

情報源符号化部 H.265 | MPEG-H HEVC 規格の概要

Ver. 1.0

2013年5月23日

社団法人 電波産業会
デジタル放送システム開発部会
映像符号化方式作業班

目 次

1. 標準化経緯
 2. 方式の概要
 3. 符号化データ構造
 - 3.1 ビットストリーム構造
 - 3.2. ピクチャとサブピクチャ
 - 3.2.1. タイル
 - 3.2.2. WPP
 4. 整数変換と量子化
 - 4.1 整数精度変換
 - 4.2 量子化
 5. インター予測
 - 5.1 複数参照フレーム
 - 5.2 インターPB サイズ
 - 5.3 小数画素精度動き補償
 - 5.4 動きベクトルの予測
 - 5.5 マージモード
 6. イントラ予測
 - 6.1 輝度信号のイントラ予測符号化
 - 6.2 色差信号のイントラ予測符号化
 7. エントロピー符号化
 - 7.1 CABAC 同期処理
 - 7.2 トランケーテッド・ライス二値化処理
 8. ループ内フィルタ
 - 8.1 デブロッキング・フィルタ
 - 8.2 画素適応オフセット
 9. プロファイルとレベル
 - 9.1 プロファイル
 - 9.2 レベル
- 参考文献

1. 標準化経緯

H.265 | MPEG-H HEVC(High Efficiency Video Coding、以下 HEVC)は、H.264 | MPEG-4 AVC(Advanced Video Coding、以下 AVC)に対して高効率の符号化を可能とする映像圧縮符号化方式である。標準化作業は、ITU-T (International Telecommunications Union - Telecommunication Standardization Sector : 国際電気通信連合 電気通信標準化部門) の Working Party(WP) 3/16 と ISO/IEC (International Organization for Standardization : 国際標準化機構 / International Electrotechnical Commission : 国際電気標準会議) の MPEG (Moving Picture Experts Group : 動画像符号化専門家グループ)によって2010年1月に設立されたJCT-VC(Joint Collaborative Team on Video Coding Standard Development)において進められた。

2つの呼称は各機関におけるこれまでの映像圧縮符号化技術の勧告もしくは標準化作業を受け継いだものである。「H.265」は1990年のH.261に始まるITU-Tの映像圧縮符号化技術に係る勧告を指し、一方、「MPEG-H HEVC」はISOとIECが合同で設立した委員会JTC1 (Joint Technical Committee 1 : 第1合同技術委員会)の作業グループの通称でもあるMPEGの名で標準化された映像圧縮符号化方式を指す。MPEGでは、MPEG-Hシリーズ(ISO/IEC 23008)のビデオパート(part2)として位置づけられる。

ITU-T、MPEGの共同作業によりAVCの規格化が行われた後、WP3/16傘下の映像符号化専門家グループVCEG(Video Coding Experts Group)では、2004年頃より、AVCの符号化性能を超える新しい符号化技術の検証が開始された。この活動は"Key Technical Areas" (KTA) と呼ばれる。MPEGにおいても、2008年頃より新しい映像符号化技術の標準化に向けた要素技術検証(Exploration)が開始された。これら技術検証においては、当初より次世代のアプリケーションに備えてHD解像度以上の高精細な映像信号がターゲットに据えられた。2009年4月、MPEGはAVCを超える圧縮性能を有する技術のエビデンスを確認することを目的として技術公募(Call for Evidence)をかけ、その結果を受けて新たな標準方式の規格化に向けた活動を本格化させた。その過程において、KTAを推進するITU-T VCEGとも活発な意見交換が行われ、新標準方式に対する要求条件において、UHDTV映像への対応が共通の重要スコープとして位置づけられるに至った。

2010年1月、ITU-T WP3/16、MPEGはJCT-VCの結成と、JCT-VCによる新標準方式の規格化プロジェクトの名称をHEVCとすることについて合意し、正式に技術公募(Call for Proposals)を発行した。同年4月の第1回JCT-VC会合において提案された符号化技術を評価し、UHDTVを含む評価映像に対してAVCの2倍の圧縮性能を実現できる見通しが確認された。その後、同年10月に標準化の出発点となる作業ドラフト(Working Draft)仕様と参照ソフトウェアモデルが策定され、2012年1月の標準化委員会草案(Committee Draft)を経て、2013年1月の第12回会合にて最終規格草案(Final Draft International

Standard)を発行して、HEVC 第1版の標準化を完了した。第1版の仕様では、4:2:0/8 ビットの映像信号を対象とし放送、通信、パッケージ等幅広いアプリケーションに適用可能な Main プロファイル、高品質な 10 ビット映像信号に対応する Main 10 プロファイル、4:2:0/8 ビットの静止画像信号の圧縮に対応する Main Still Picture プロファイル、の3つのプロファイルが策定されている。

この第1版発行に引き続いて、より一層の高画質化やスケーラブル符号化を目的とした追加検討が行われている。高画質化の方向については、12 ビットあるいは14 ビットでサンプリングした 4:2:2 や 4:4:4 の映像信号を業務用途の制作・伝送の目的で圧縮する要求条件が考慮され、Range Extensions(RExt)の名称で新たなプロファイルの策定を進めている。この追加仕様については2014年1月の標準化完了を目指している。一方、スケーラブル符号化の方向に関しては、既存サービスとの互換性を維持しながら新しい映像サービスを提供したり、モバイル回線や光回線など多様なネットワークへコンテンツを効率的に流通させることを目的として、1本のビットストリームで複数の品質の映像信号を符号化できる仕様を策定中である。スケーラブル符号化機能については2014年6月の標準化完了を目指している。

2. 方式の概要

図2-1にHEVCエンコーダの概略構成を示す。基本的なブロック構成はMPEG-2やAVCに類似しているが、予測・変換ブロックサイズの拡張・適応化をはじめ、符号化効率を上げるための種々の工夫がなされている。図2-1に描かれた吹き出し項目は、主に従来方式にはない新たな要素技術である。

本章では、HEVC符号化を構成する主な要素技術について、既存のデジタル放送で採用されているMPEG-2、AVCなどの従来の符号化方式と比較しながら簡単に説明し、さらに詳細については3章以降に記述する。

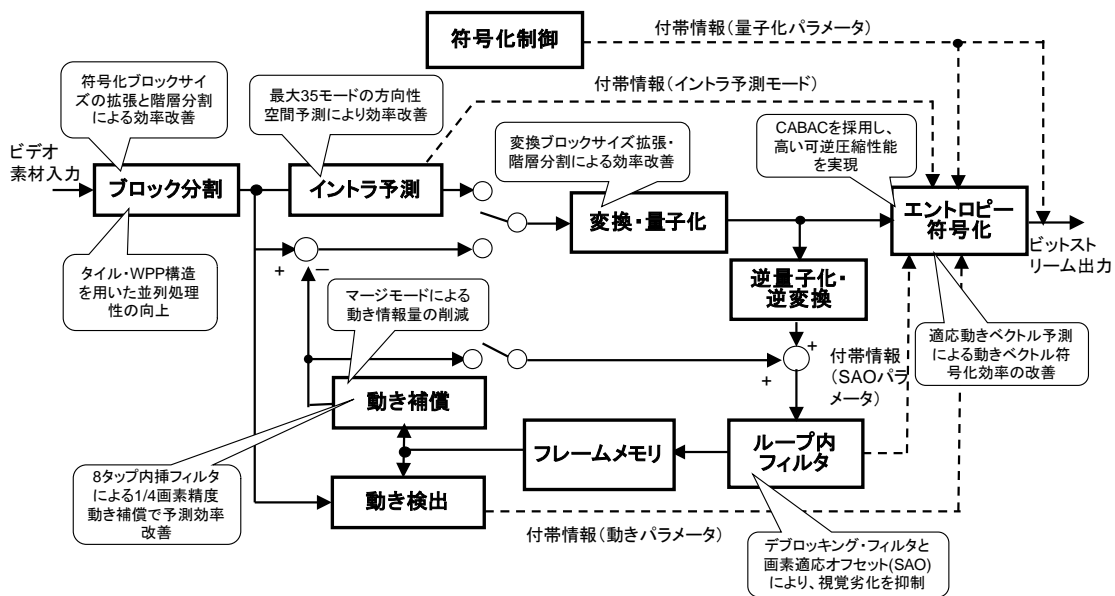


図 2-1 HEVC エンコーダの概略構成

-符号化・予測・変換ブロックサイズの拡張と適応化

HEVC では各ピクチャの符号化処理を、CTU(Coding Tree Unit)と呼ばれる正画面素ブロックの単位で行う。CTUは MPEG-2 や AVC におけるマクロブロック (Macroblock) に相当する処理単位で、輝度信号ブロックと対応する色差信号ブロックから構成される。CTU のサイズはマクロブロックのように固定ではなく、シーケンス(3.1 節参照)ごとにエンコーダが選択・シグナリングできる(最大 64x64 画素ブロック)。

イントラ・インター予測、変換・量子化、エン트로ピー符号化などの一連の符号化処理は、CTU を再帰的に四分木分割することによって得られる CU(Coding Unit)の単位で行う。CTU 内を多様な大きさの CU に分割して符号化を行うことができるため、映像信号の局所的な特性に対して符号化処理を適応化することが可能となる。例えば、複雑な動きや絵柄が存在する領域では CU のサイズを小さくして動きベクトルやイントラ予測モードなどの予測パラメータに符号量を多く配分することで予測性能を高め、平坦な絵柄や一様な動きの領域については CU のサイズを大きくして予測パラメータの符号量を抑制するなどの制御をおこなうことで、従来方式に比べて少ない符号量でも高画質化を図ることが可能である。一方、CU サイズのバリエーションや予測パラメータの選択肢が増えたことによって符号化処理の演算量が増加するため、高精細な映像信号をリアルタイムで高効率に符号化するためには、エンコーダの設計に従来以上の工夫を要する。

CU の符号化処理に際しては、イントラ・インター予測については、CU 内を PU(Prediction Unit)と呼ばれるブロックに分割して行われる。予測差分信号の変換・量子化処理は、TU(Transform Unit)と呼ばれるブロック単位で行われる。

TUもCTU内のCU分割と同様、CU内部で再帰的に四分木分割を行うことが可能である。なお、CTU、CU、PU、TUを構成する各色成分信号のブロックをそれぞれCTB(Coding Tree Block)、CB(Coding Block)、PB(Prediction Block)、TB(Transform Block)と呼ぶ。図2-2にCTU、CU、PU、TUの関係、およびユニット(Unit)とブロック(Block)の関係を示す。

