

# H.264/AVC 動画像符号化標準

H.264/AVC Video Coding Standard

高橋 真毅\*  
Maki Takahashi

## 要 旨

H.264/AVCは、ITU-Tのビデオ符号化専門家グループ(VCEG)とISO/IECの動画像符号化専門家グループ(MPEG)によって開発された動画像符号化の最新の国際標準である。H.264/AVCは、TV電話のように低解像度・低ビットレートアプリケーションから、HDTVのように高解像度・高ビットレートアプリケーションにまで適応できるように設計されており、さまざまなアプリケーションでの利用が期待されている。本稿では、H.264/AVCの技術的特長の概要について解説する。

H.264/AVC is the latest international video coding standard, which was jointly developed by the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG). It was designed and is expected to be used in applications of various resolutions and bit-rates (e.g. video phones at low end and HDTV at high end). This article gives an overview of the new video coding standard.

### まえがき

H.264/AVCは、1998年後半にITU-TのVCEGで検討が始まったH.26Lをベースに、ITU-TのVCEGとISO/IECのMPEGの2つの団体により結成されたJVT (Joint Video Team)によって開発が進められた動画像符号化方式である<sup>1) 2)</sup>。

H.264/AVCおよびその原型であるH.26Lでは、従来の動画像符号化方式との互換性は考慮せず、符号化効率を追求する開発が行われた。したがって従来の動画像符号化方式では複雑度の観点から採用されなかった技術も積極的に取り入れられた。その結果H.264/AVCは、従来の動画像符号化方式であるMPEG-2/4等と比較して、約2倍の符号化効率で同程度の画質を実現している。

### 1. プロファイルと想定アプリケーション

H.264/AVCは、TV電話のように低解像度・低ビットレートアプリケーションからHDTVのように高解像度・高ビットレートアプリケーションまでカバーする汎用的な動画像符号化方式を目指し、想定さ

れるアプリケーションの特性に応じて、図1に示す3つのプロファイルが策定されている。以下に3つのプロファイルとその特徴を示す。

#### (1) ベースラインプロファイル

主にテレビ電話やテレビ会議システム等の双方向通信アプリケーションや、携帯電話等のワイヤレス/モバイルアプリケーションでの利用を想定したプロファイルであり、地上デジタル放送における携帯端末向けの「1セグメント放送」にも採用されることが決定している。

なお、双方向通信アプリケーションでのリアルタイム性を重視して双方向予測のように符号化遅延の大きい符号化ツールは省略されている。また、モバイルアプリケーションでのリソースの制約も重視され、複雑度の高い符号化ツールも省略された構成となっている。例えばエン트로ピー符号化用の符号化ツールとしては、CAVLC (Context-Based Adaptive Variable Length Coding) のみが利用可能な構成となっている。

その一方で通信路でのエラーの発生を考慮して、FMO (Flexible Macroblock Ordering), ASO (Arbitrary Slice Order), 冗長スライスといったエラー耐性向けの符号化ツールが採用された構成となっている。

\* デジタル家電開発本部 次世代商品開発センター 第3開発室

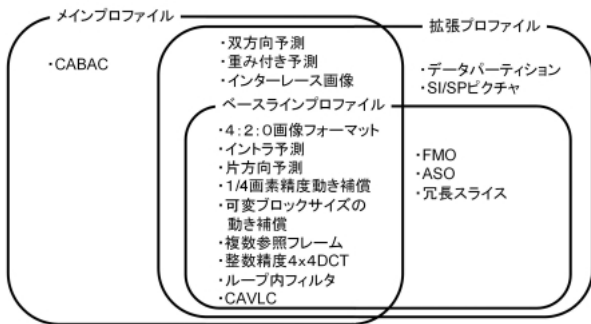


図1 プロファイルと符号化ツール  
Fig. 1 Profiles and coding tools.

