8]。Intra-only MPEGはランダムアクセスや編集性の面での利点があり、技術的には既存規格でサポートされている。しかしながらプロモーションが不十分であり、認知度が低いとMPEG委員会は考えている。この状況を鑑み、新しいプロファイルや技術報告書の作成、新たな技術検討などの必要性を検討するグループを設立した[9]。現在は、既存MPEG規格ならびにJPEG規格のIntra-onlyビデオの性能ならびにコスト比較を実施している。

4. まとめ

MPEG-4新規格であるAVC|H.264の基本技術と現在MPEGビデオにて審議されている検討項目について報告した。MPEG-4ビデオとしての活動は収束方向にあるが、新しい標準化の動きが既に始まっている。

文 献

- [1] "Text of ISO/ICE 14496-2 Third Edition", March 2003.
- [2] "Text of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 Advanced Video Coding)", May 2003.
- [3] "Report Of The Formal Verification Test on AVC (ITU-T Rec. H.264 | ISO/IEC 14496-10)", Dec. 2003.
- [4] "Call For Proposals on Scalable Video Coding Technology", MPEG2003/N6193, Dec. 2003.
- [5] "The Status of Interframe Wavelet Coding Exploration in MPEG", MPEG2003/N4928, July 2002.
- [6] "Call For Comments on 3DAV", MPEG2003/N6051, Oct. 2003.
- [7] "Report on 3DAV Exploration", MPEG2003/N5878, July. 2003.
- [8] Kohtaro Asai, "Suggestion of intra-only MPEG", MPEG2003/M9947, July 2003.
- [9] Kohtaro Asai, "AHG on Intra-only coding in MPEG", MPEG2003/M10246, Dec. 2003.