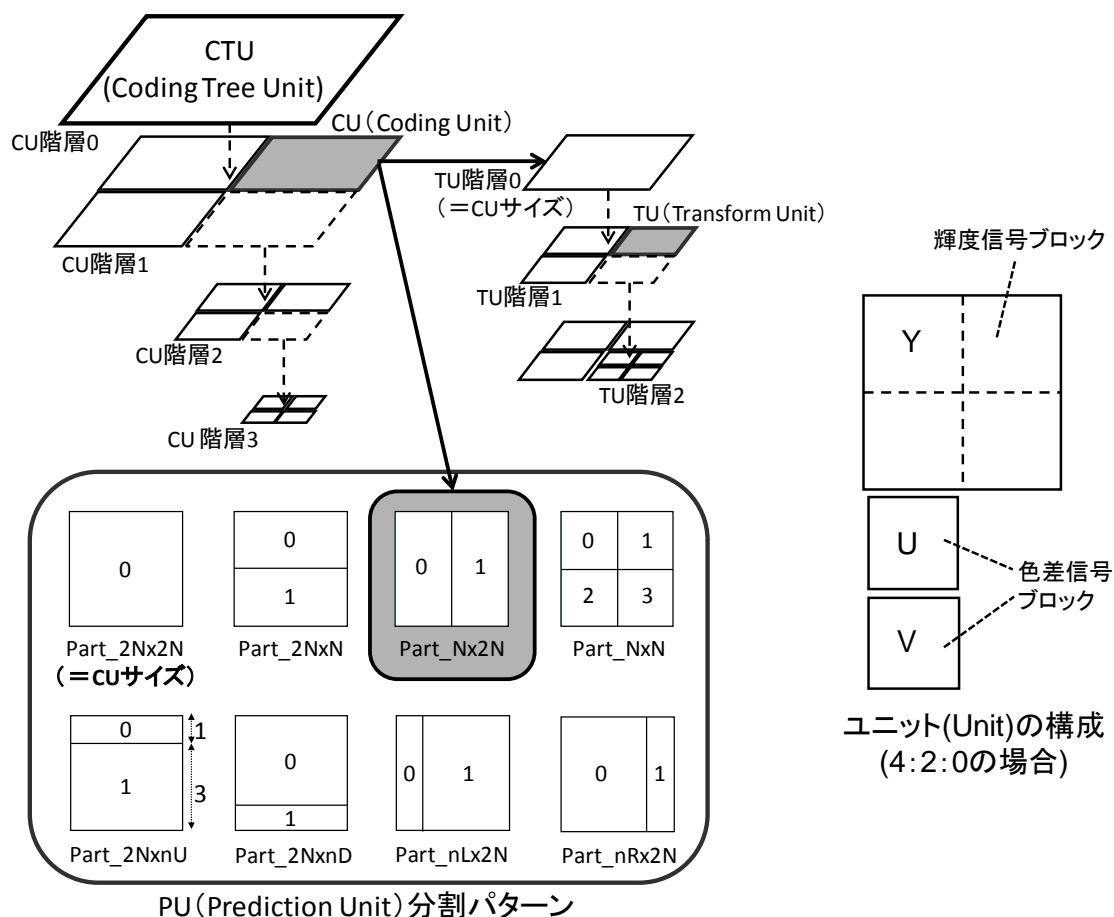


図 2-2 HEVC のブロック構造

-イントラ予測

AVCと同様のコンセプトで、PBごとにピクチャ内の近傍の復号済みの画素（予測参照画素）を用いて空間予測を行い、予測差分信号を変換符号化する。輝度信号では、4x4~64x64画素ブロックの各PBサイズに対し、いずれのサイズにおいても35種類の予測モードの中から予測効率に優れたモードを選択可能である。色差信号については、輝度信号と予測モードを共通化するモードを含め、5種類の予測モードを輝度信号とは独立に選択できる。なお、イントラ予測効率をさらに高める工夫として、予測参照画素に対する適応的な平滑化フィルタ処理、DC予測や水平・垂直方向予測等のシンプルな予測モードに対する予測参照画素と予測画像との間に生じる不連続性を抑制するための境界フィルタ処理、斜め方向の方向性予測に対する疑似輪郭抑制フィルタ処理等が導入されて

いる。

-インター予測

インター予測においては、CBサイズを上限に、最小8x4ないし4x8から最大64x64画素のPBサイズでフレーム間動き補償予測を行うことができる。AVCと同様、フレームメモリ内に複数枚の参照画像を格納して、PBごとに参照画像を特定して予測を行うことができる。レベルによって異なるが最大16フレームを参照画像として用いることができる（HDTVに対応するレベル4の場合、最大4フレームを使用可能）。複数の参照画像を用いて多様なブロックサイズで動き補償予測を行う処理は、演算量・メモリ量・メモリバンド幅の増大要因となるが、AVCに比べて予測差分信号の符号量を大きく削減することが可能になる。

また、動き補償予測の精度は、AVC同様1/4画素精度までとり、小数画素位置の予測画像の生成に8タップ内挿フィルタを採用して性能を高めたほか、重み付き予測も採用されている。

その他、動きベクトルの予測値を空間・時間近傍の動きベクトルから適応的に選択することによる動きベクトルの効率的な符号化手法や、近傍から動きベクトルをはじめとする予測パラメーター式を流用して予測を行うことができるマージモードの導入などの工夫により、インター予測に関するサイド情報の符号量を低減している。

-変換

MPEG-2では8x8離散コサイン変換、AVCでは4x4と8x8の2通りのブロックサイズの整数変換を用いていたのに対して、HEVCでは、CBサイズを上限として、最小4x4から最大32x32画素のTBサイズを用いた整数変換を選択することができる。16x16や32x32といった大きなブロックサイズでの変換を使用することによって、予測差分信号に残存する空間冗長度を効率的に削減できる。また、HEVCでは、変換処理として離散コサイン変換だけでなく、離散コサイン変換や変換スキップモード等も採用されている。

-エントロピー符号化

HEVCでは、CTU以下の符号化データに対しては、全てのプロファイルにおいて圧縮効率の高いCABAC (Context-based Adaptive Binary Arithmetic Coding)が採用されている。CABACはAVCのMain, Highプロファイルに採用されているが、HEVCではコンテキスト適応化処理の簡易化を行ったり、並列処理との親和性を高めることによりスループットの向上が図られている。シーケンス・ピクチャレベルのパラメータセットやスライスヘッダ等の上位シンタックスに対しては、固定長符号ないし指数Golomb符号(Exponential Golomb Coding)が用いられる。

-デブロッキング・フィルタ

AVC同様、ループ内フィルタの一つとして、エンコーダおよびデコーダ内において再構成画像作成時にデブロッキング・フィルタが用いられる。本フィルタはPUやTUの境界画素に対して実行するため、大きなPU・TUサイズが選

扱される箇所では、AVC に比べて演算量が低減される。

-画素適応オフセット

第 2 のループ内フィルタとして、符号化歪みが重畳された画像信号波形を、画素単位のオフセット値加算によって補正する画素適応オフセットが導入されている。オフセットパラメータは CTU ごとにエンコーダで決定され、ビットストリームに多重してデコーダへ伝送される。デブロッキング・フィルタがブロック境界の主観歪みを低減するのに対して、画素適応オフセットはブロック内部のリングング歪みの除去に効果がある。

表 2-1 に、MPEG-2 の Main プロファイル、AVC の High プロファイルと、HEVC の Main プロファイルの要素技術比較表を示す。

表 2-1 要素技術比較表

要素技術	MPEG-2 (Main プロファイル)	AVC (High プロファイル)	HEVC (Main プロファイル)
イントラ予測	なし	4×4 ブロックで 9 種類、 8×8 ブロックで 9 種類、 16×16 ブロックで 4 種類の 予測モード	4×4～64×64 ブロックで 35 種類の 予測モード
インター予測	16×16 または 16×8 ブロック 直前および直後の フレームより予測 1/2 画素精度予測	16×16 を 7 種類のモード で動き予測 最大 DPB ^{注1} サイズ=16 1/4 画素精度予測	8×4/4×8～64×64 ブロックで動き予測 最大 DPB ^{注1} サイズ=16 1/4 画素精度予測
変換	8×8 離散コサイン 変換	4×4 または 8×8 の 整数変換	4×4～32×32 ブロックで 整数変換
エントロピー 符号化	ハフマン符号	コンテキスト適応可変長 符号(CAVLC)または コンテキスト適応算術符号 (CABAC)	コンテキスト適応算術 符号(CABAC)
デブロッキング ・フィルタ	なし	あり	あり
画素適応 オフセット	なし	なし	あり

注1) 復号ピクチャバッファ(Decoded Picture Buffer)。5.1 節参照。

この章のまとめとして最後に、AVC と比較して、HEVC のメリットとデメリットについて記述する。

- HEVC のメリット

- 新規方式による生成符号量の削減
多様な予測・変換ブロックサイズ、インター/イントラ予測や動きパラメータ削減技術などの新規方式により生成符号量が削減できる。
- 符号化歪みの少ない復号画像

再構成画像作成時にデブロッキング・フィルタや画素適応オフセットが適用されることにより、符号化歪みの少ない復号画像が得られる。

- HEVC のデメリット

- 演算量の増大
高精度のインター/イントラ予測および画素適応オフセットなどの新規方式により、エンコーダの演算量が増大する。(デコーダの演算量は AVC と同等のオーダーに抑えられている)

3. 符号化データ構造

3.1 ビットストリーム構造

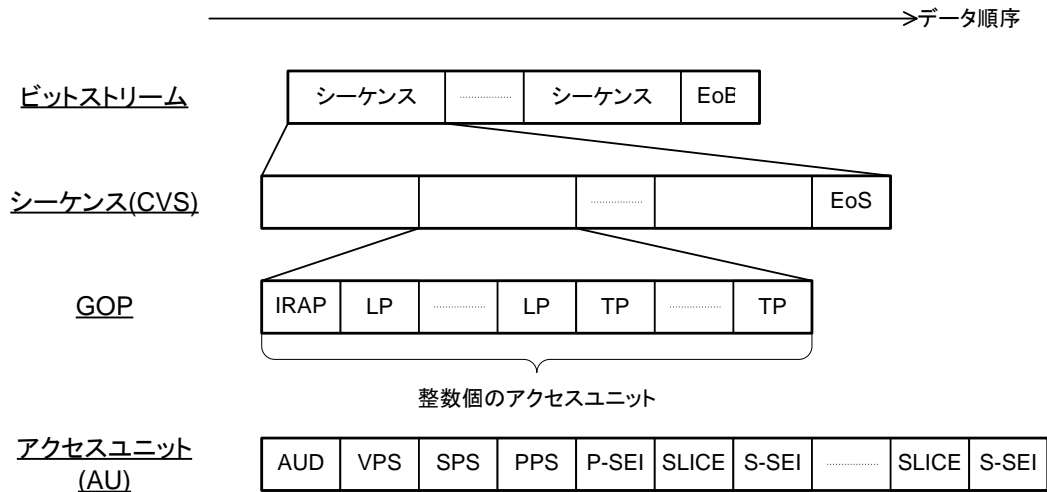


図 3-1 HEVC ビットストリームの構造

1つのHEVCビットストリームは、複数のシーケンス(Coded Video Sequence, CVS)と、End of sequence NAL unit (EoB)から構成される。CVSは、複数のGroup of Pictures (GOP)¹と、End of Sequence NAL unit (EoS)から構成される。GOPは、IRAP (Intra Random Access Point) Access Unit (AU)を先頭とし、後続する複数のLP (Leading Picture) AU及び、TP (Trailing Picture) AUから構成される。