(2) メインプロファイル

主にデジタルTV放送やデジタルビデオレコーダ、パッケージメディア等での利用を想定したプロファイルである。

これらのアプリケーションでは、誤り訂正符号等の利用によって実質的にデータ中のエラーの存在が無視できる。したがって、他のプロファイルでは採用されているエラー耐性向けの符号化ツールが省略された構成となっている。

その反面、最大限符号化効率を重視した構成となっており、エントロピー符号化用の符号化ツールとして、CAVLCより複雑度は高いが符号化効率の高いCABAC (Context-based Adaptive Binary Arithmetic Coding) も利用可能な構成となっている。

(3) 拡張プロファイル

主にインターネットでのストリーミングアプリケーションでの利用を想定したプロファイルである。ストリーミングアプリケーションでは、オフラインで符号化を行うことができるため、符号化にリアルタイム性は要求されない。したがって、ベースラインプロファイルに比べ、双方向予測等の符号化効率を重視したツールを利用可能な構成となっているが、PCでの利用が想定されるため、複雑度が高くデコーダで専用ハードウェアの必要性が見込まれるCABACは省略された構成となっている。

またJVTでは現在、FRExt (Fidelity Range Extensions) と呼ばれる、高解像度・高画質向けの拡張についての標準化作業が進められており、HD-DVD, Blu-ray Discでの採用が見込まれている。

FRExtでは、既存のメインプロファイルをベースに、

- (1) 4:2:2, 4:4:4 画像フォーマットのサポート
- (2) 10,12ビットサンプリング画像のサポート
- (3) 整数精度 8x8DCT
- (4) 量子化マトリクス

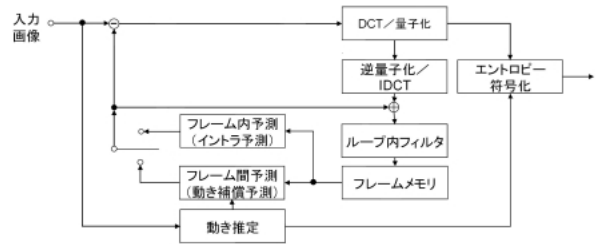


図2 H.264/AVC エンコーダの基本構成  
Fig. 2 Basic structure of H.264 encoder.

等の機能拡張したプロファイルを、上述の3つのプロファイルに加えて新設することが決定している。

2. H.264/AVC の技術的特長

図2は、H.264/AVCにおけるエンコーダの基本構成を示す図である。図2から明らかなように、H.264/AVCの基本アルゴリズムは、従来のMPEG-2/4等の動画像符号化方式と同様の動き補償予測とDCTを用いたハイブリッド符号化である。

その一方で、個々の符号化ツールには様々な改良が加えられており、H.264/AVCの高い符号化効率は、これらの改良の積み重ねによって実現されている。表1に従来の動画像符号化方式であるMPEG-4 (Simple Profile) とH.264/AVCの各符号化ツールの比較を示す。また、MPEG-4からの改良点を中心にH.264/AVCの主な符号化ツールの詳細について紹介する。

2・1 フレーム間予測 (動き補償予測)

フレーム間予測についてはMPEG-4に比べ、さまざまな点で改良が加えられている。主な違いは以下の4点である。

(1) 可変ブロックサイズ動き補償

MPEG-4では、1マクロブロック (16x16画素) 単位、またはマクロブロックを4分割した8x8画素単位で動き補償を行うのに対し、H.264/AVCでは、1マクロブロック (16x16画素), 16x8画素, 8x16画素, 8x8画素, 8x4画素, 4x8画素, 4x4画素単位の計7種類のブロックサイズで動き補償を行うことができる。

小さなブロックで動き補償を行うことで、単位面積当たりの予測誤差を小さく抑えることができる。その反面、符号化する動きベクトル数が増加し、符号量が増加する可能性もある。

(2) 予測精度

MPEG-4の1/2画素精度に対し、H.264/AVCでは1/4画素精度に予測精度が向上している。

(3) フィルタ処理

予測画像生成のためのフィルタ処理において、

表1 MPEG-4 と H.264/AVC の符号化ツールの比較

Table 1 Comparison of coding tools of MPEG-4 and H264/AVC.