IRAP AUは、ランダムアクセスポイントとなるAUである。LPAUは、直前のIRAP AUよりも表示時刻が早いAUである。LPAUは、参照ピクチャとしてIRAP AUよりも復号時刻が早いAUを参照することが可能であり、この時、デコーダがIRAP AUから復号を開始した場合に正常復号されない。他方、TP AUは表示時刻がIRAP AUよりも遅いAUである。HEVC規格では、IRAP AUから復号を開始した場合にTP AUの正常再生を保証する。逆に言うと、TP AUの正常再生を保証しないIntra AUはIRAP AUとはならない。

¹ HEVC規格上ではGOPという概念を導入していないが、本資料では説明上GOPと呼ぶ。

AU は、複数の Network Abstraction Layer Unit (NALU) から構成される。AU は必ず 1 つ以上の Slice Segment NALU を含む必要がある。表 3-1 に、HEVC で定義される NALU 種別を示す。なお、AU の種類は、内包する Slice NALU (VCL NALU) によって区別する。更に、AVC とは異なり AU が参照ピクチャであるかどうかは、NALU 種別によって識別する。

表 3-1 H.265 の NALU 種別(nal_unit_type) (Reserved は除く)

nal_unit_type 名称	種別	説明
TRAIL_N TRAIL_R	VCL	Trailing Picture (TP) Slice “_N”は非参照ピクチャ、“_R”は参照ピクチャを意味する。
TSA_N TSA_R	VCL	Temporal Sub-layer Access (TSA) Slice 時間方向階層符号化ビットストリームにおける、復号レイヤ切り替えが可能な TP AU (図 2 参照)。TP AU の一種。
STSA_N STSA_R	VCL	Step-wise Temporal Sub-layer Access (STSA) Slice TSA AU の更に特殊な場合 (図 3-2 参照)。
RADL_N RADL_R	VCL	Random Access Decodable Leading (RADL) Slice 直前の IRAP AU から復号を開始した場合にも正常復号が可能な LP AU (IRAP AU よりも復号時刻が先のピクチャを参照しない)
RASL_N RASL_R	VCL	Random Access Skipped Leading (RASL) Slice 直前の IRAP AU から復号を開始した場合に正常復号できない LP AU (IRAP AU よりも復号時刻が先のピクチャを参照する)
BLA_W_LP BLA_W_RADL BLA_N_LP	VCL	Broken Link Access (BLA) Slice IRAP AU の一つであり、ビットストリーム編集により、POC の連続性が保証されない AU “_W_LP”は、RADL と RASL の両 LP が存在しうること、“_W_RADL”は RADL のみ存在しうること、“_N_LP”は LP が存在しないことを意味する。
IDR_W_RADL IDR_N_LP	VCL	Instantaneous Decoding Refresh (IDR) Slice IRAP AU の一つであり、POC が 0 にリセットされる AU LP として RASL を含まない
CRA_NUT	VCL	Clean Random Access (CRA) Slice IRAP AU の一つ。 CRA AU を先頭とするビットストリーム編集を行った場合、CRA AU を BLA AU に置き換える。
VPS_NUT	non-VCL	Video Parameter Set (VPS) HEVC で新しく導入されたパラメータセット 階層符号化時に、各レイヤの共通情報や関係情報を記述する。 非階層符号化時にも必須とされるが、デコーダは読み捨てる。

nal_unit_type 名称	種別	説明
SPS_NUT	non-VCL	Sequence Parameter Set (SPS) AVCと同様に、シーケンスで共通なパラメータを記述する。
PPS_NUT	non-VCL	Picture Parameter Set (PPS) AVCと同様に、ピクチャで共通なパラメータを記述する。
AUD_NUT	non-VCL	Access Unit Delimiter (AUD) AVCと同様に、AU 境界を示す。
EOS_NUT	non-VCL	End of Sequence (EoS) シーケンス境界を示す。
EOB_NUT	non-VCL	End of Bitstream (EoB) ビットストリーム境界を示す。
FD_NUT	non-VCL	Filler Data (FD) フィラーデータ。
PREFIX_SEI_NUT SUFFIX_SEI_NUT	non-VCL	Supplemental Enhancement Information (SEI) AVC の SEI とは異なり、AU 内先頭 VCL NALU の前に出現する Prefix SEI (P-SEI) と、VCL NALU の後に出現する Suffix SEI (S-SEI) の二つを定義する。これらの区別は、SEI の内容を Slice の復号前に参照するか、もしくは復号後に参照するかに基づく。

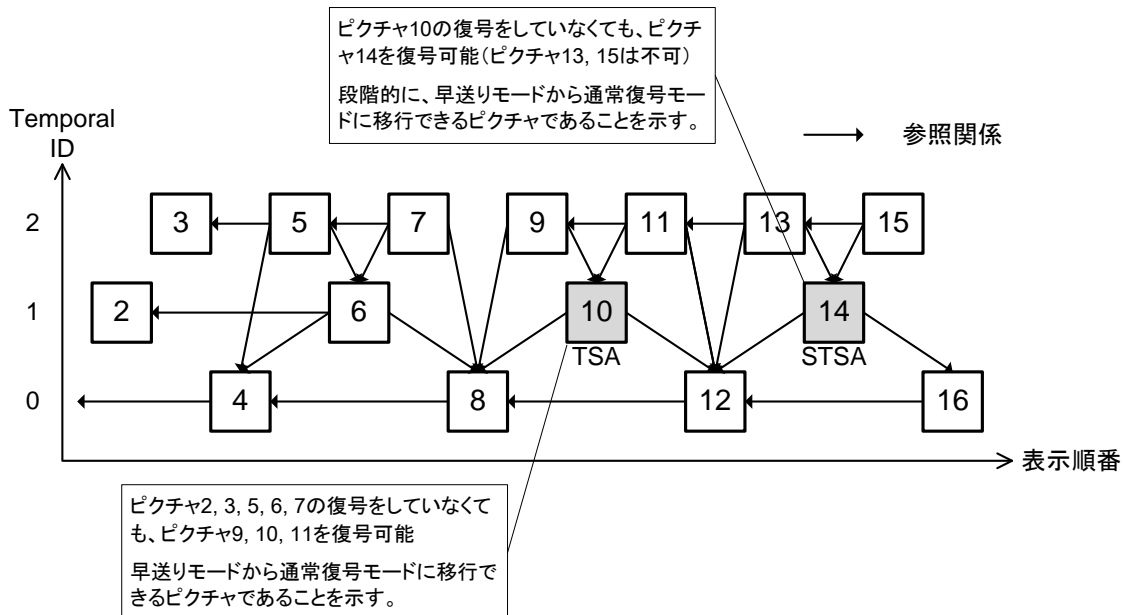


図 3-2 TSA AU, STSA AU の説明

3.2. ピクチャとサブピクチャ

AVC 規格と同様に、符号化ピクチャは複数のスライス (サブピクチャ) に分割される。スライスは複数の CTU から構成される (図 3-3)。

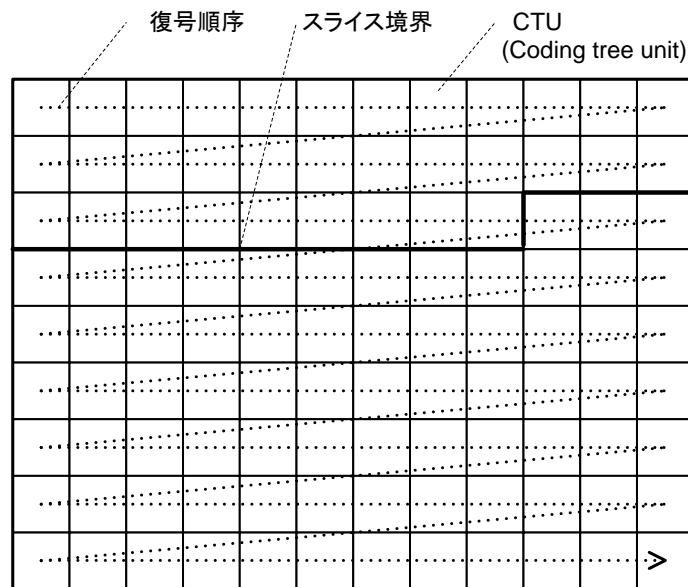


図 3-3 ピクチャの構造

スライスには、AVC 規格相当のスライス（他のスライスと独立して復号可能）と、従属スライスの 2 種類がある。従属スライスは、CABAC を含め、復号処理が（同一スライス内の）別従属スライスに依存し、スライスヘッダの情報は簡略化される。

3.2.1. タイル

タイルは、図 3-4 のように画面内を縦・横方向に CTU 単位で分割し、各分割領域（タイル）を独立して符号化・復号するための仕組みである。AVC のスライスグループと同様のコンセプトであるが、各タイル先頭でのスライスヘッダ挿入が不要、という特徴がある。スライス境界でのループ内フィルタの有効・無効化は PPS で設定可能である。各 CTU の符号化・復号順序は、タイル内ラスタスキャン順となる。また各タイルの先頭バイト位置のリストをスライスヘッダに記述することで、デコーダ側での並列復号を容易にしている。

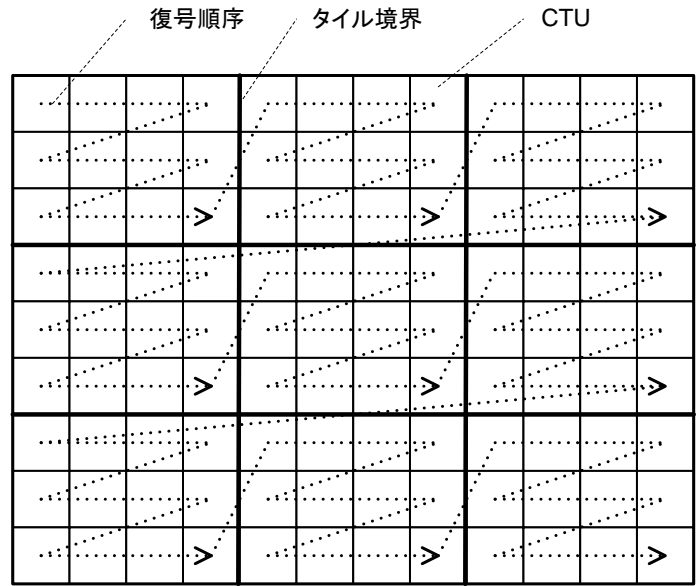


図 3-4 タイル構造

3.2.2. WPP

WPP (Wavefront parallel processing)は、符号化効率の低下を抑制しつつ、各 CTU ラインを並列符号化・復号するための仕組みである。