符号化ツール	MPEG-4 (Simple Profile)	H.264/AVC	
フレーム間 符号化	ブロックサイズ	16x16画素または8x8画素の2パターン	最大16x16画素, 最小4x4画素の7パターン
	予測精度	1/2画素精度	1/4画素精度
	フィルタ処理	2タップ	6タップ
	参照フレーム	1フレームのみ	複数フレーム可
フレーム内符号化	DCT係数レベルの予測	画素レベルの予測	
直交交換	実数精度8x8DCT	整数精度4x4DCT 整数精度8x8DCT (FRExtのみ可)	
デブロックフィルタ	ポストフィルタ (規格外)	ループ内フィルタ	
エントロピー符号化	3次元VLC	CAVLCまたはCABAC	

MPEG-4の2タップフィルタ処理に対し、H.264/AVCでは6タップフィルタ処理が用いられる。高次数のフィルタ処理を用いることで、予測画像における高周波成分の再現性が高くなっている。

#### (4) 複数参照フレーム

MPEG-4では片方向予測の場合、1フレームのみが参照フレームとして利用可能である。一方H.264/AVCでは、複数の参照フレームを用いることが可能である。

例えば、ある物体が別の物体の陰に隠れた後、再び現れるようなシーンでは、参照フレームが複数存在することで、符号化対象のブロックにより近い画像を見つけられる可能性が高くなり、符号化効率の向上が期待できる。その一方で、複数の参照フレームを保持するためのフレームメモリが必要となる欠点も存在する。

このようにH.264/AVCにおけるフレーム間予測は、MPEG-4に比べ自由度の高い設計となっている。したがってフレーム間予測における符号化効率の改善は

は、ブロックサイズ、参照フレーム等の多様な符号化パラメータの組み合わせから適切な組み合わせを選択することが必要になると共に、最適なパラメータの決定のために膨大な演算が必要である。

## 2・2 フレーム内予測 (イントラ予測)

MPEG-4では、DCT係数における直流成分と一部の交流成分の予測によりフレーム内予測が実現されていたのに対し、H.264/AVCでは、隣接ブロックからの画素レベルでの予測によりフレーム内予測が実現されている。

輝度信号の場合は図3に示すように、4x4画素のブロックに対し、9通りの予測モードが定義されている。例えば図3(a)の場合、予測対象の4x4画素ブロックの上側に隣接する復号済みの画素データを用いて、垂直方向の予測を行い、予測画像を生成する。この予測モードは、予測対象の画像領域に垂直方向エッジがある場合に有効な予測モードである。図3(a)以外の予測モードも同様にそれぞれが特定方向のエッジ

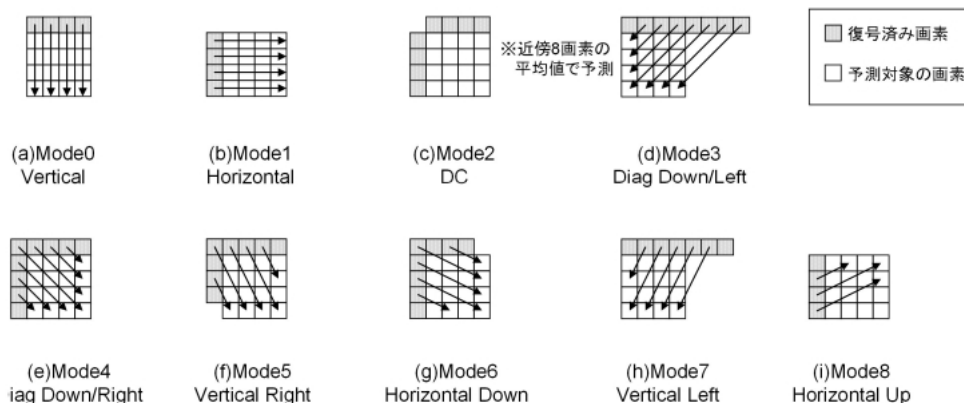


図3 イントラ予測モード (輝度信号 4x4 ブロック)

Fig. 3 Intra prediction modes (4x4 block of luma sample).