“Wavefront”の名称通り、各 CTU ラインの符号化・復号処理を、それぞれ 2 CTU 時間だけシフトさせて実行することを想定している (図 3-5)。これにより、CTU ライン間の従属性を維持しつつ並列処理が可能になる。

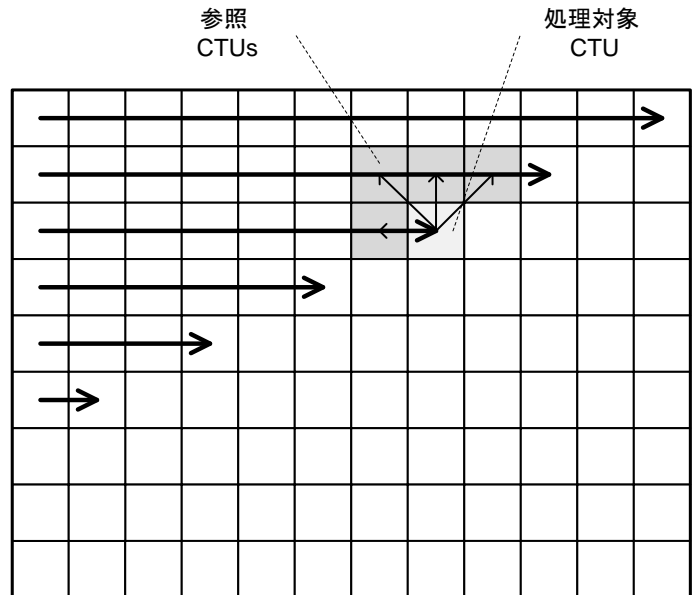


図 3-5 WPP のコンセプト

各 CTU ラインの CABAC 符号化を独立にするためには、CTU ラインの終端で CABAC をリセットする必要がある。コンテキスト初期化による符号化効率低下

を抑制するために、各 CTU ラインの先頭にて、上 CTU ラインの右から 2 番目の CTU 処理後のコンテキストをコピーする仕組みが導入されている(図 3-6)。

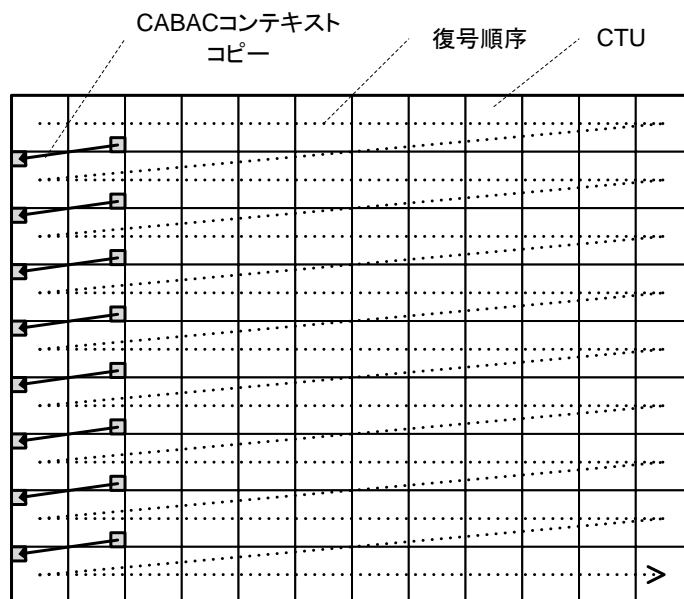


図 3-6 WPP 動作

タイルと同様に、各 CTU ラインの先頭バイト位置のリストをスライスヘッダに記述することで、デコーダ側での並列復号を容易にしている。なお、タイルと WPP とは排他的に使用される。

4. 整数変換と量子化

4.1 整数精度変換

MPEG-2 では、 8×8 の実数精度 DCT (Discrete Cosine Transform : 離散コサイン変換) によって映像信号を周波数領域に変換 (直交変換) する。しかし、実数精度の演算はハードウェアごとに演算精度が異なることから環境によって再生映像が完全一致しないという問題が生じた。そこで AVC では、 4×4 および 8×8 の整数精度 DCT が用いられた。直交変換が整数精度であるため、浮動小数点演算に似た演算の丸め処理がなく、環境に依存した変換の不一致が生じず、かつ高速に処理できる。

HEVC における変換処理もまたこれを踏襲し、整数精度の直交変換を用いることにより環境依存の演算誤差は生じない。変換処理では整数精度 DCT に加え、新たに整数精度 DST (Discrete Sine Transform : 離散サイン変換) と変換スキップが用いられている。これらのモードは、イントラ/インターの予測方法やブロックサイズ、輝度/色差の信号の別によりそれぞれのモードを適応的に選択される。

整数精度 DST はイントラ CU において輝度信号の TB サイズが 4×4 の場合にのみ選択可能な変換である。他の TB サイズおよび色差信号に対しては適用さ

れない。変換マトリックスを以下に示す。

$$\begin{bmatrix} 29 & 55 & 74 & 84 \\ 74 & 74 & 0 & -74 \\ 84 & -29 & -74 & 55 \\ 55 & -84 & 74 & -29 \end{bmatrix}$$

図 4-1 4x4 DST 変換行列

変換スキップは輝度／色差信号、イントラ／インターCU のいずれの場合においても TB サイズが 4x4 である場合に選択可能なモードである。前述の通り、輝度信号の 4x4 イントラ TB では整数精度 DST と切り替えて適用され、それ以外の 4x4TB の場合においては整数精度 DCT と切り替えて適用される。変換スキップは予測差分信号に対して変換処理を行わずに、量子化、エントロピー符号化処理を行うモードである。変換処理を用いないため、変換係数の量子化によって生じるエッジの劣化が生じにくく、CG や PC 画面などの急峻なエッジを含む映像の再現性が高い。

4x4TB 以外の場合においては整数精度 DCT のみが適用される。HEVC における整数精度 DCT は Chen の高速バタフライ演算を利用できるように基底が設計されており、大サイズの変換行列は小サイズの変換行列を内包する。図 4-2、4-3 に 4x4 と 8x8 の整数精度 DCT の変換行列を示す。

$$\begin{bmatrix} 64 & 64 & 64 & 64 \\ 83 & 36 & -36 & -83 \\ 64 & -64 & -64 & 64 \\ 36 & -83 & 83 & -36 \end{bmatrix}$$

図 4-2 4x4 変換行列

$$\begin{bmatrix} 64 & 64 & 64 & 64 & 64 & 64 & 64 & 64 \\ 89 & 75 & 50 & 18 & -18 & -50 & -75 & -89 \\ 83 & 36 & -36 & -83 & -83 & -36 & 36 & 83 \\ 75 & -18 & -89 & -50 & 50 & 89 & 18 & -75 \\ 64 & -64 & -64 & 64 & 64 & -64 & -64 & 64 \\ 50 & -89 & 18 & 75 & -75 & -18 & 89 & -50 \\ 36 & -83 & 83 & -36 & -36 & 83 & -83 & 36 \\ 18 & -50 & 75 & -89 & 89 & -75 & 50 & -18 \end{bmatrix}$$

図 4-3 8x8 変換行列

8x8 変換行列は変換行列の左側 0, 2, 4, 6 行目の変換係数が 4x4 変換行列と一致しており、大サイズの変換を実現する際に小サイズの変換を再利用することができる。同様に 16x16、32x32 変換行列もそれぞれ小サイズの変換行列を内包しており、実装時において小サイズの変換を利用することによって少ないリソースで高速に大サイズの変換を実現することができる。AVC の変換処理では基底ごとに、異なる正規化処理が行われていたため量子化の際に係数ごとに異なるスケーリング項を乗算する必要があったが、HEVC においてはそれぞれの基底で同一の正規化処理が行われるため基底ごとに異なる処理の必要はない。

4.2 量子化

HEVC における量子化パラメータは、AVC 同様、量子化ステップを直接伝送するのではなく、量子化ステップの対数と比例する値を伝送する。そのため量子化パラメータが 6 増加すると量子化ステップが 2 倍となるのもまた AVC と同じである。整数精度 DCT、DST の別なく、共通の量子化処理が適用される。変換スキップモードにおいては予測差分信号を 5 bit シフトアップしたうえで量子化/逆量子化処理が行われる。

視覚特性を反映した量子化処理として従来方式と同様、量子化マトリックスによる量子化処理を行うことができる。量子化マトリックスを用いることにより、周波数成分ごとに品質制御が可能である。量子化マトリックスはイントラ/インターで個別に伝送することができるが、伝送できる量子化マトリックスは 4x4 および 8x8 までとなる。16x16 および 32x32 の量子化マトリックスは 8x8 の量子化マトリックスをスケーリングして使用する。

$$\begin{bmatrix} 16 & 16 & 16 & 16 & 17 & 18 & 21 & 24 \\ 16 & 16 & 16 & 16 & 17 & 19 & 22 & 25 \\ 16 & 16 & 17 & 18 & 20 & 22 & 25 & 29 \\ 16 & 16 & 18 & 21 & 24 & 27 & 31 & 36 \\ 17 & 17 & 20 & 24 & 30 & 35 & 41 & 47 \\ 18 & 19 & 22 & 27 & 35 & 44 & 54 & 65 \\ 21 & 22 & 25 & 31 & 41 & 54 & 70 & 88 \\ 24 & 25 & 29 & 36 & 47 & 65 & 88 & 115 \end{bmatrix}$$

a. 8x8 デフォルト量子化マトリックス

16	16	16	16	16	16	16	16	17	17	18	18	21	21	24	24
16	16	16	16	16	16	16	16	17	17	18	18	21	21	24	24
16	16	16	16	16	16	16	16	17	17	19	19	22	22	25	25
16	16	16	16	16	16	16	16	17	17	19	19	22	22	25	25
16	16	16	16	17	17	18	18	20	20	22	22	25	25	29	29
16	16	16	16	17	17	18	18	20	20	22	22	25	25	29	29
16	16	16	16	18	18	21	21	24	24	27	27	31	31	36	36
16	16	16	16	18	18	21	21	24	24	27	27	31	31	36	36
17	17	17	17	20	20	24	24	30	30	35	35	41	41	47	47
17	17	17	17	20	20	24	24	30	30	35	35	41	41	47	47
18	18	19	19	22	22	27	27	35	35	44	44	54	54	65	65
18	18	19	19	22	22	27	27	35	35	44	44	54	54	65	65
21	21	22	22	25	25	31	31	41	41	54	54	70	70	88	88
21	21	22	22	25	25	31	31	41	41	54	54	70	70	88	88
24	24	25	25	29	29	36	36	47	47	65	65	88	88	115	115
24	24	25	25	29	29	36	36	47	47	65	65	88	88	115	115

b. 16x16 デフォルト量子化マトリックス
(8x8 用の量子化マトリックスをスケーリング)

図 4-4 量子化マトリックス (イントラ用)

5. インター予測

HEVC におけるインター予測 (フレーム間予測) では、従来方式と同様にブロック単位で動き補償予測を行うが、AVC に比べてブロックサイズの大型化 (最大 64×64 画素) や小数画素精度動き補償フィルタおよび動きベクトル予測の高性能化などにより、動き補償予測の効率を向上させている。一方、デコーダにおける参照ピクチャ管理プロセスを単純化するなど、複雑さの軽減も図られている。

5.1 複数参照フレーム

AVC と同様に、HEVC では、復号ピクチャバッファ (Decoded Picture Buffer) に保存されているフレームから、過去、未来に関わらず 2 枚の参照フレームを選択し、インター予測を行うことができる。また、より最適な参照フレームの選択のため、復号ピクチャバッファに保持するフレームをエンコーダが指定することも同様に可能であるが、HEVC では、インター予測で参照するフレームを参照ピクチャ情報 (Reference Picture Set) として明示的にデコーダへシグナリングし、デコーダにおける参照ピクチャ管理プロセスが単純化された。

図 5-1 に参照ピクチャ情報の一例を示す。

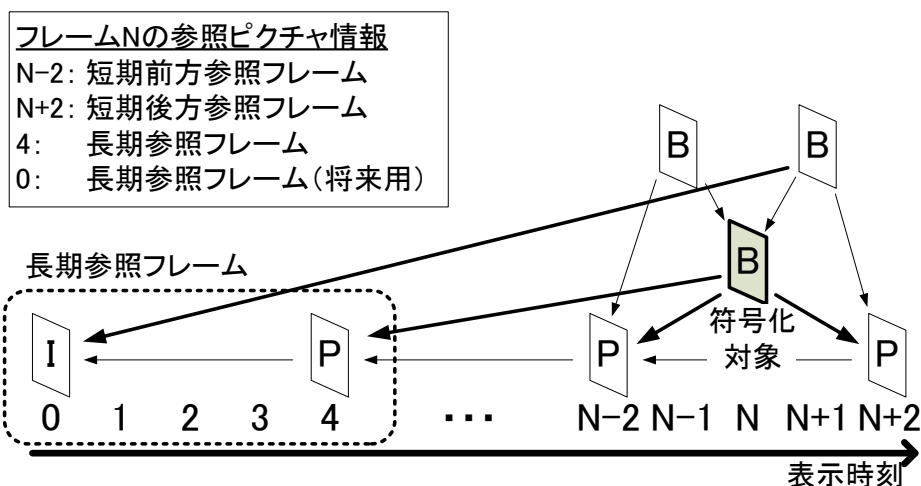


図 5-1 参照ピクチャ情報の一例

5.2 インターPB サイズ

AVC では 4×4 から 16×16 までの PB サイズを使用できたが、HEVC ではより大きな PB サイズや非対称な PU 分割が導入された。大きな PB サイズを使用することで、動き補償に関する情報をまとめて符号化することができ、その結果インター予測符号化の効率が向上している。

表 5-1 インターPB サイズの比較

AVC	HEVC	
	対称 PU 分割	非対称 PU 分割
$16 \times 16, 16 \times 8, 8 \times 16,$ $8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4$	$64 \times 64, 64 \times 32, 32 \times 64,$ $32 \times 32, 32 \times 16, 16 \times 32,$ $16 \times 16, 16 \times 8, 8 \times 16,$ $8 \times 8, 8 \times 4, 4 \times 8$	$64 \times 16 : 64 \times 48,$ $16 \times 64 : 48 \times 64,$ $32 \times 8 : 32 \times 24,$ $8 \times 32 : 24 \times 32,$ $16 \times 4 : 16 \times 12,$ $4 \times 16 : 12 \times 16$

5.3 小数画素精度動き補償

AVC では参照ピクチャから 6 タップフィルタを用いて $1/2$ 画素精度の予測画像を生成し、更に 2 タップフィルタを用いて $1/4$ 画素精度の予測画像を生成していた。HEVC では、フィルタを 8 タップ化して性能を向上すると共に、参照ピクチャから直接 $1/4$ 画素精度の予測画像を生成して演算誤差を抑制することで、動き補償予測の効率を向上させている。

5.4 動きベクトル予測

HEVC においては、エンコーダ、デコーダの双方で所定の導出方法により2つの候補ベクトルを決定し、エンコーダが動きベクトルの予測に用いた候補ベクトルの識別子をデコーダへシグナリングする。識別子追加による符号量増加に比べて動きベクトルの予測効率向上が大きく、全体として符号量が削減される。

図 5-2 は、予測動きベクトルの候補を示している。候補ベクトル mvLXA, mvLXB, mvLXCol がそれぞれの候補から決定され、可用性等に基づきエンコーダ、デコーダの双方で使用可能な2つの候補ベクトルに絞り込まれる。

なお、mvLXCol 導出用の参照フレーム(Collocated picture)は、AVC では特定の参照ピクチャインデックスが指す参照フレームに固定されていたが、HEVC ではどの参照フレームを用いるか明示的に示す方式が導入された。

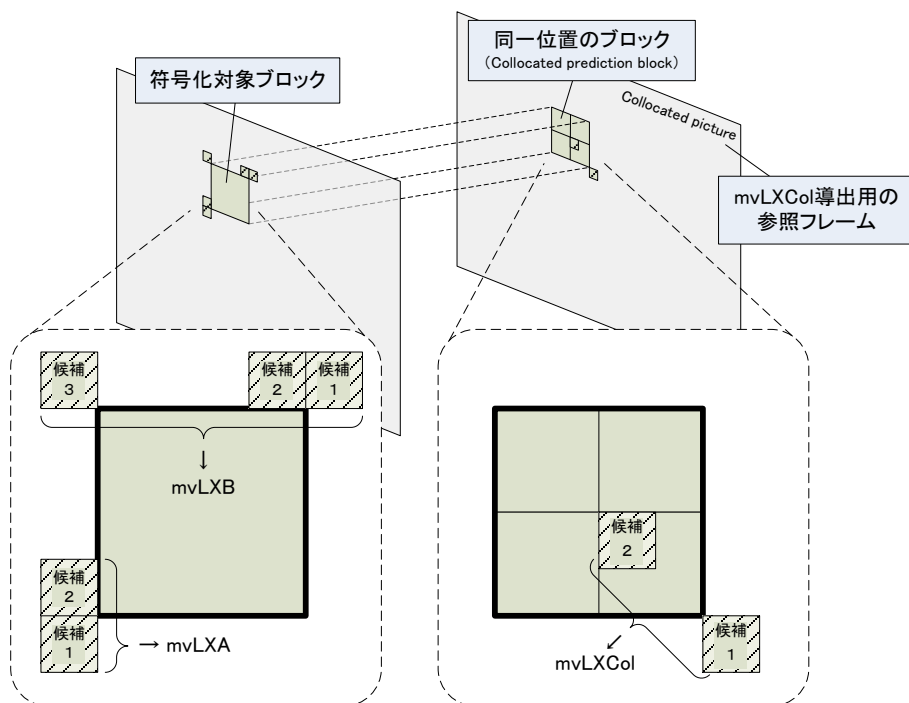


図 5-2 予測動きベクトルの候補

5.5 マージモード

AVC では、動き情報を符号化せずに符号化済み PB の動き情報から予測を行うダイレクトモードを使用することができた。HEVC では、ダイレクトモードに代えて近傍ブロックと動き情報を統合するマージモードという方式が採用された。マージモードでは、エンコーダ、デコーダの双方で所定の導出方法により 5 つの統合ブロック候補を決定し、エンコーダが選択した統合ブロック候補の識別子をデコーダへシグナリングする。

図 5-3 は、統合ブロック候補の位置を示している。図中の7つの候補ブロックの動き

情報の可用性等に基づき、エンコーダ、デコーダの双方で使用可能な 5 つの候補ブロックに絞り込まれる。絞り込みにより候補が 5 つに満たない場合には、7つの候補ブロックの動き情報を組み合わせたブロックや(0,0)ベクトルのブロックなどを仮想的に生成し、候補ブロックとして追加する。

なお、HEVC のスキップモードにおいては、このマージモードにより動き情報を決定する。

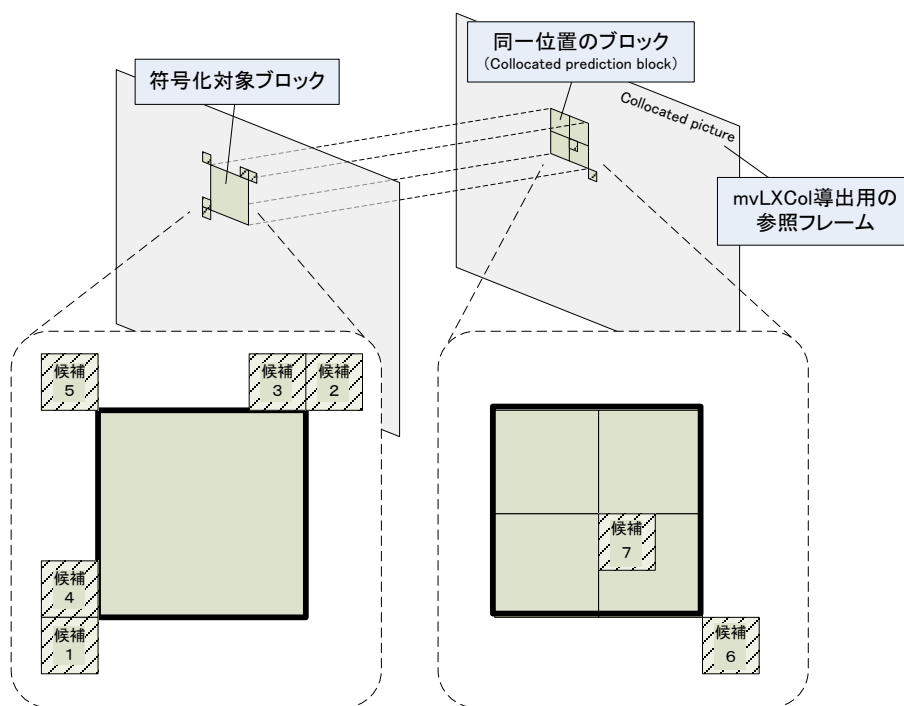


図 5-3 動き情報を統合するブロック位置の候補

6. イントラ予測

HEVCでは、フレーム内符号化されるブロックに対して、符号化済みの隣接ブロックの画素値から予測画像を生成し、原画像との差分を符号化する画面内予測(イントラ予測)が採用されている。