に対して有効な予測モードとなっており、隣接ブロックの復号済み画素を基に予測画像を生成する。また輝度信号の場合は、4x4画素ブロックに対してだけでなく、16x16画素ブロックに対しても同様に4通りの予測モードが定義されており、計13通りの予測モードが存在する。したがって、これら13通りの予測モードの中から最適な予測モードを選択し符号化に用いる。また色差信号については、8x8画素ブロックに対し、4通りの予測モードが定義されており、輝度信号とは独立に予測モードを符号化することが可能である。

このようにH.264/AVCにおけるフレーム内予測もフレーム間予測同様MPEG-4に比べ、多様な予測モードを備えることで自由度が高い設計となっており、フレーム間予測における符号化効率の改善には、エンコーダが適切な予測モードを選択することが必要になる。

### 2・3 直交変換

MPEG-4では、直交変換に実数精度8x8DCTを用いたのに対し、H.264/AVCでは、整数精度4x4DCTが用いられる。

変換ブロックサイズが小さくなったことで、ブロックノイズ、モスキートノイズが目につきにくいといった利点がある。その反面、周波数分解能が低くなったため、広範囲に渡って相関の高い画像では、符号化効率が低下する場合がある。したがって、HDTV等の高解像度の画像を主なターゲットとするFRExtでは、整数精度4x4DCTに加え、整数精度8x8DCTも選択可能である。

### 2・4 デブロッキングフィルタ

MPEG-4では、デブロッキングフィルタは採用されていない符号化ツールである。しかしながら、規格外の処理として、デコーダのためのポストフィルタ処理で一般に用いられる。

一方H.264/AVCでは、ループ内フィルタとして符号化ループに組み込まれている。デブロッキングフィルタが符号化ループに組み込まれる利点は、ブロックノイズの除去された画像を参照画像として用いることができることである。参照画像からブロックノイズが除去されることにより、動き補償予測による予測誤差からブロックノイズの影響が除かれ、符号化効率が向上する。

なおフィルタ強度は量子化パラメータやDCT係数の有無、動きベクトルの大きさ等、符号化の状況に応じて決定され、適応的なフィルタ処理が行われるため、ループ内フィルタの占める演算量が大きいという欠点が存在する。実装方法にも依存する問題ではあるが、デコーダの場合、全体の50%程度の演算量を占める場合も存在する。

### 2・5 エントロピー符号化

MPEG-4では、DCT係数の符号化にラスト、ラン、レベルの3つの要素に対し、1つの固定的な符号語を割り当てる3次元VLCが用いられる。一方、H.264/AVCでは、CAVLCとCABACの2種類のエントロピー符号化のいずれか一方を用いて符号化を行う。どちらの方式を用いた場合でも、周囲のブロックの状況に応じて適応的な符号語の割り当てを行うため、MPEG-4に比べて高い符号化効率を実現することができる。

なおCABACでは算術符号化が用いられるため、CAVLCに比べ効率的な符号割り当てが可能である。したがって、CABACを使用した場合、CAVLCに比べ約10%の符号化効率の改善が可能であるが、その一方で、膨大な演算量が必要になるといった欠点がある。

### むすび

以上紹介したように、H.264/AVCでは、従来の動画像符号化方式と同様の基本アルゴリズムを採用しながら、個々の符号化ツールの改良により高い符号化効率を実現する動画像符号化方式となっている。その一方で、符号化/復号に必要な演算量は、従来の動画像符号化方式に比べ、数倍から数十倍の演算量が必要になっている。特にエンコーダでは多くの演算量を必要とするため、多様な符号化パラメータから最適な符号化パラメータを効率良く求めるアルゴリズム開発が重要である。

### 参考文献

- 1) ITU-T Recommendation H.264: "Advanced Video Coding for generic audiovisual services" (2003).
- 2) ISO/IEC 14496-10: "Coding of audiovisual objects - Part 10: Advanced Video Coding" (2003).

(2004年9月17日受理)