予測モードは前述の PU ごとにシグナリングされるが、予測画像は TU ごとに生成される。例えば、CU サイズ=PU サイズ=TU サイズであれば、予測モードと予測画像の処理サイズは一致している。ところが、PU サイズ>TU サイズの場合には、予測モードは 1 種類だが、予測画像の生成、予測差分信号の符号化・復号、復号画像の生成といったローカル復号画像の生成までを TB ごとに繰り返し行う。さらに、CU サイズが最小である場合に限り PU は 4 分割可能 (CU サイズ>PU サイズ、PART_NxN) であり、輝度と色差とで PU 構造が異なる。すなわち、輝度 CB には 4PU 含まれるが、色差 CB には 1PU しか含まれない。

輝度イントラ予測方式は、AVC と比較して、9 種類から 35 種類に増加している。ま

た、色差イントラ予測方式も同様に、4 種類から輝度と同じ程度まで増加しているが、詳細は後述する。予測方向を図 6-1 に示す。

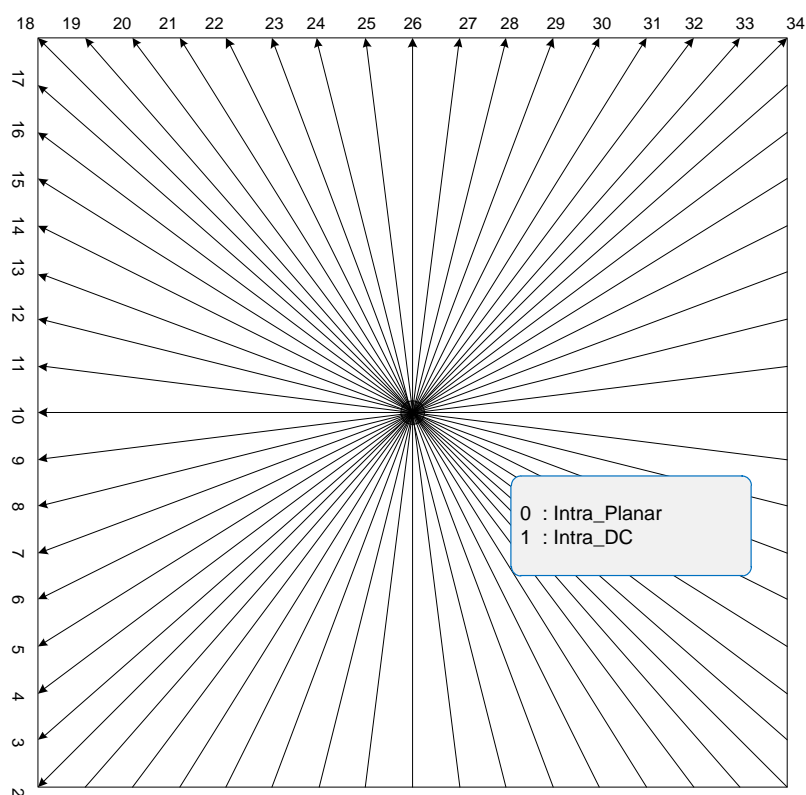


図 6-1 イントラ予測モードの予測方向

6.1 輝度信号のイントラ予測符号化

AVC と比較して、シグナリングは大きく変更されている。35 種類の予測モードを効率的に符号化するため、最確モード (most probable mode) 方式が導入された。

輝度信号のイントラ予測は、上述したように TB ごとに予測画像を生成する。ある TB の予測画像を生成する場合、符号化済みの CB から参照画素を特定する。次に、予測モードに応じた平滑化フィルタを参照画素へ適用する。最後に、予測モード (Planar、DC、方向性) に応じて予測画像を生成する。なお、平滑化フィルタを除いて、この処理手順は色差イントラ予測でも共通である。

まず、参照画素は図 6-2 に示すような左側と上側に TB サイズの 2 倍の個数と、左上に一つあると見なす。実際には、符号化済みで参照可能であれば該当する画素値を用い、符号化前などで参照できなければ近傍の画素値で代用する。

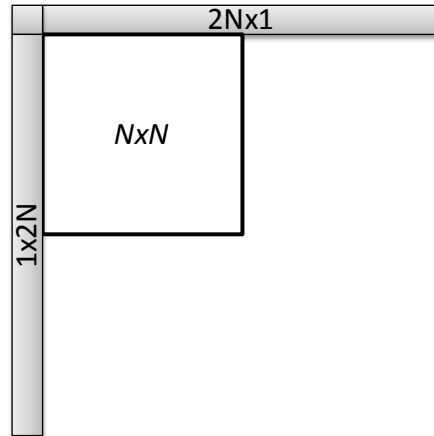


図 6-2 イントラ予測対象ブロック $N \times N$ と参照画素の関係

次に、平滑化フィルタは **TB** サイズが大きく水平方向や垂直方向から遠い場合ほど適用する。逆に、予測モードが **DC**、水平、垂直予測モード (**TB** サイズに依存しない) と、**TB** サイズが小さく水平方向や垂直方向に近い場合ほど平滑化フィルタを適用しない。イントラ予測により生じる擬似輪郭を抑制するための特別な条件を除いて、平滑化は **1:2:1** で重み付けした **3** タップフィルタで実現する。

最後に、具体的な予測画像の生成手順を述べる。

Planar 予測は、図 6-3 に示す 4 点を用いて、内挿予測により予測画素値を生成する。ここで、黒画素が参照画素、グレー画素が予測対象画素である。

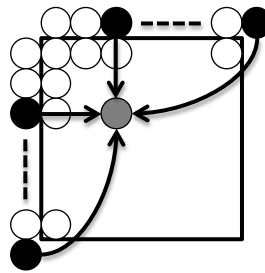


図 6-3 **Planar** 予測方式

DC 予測は、**TB** サイズが $N \times N$ の場合、図 6-2 に示すような周囲にある参照画素のうち左上を除いて隣接する $2N$ 個の平均値を用いて、**TB** 内の予測画素値を生成する。**TB** サイズが **32** 未満の時には、参照画素に隣接している予測対象画素に限定して、参照画素値と予測画素値との重み付け和を新しい予測画素値とする。

方向性予測は、参照画素をある方向に沿って外挿して予測画素値を生成する。予測モードが **18** 以上、すなわち予測方向が垂直方向の場合と、予測モードが **18** 未満、すなわち予測方向が水平方向の場合とに分けられる。ただし、水平方向と垂直方向は転置の関係にあるため、ここでは垂直方向のみを説明する(図 6-4)。

予測モード 26（垂直方向）を基準として、予測モード 25 以下の場合には左図のような参照画素を準備し、予測モード 27 以上の場合には、右図のような参照画素を準備する。なお、左図の矢印は、黒画素の画素値がグレー画素の位置にあると仮定して、グレー画素を参照することを表している。

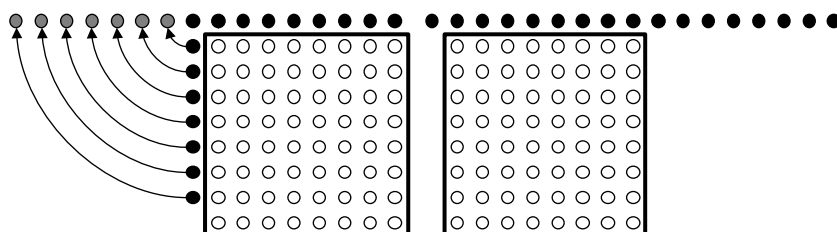


図 6-4 方向性予測（垂直方向）の参照画素

あらかじめ定義された角度に基づき、1/32 画素精度で参照画素位置を算出する。小数画素位置を指し示す場合には、左右の画素値を線形補間した値を予測画素値とする。整数画素位置を指し示す場合には、参照画素値を予測画素値とする。

予測モード 26（垂直方向）で TB サイズが 32 未満の場合には、左端 1 列の画素値は、隣接する左側の参照画素と最も左上の参照画素とを用いて補正する。

6.2 色差信号のイントラ予測符号化

色差信号のイントラ予測は、AVC と比較してシグナリングが変更されている。具体的には、色差 PU の予測モードは CU 内の左上輝度 PB の予測モードと同一であるというシグナリングが追加された。この結果、色差信号のイントラ予測の予測モード数は輝度信号のそれと同レベルにまで増加している。また、輝度と色差とで予測モードが同一でない場合には、Planar、垂直、水平、DC（、斜め）からが定義されている。上述したように、予測画像の生成手順は全ての平滑化処理を除いて同一である。

7. エントロピー符号化

HEVC では、ブロック層データの符号化に CABAC（Context-based Adaptive Binary Arithmetic Coding：コンテキスト適応バイナリ算術符号化）を使用する。AVC では、アプリケーションの特性に応じて、CAVLC（Context-based Adaptive Variable Length Coding：コンテキスト適応可変長符号化）を選択できたが、プロセッサの処理能力向上などを受けて HEVC では CABAC に一本化された。

CABAC は、シンタックス要素を所定の方法により二値化（Binarize）し、二値化ビット列のビット毎に出現頻度モデル（コンテキスト）を切り替えて算術符号化する。コンテキストは、符号化対象および周囲の状況に応じて、適用するモデルを動的に切り替えながら二値化ビット列に符号を割り当てる。

HEVC の CABAC 処理の基本構成は AVC の CABAC と同様であるが、コンテキストの共用化や上側参照削減によるメモリ削減に加え、CABAC 同期処理やトランケーテッド・ライス(Truncated Rice) 二値化処理の導入による並列実装の容易化など、符号化効率を維持しつつ CABAC の実装を容易にする改良が施された。

7.1 CABAC 同期処理

AVC では、コンテキストおよび算術符号は両者ともスライス単位で初期化／終端されていたが、HEVC では、表 7-1 に示すように、算術符号が初期化される場合にもコンテキストは必ずしも初期化されず、WPP 有効時の CTU 列先頭や従属スライスセグメントの先頭においては、コンテキストが所定ブロックの処理終了時の状態に同期される。この同期処理により、算術符号を終端／初期化した後に学習済みコンテキストを用いて算術符号を開始できるため、符号化効率を低下させることなく並列処理が実現可能となる。

表 7-1 CABAC の初期化

位置	コンテキスト	算術符号
独立スライスセグメント またはタイルの先頭	初期化	初期化
WPP 有効時の CTU 列 の先頭	上側 CTU 列の 2 番目の CTU 処理終了後の コンテキストと同期 (上側 CTU 列を参照不能の場合は初期化)	初期化
従属スライスセグメント の先頭	直前のスライスセグメントの最終 CTU 処理 終了後のコンテキストと同期	初期化

7.2 トランケーテッド・ライス二値化処理

AVC では、トランケーテッド・ユーナリと指数ゴロムを用いて、全ての変換係数のレベルが同じ方法で符号化されていた。HEVC では、トランケーテッド・ユーナリの代わりにトランケーテッド・ライス(TR)を用いて、符号化済み変換係数レベルの大きさに応じて二値化処理を切り替える方法が採用された。これにより、変換係数レベルの二値化列ワースト長がトランケーテッド・ユーナリより短縮され、特に高いビットレートにおいて 1 画素あたりの二値化列長が抑制される。二値化列長が短縮されたことにより、CABAC の処理量が軽減され、更には、TR による二値化列をバイパスモード(固定確率モード)により算術符号化することで符号列の一括処理実装が容易となり、CABAC 実装におけるスループットの向上が期待されている。

表 7-2 は、TR による二値化表現の一覧である。変換係数レベルは、レベルが小さい高周波成分から順に二値化され、小さなレベルを表現する際の二値化列長が短い、ライスパラメータ 0 が最初に用いられる。所定の閾値を超える変換係数レベルを二値化する度に、より大きなレベルを表現するのに適した一つ大きなライスパラメータの二

値化表現へ移行する。TRの最大値を超える大きな変換係数レベル値については、超える部分を指数ゴロムを用いて二値化する。

表 7-2 トランケーテッド・ライスの二値化表現

レベル	ライスパラメータ 0		ライスパラメータ 1		ライスパラメータ 2		ライスパラメータ 3		ライスパラメータ 4	
	Prefix	Suffix	Prefix	Suffix	Prefix	Suffix	Prefix	Suffix	Prefix	Suffix
0	0		0	0	0	00	0	000	0	0000
1	10		0	1	0	01	0	001	0	0001
2	110		10	0	0	10	0	010	0	0010
3	1110		10	1	0	11	0	011	0	0011
4			110	0	10	00	0	100	0	0100
5			110	1	10	01	0	101	0	0101
6			1110	0	10	10	0	110	0	0110
7			1110	1	10	11	0	111	0	0111
8					110	00	10	000	0	1000
9					110	01	10	001	0	1001
10					110	10	10	010	0	1010
11					110	11	10	011	0	1011
12					1110	00	10	100	0	1100
13					1110	01	10	101	0	1101
14					1110	10	10	110	0	1110
15					1110	11	10	111	0	1111
16							110	000	10	0000
17							110	001	10	0001
:							:	:	:	:
:							:	:	:	:
30							1110	110	10	1110
31							1110	111	10	1111
32									110	0000
33									110	0001
:									:	:
:									:	:
46									110	1110
47									110	1111
48									1110	0000
49									1110	0001
:									:	:
:									:	:
62									1110	1110
63									1110	1111

8. ループ内フィルタ

ループ内フィルタは、復号画像の符号化ノイズを低減するフィルタであり、画質劣化がフレーム間で伝播することを防ぐ。視覚的に目立つブロック歪みやリングング歪みを直接低減することが可能であり主観性能を大きく向上させる。また、客観性能への寄与も大きい。

HEVC のループ内フィルタは、デブロッキング・フィルタ (DF: Deblocking Filter) と画素適応オフセット (SAO: Sample Adaptive Offset) から構成される。ループ内フィルタは比較的複雑な処理であるため DF、SAO とともに独立にオン・オフを制御できる。また、スライスやタイル境界へのループフィルタ適用もオン・オフ制御可能である。これにより、必要に応じて並列処理後のループ内フィルタ適用を省略できる。なお、画質劣化のない PCM 符号化やロスレス符号化された画素のフィルタは省略される。

8.1 デブロッキング・フィルタ

予測、変換、量子化はブロック単位で行われるため、ブロック境界に不連続な歪み (ブロック歪み) が生じる。デブロッキング・フィルタ (Deblocking Filter; DF) は、ブロック境界線 (垂直境界と水平境界) に沿って、境界線の近傍画素を用いて平滑化処理を行う。

HEVC では、画面内の全ての垂直境界 (もしくは水平境界) の並列処理が可能のようにフィルタ順が規定されている。したがって、画面単位で垂直境界のフィルタを施した後、そのフィルタ後画像を用いて水平境界のフィルタ処理を施す (図 8-1)。AVC 同様、ブロック単位処理も可能である。

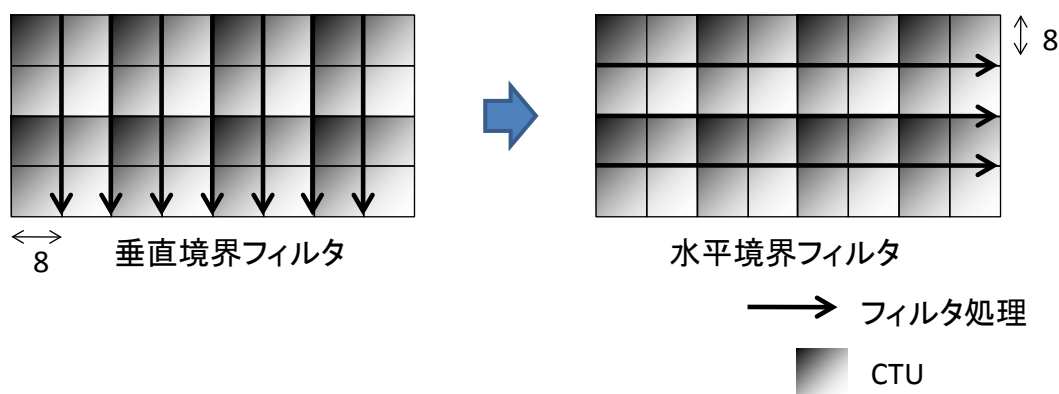


図 8-1 ピクチャ単位フィルタ処理

フィルタ処理は、TB 境界と PB 境界に対して適用される。両境界が重なる場合も 1 度だけフィルタ処理が適用される。TB および PB の最小サイズは 4 画素であるが、小ブロックではフィルタ効果が比較的小さいことから、複雑度削減を目的として、8 画素単位の境界にのみフィルタ処理が適用される。

ブロック歪みの大きさや範囲は境界に応じて異なること、平滑化が強すぎるとボケ等の弊害があることから、境界付近の符号化パラメータで定まるブロック境界強度 bS (0~2) と、画素値で定まるエッジ分類 dE (0~2)、量子化ステップと bS で定まるクリップ制御値 tc 、平坦度判定の閾値 $beta$ を用いて、フィルタを制御する。 bS は、AVC 同様、イントラ予測や残差の有無、参照ピクチャや動きベクトルの違いから導出される。 bS による分類に加えて、 dE による強フィルタ、弱フィルタの分類がある。なお、閾値 tc 、 $beta$ は、AVC 同様、スライス単位で変更可能であり、フィルタ強度の調整に利用できる。

図 8-2 は、垂直境界の場合の画素値を示すものである。境界付近の前後の画素値 4 画素 $p_0 \sim p_3$ 、 $q_0 \sim q_3$ を参照し、以下のフィルタ処理が適用される。

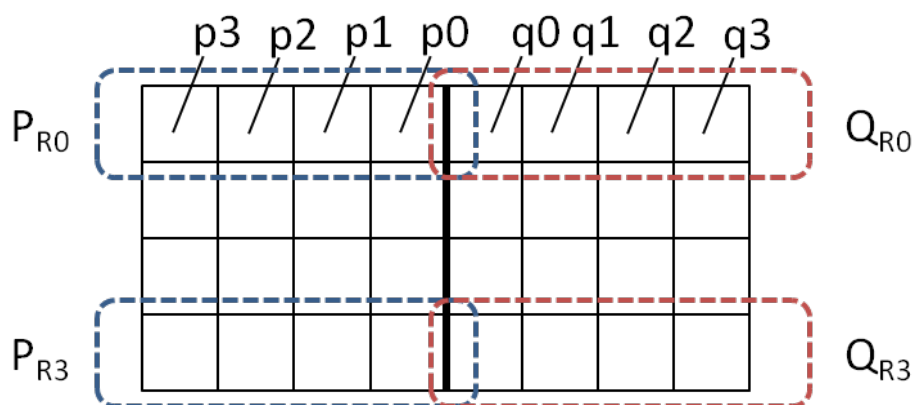


図 8-2 境界画素

(1) bS が 0 の場合、フィルタは施されない (フィルタオフ)。

(2) dE が 0 の場合も、フィルタオフ。 dE が 1 (弱フィルタ)、 dE が 2 (強フィルタ) の場合は、(3) 以降に示す手順でフィルタが適用される。図 8-2 に図示した領域 P_{R0} 、 P_{R3} 、 Q_{R0} 、 Q_{R3} のアクティビティの和が閾値 $beta$ 未満であれば dE は 1 以上の値に設定される。さらに、領域 P_{R0} と Q_{R0} および領域 P_{R3} と Q_{R3} それぞれの組み合わせ (P_{Ri} 、 Q_{Ri} と表記) について次の条件全てを満たす場合は dE に 2 が設定される: (a) P_{Ri} と Q_{Ri} の領域のアクティビティの和が閾値 $beta$ の 1/8 より小さい、(b) P_{Ri} と Q_{Ri} の領域の平坦度の和が閾値 $beta$ の 1/8 より小さい、(c) 境界をまたぐ 2 画素の差の絶対値が $2.5tc$ より小さい。

(3) dE が 1 の場合、弱フィルタが適用される。弱フィルタは、境界隣接画素 (p_0 や q_0) を変更する。

$$\Delta = \text{Clip3}(-tc, tc, (9 * (q_0 - p_0) - 3 * (q_1 - p_1) + 8) \gg 4)$$

$$p_0' = \text{Clip1}_Y(p_0 + \Delta)$$

$$q_0' = \text{Clip1}_Y(q_0 - \Delta)$$

なお、 $\text{Clip3}(-t_c, t_c, A)$ は、 A の値の $[-t_c, t_c]$ の範囲へのクリップ、 $\text{Clip1}_Y(A)$ は、 A の値の画素値の範囲へのクリップを意味する。

さらに、弱フィルタにおいて、 P_{R0} 、 P_{R3} のアクティビティの和が閾値 (beta の $3/16$) より小さい場合、ブロック境界から 1 画素目の画素 (p_1 の列の画素) も変更する。(q1 の列の画素も同様。)

$$\Delta p = \text{Clip3}(- (t_c \gg 1), t_c \gg 1, (((p_2 + p_0 + 1) \gg 1) - p_1 + \Delta) \gg 1)$$

$$p_1' = \text{Clip1}_Y(p_1 + \Delta p)$$

(4) dE が 2 の場合には、強フィルタが施され、境界前後 3 画素 ($p_0 \sim p_2$ 、 $q_0 \sim q_2$ の列の画素) を変更する。 $p_0 \sim p_2$ の画素の式を示す。(q0~q2 も同様。)

$$p_0' = \text{Clip3}(p_0 - 2 * t_c, p_0 + 2 * t_c, (p_2 + 2 * p_1 + 2 * p_0 + 2 * q_0 + q_1 + 4) \gg 3)$$

$$p_1' = \text{Clip3}(p_1 - 2 * t_c, p_1 + 2 * t_c, (p_2 + p_1 + p_0 + q_0 + 2) \gg 2)$$

$$p_2' = \text{Clip3}(p_2 - 2 * t_c, p_2 + 2 * t_c, (2 * p_3 + 3 * p_2 + p_1 + p_0 + q_0 + 4) \gg 3)$$

8.2 画素適応オフセット

画素適応オフセット (Sample Adaptive Offset; SAO) は HEVC で新たに導入されたループ内フィルタであり、DF 適用後の画像を画素単位で分類し、各画素値に分類に応じたオフセットを加算することで、主観性能および客観性能を向上させる。特に、リングング歪み除去に効果がある。画素単位の分類が必要な非線形フィルタであるが、CTU 共通パラメータ、簡易な分類方法、オフセットフィルタ等の限定により、比較的演算量が小さい。

分類の概略と伝送

画素の分類方法には、バンドオフセット (BO: Band Offset)、エッジオフセット (EO: Edge Offset) の 2 種類がある。CTU 単位で BO、EO、オフの何れかが輝度と色差で独立に選択できる。

BO では、画素値の属する値域 (バンド) により画素を分類する。分類の補助情報として、オフセットを適用するバンドを指定する「バンド位置」が伝送される。加えて、バンド位置の示す 4 つのバンドに属する画素に適用するオフセット値 (整数) が伝送される。

EO では、隣接 3 画素の大小関係により判定されるエッジの状態画素を分類する。分類の補助情報として、判定に用いる隣接 3 画素の方向を示す「エッジオフセットクラス」が伝送される。加えて、エッジの状態に応じて適用するオフセット値 (非負整数) が伝送される。

BO による分類とオフセット適用

画素値の値域を等分して 32 個のバンドを設定し、画素値が属するバンドにより分類する (図 8-3)。画素毎に、属するバンドに対応するオフセットが加算される。なお、一つの CTU 内では、32 個のバンドのうち連続する 4 個のバンドにのみオフセットを適用できる。4 個のバンドを特定するために、4 個のバンドの先頭位置を示す情報がバンド位置 (Band Position) として用いられる。

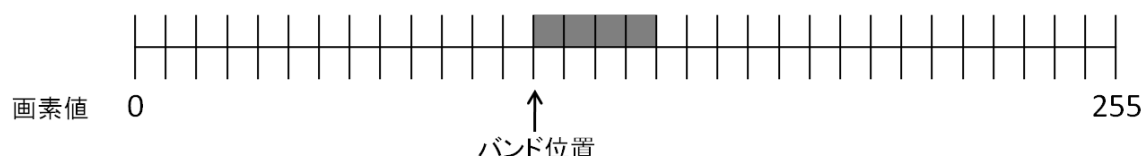


図 8-3 画素値の値域による分類とバンド位置

EO による分類とオフセット適用

エッジオフセットクラス (EO Class) として指定された方向 (水平、垂直、45 度、135 度のいずれか) に連続する 3 画素 (図 8-4) の大小関係によりエッジ状態进行分类する。エッジ状態の分類 (分類番号) に対応する 3 画素の大小関係を図 8-5 に示す。画素毎に判定されたエッジ状態に対応するオフセット値が画素値に加算される。なお、EO では、エッジが平坦になる方向にオフセットの符号が制限されている。それにより、意図しないエッジ強調による粒状雑音 (salt-and-pepper noise) の発生を防いでいる。

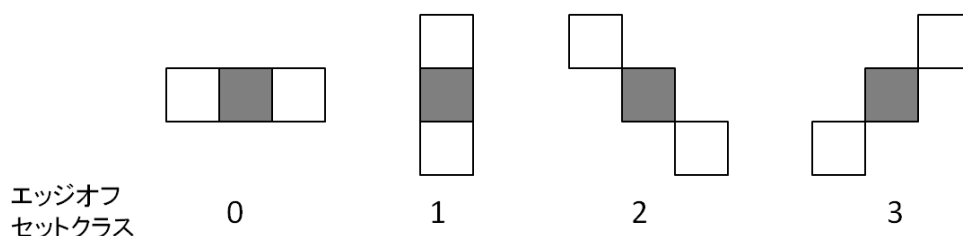


図 8-4 エッジオフセットクラスと隣接画素の関係

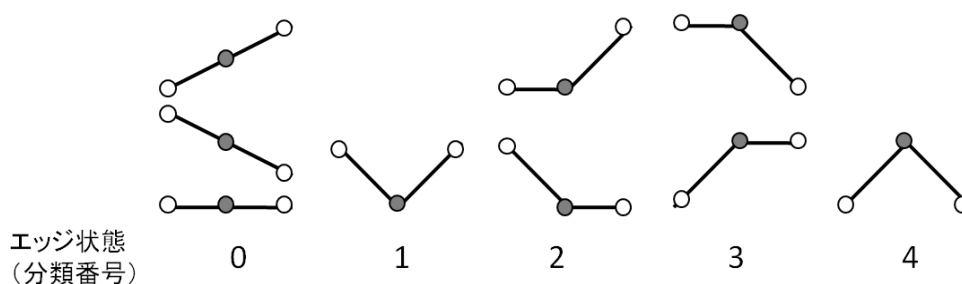


図 8-5 隣接 3 画素の大小関係によるエッジ状態の分類

9. プロファイルとレベル

HEVC は、AVC と同様に、モバイルからプロ・スタジオなどの広範な用途に適用できる高品質の汎用映像符号化方式である。デコーダに実装する機能や処理対象の映像等に制約を設けることにより、各用途において経済的、効率的な実装を可能とすると共に相互接続性を確保しており、この目的のために、使用するツールと処理範囲等を定めるプロファイルと、映像解像度やビットレートといったパラメータの上限を定めるレベルが規定されている。

HEVC 規格の第1版は、コンシューマ向けの映像フォーマット(4:2:0 色差フォーマット、8-bit/10-bit 画素値精度)のみに対応しており、対応する3種類のプロファイルと、13種類のレベルを定義している。

9.1. プロファイル

表 9-1 HEVC 規格第1版で定義されるプロファイル

プロファイル名	概要・想定用途	主な仕様
Main	標準的なツール群で構成される代表的なプロファイル 想定用途:放送、蓄積メディア、配信など	<ul style="list-style-type: none"> ● 4:2:0 クロマフォーマット ● 8-bit 画素サポート ● I/P/B スライス ● 8x8~64x64 CU ● 4x4~32x32 DCT, 4x4 DST ● 35 モードのイントラ予測 ● 適応イントラ予測フィルタ ● CABAC ● デブロッキング・フィルタ ● 画素適応オフセット ● 量子化マトリクス ● 変換スキップ ● スライス/タイル構造 ● WPP ● 4x8/8x4~64x64 PU ● 1/4 画素精度動き補償 ● 重み付き予測 ● 適応動きベクトル予測 ● 時間方向階層符号化
Main 10	Main プロファイルに対し、画素値精度を10-bit に引き上げたプロファイル 想定用途:放送、蓄積	<ul style="list-style-type: none"> ● 10-bit 画素サポート

プロファイル名	概要・想定用途	主な仕様
	メディア、配信など	
Main Still Picture	静止画像に特化したプロファイル 想定用途: デジタルカメラ、静止画像配信など	Main プロファイルから、動き補償予測関連のツールを除いたもの

9.2. レベル

HEVC では、最大ビットレート規定において、層(tier)という概念を導入している。主にコンシューマ用途向けの Main tier と、プロフェッショナル用途向けの High tier が存在し、それぞれの中で、レベルと最大ビットレートとが比例関係になるように設定されている。

表 9-2 HEVC 規格第 1 版で定義されるレベル

レベル	輝度画素数最大値		最大ビットレート [kbps]		最大 CBP サイズ [kbits]		ピクチャ内最大スライス数	最大タイル行数	最大タイルカラム数	最大圧縮率
	ピクチャ内	秒当たり	Main tier	High tier	Main tier	High tier				
1	36 864	552 960	128	-	350	-	16	1	1	2
2	122 880	3 686 400	1 500	-	1 500	-	16	1	1	2
2.1	245 760	7 372 800	3 000	-	3 000	-	20	1	1	2
3	552 960	16 588 800	6 000	-	6 000	-	30	2	2	2
3.1	983 040	33 177 600	10 000	-	10 000	-	40	3	3	2
4	2 228 224	66 846 720	12 000	30 000	12 000	30 000	75	5	5	4
4.1	2 228 224	133 693 440	20 000	50 000	20 000	50 000	75	5	5	4
5	8 912 896	267 386 880	25 000	100 000	25 000	100 000	200	11	10	6
5.1	8 912 896	534 773 760	40 000	160 000	40 000	160 000	200	11	10	8
5.2	8 912 896	1 069 547 520	60 000	240 000	60 000	240 000	200	11	10	8
6	35 651 584	1 069 547 520	60 000	240 000	60 000	240 000	600	22	20	8
6.1	35 651 584	2 139 095 040	120 000	480 000	120 000	480 000	600	22	20	8
6.2	35 651 584	4 278 190 080	240 000	800 000	240 000	800 000	600	22	20	6

表 9-3 最大ピクチャレートと代表的な画像フォーマット

Level:				1	2	2.1	3	3.1	4	4.1	5	5.1	5.2	6	6.1	6.2
フォーマット名	輝度幅	輝度高	輝度 サイズ													
SQCIF	128	96	16 384	33.7	225.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
QCIF	176	144	36 864	15.0	100.0	200.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
QVGA	320	240	81 920	-	45.0	90.0	202.5	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
525 SIF	352	240	98 304	-	37.5	75.0	168.7	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
CIF	352	288	122 880	-	30.0	60.0	135.0	270.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
525 HHR	352	480	196 608	-	-	37.5	84.3	168.7	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
625 HHR	352	576	221 184	-	-	33.3	75.0	150.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
Q720p	640	360	245 760	-	-	30.0	67.5	135.0	272.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
VGA	640	480	327 680	-	-	-	50.6	101.2	204.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
525 4SIF	704	480	360 448	-	-	-	46.0	92.0	185.4	300.0	300.0	300.0	300.0	300.0	300.0	300.0
525 SD	720	480	393 216	-	-	-	42.1	84.3	170.0	300.0	300.0	300.0	300.0	300.0	300.0	300.0
4CIF	704	576	405 504	-	-	-	40.9	81.8	164.8	300.0	300.0	300.0	300.0	300.0	300.0	300.0
625 SD	720	576	442 368	-	-	-	37.5	75.0	151.1	300.0	300.0	300.0	300.0	300.0	300.0	300.0
480p (16:9)	864	480	458 752	-	-	-	36.1	72.3	145.7	291.4	300.0	300.0	300.0	300.0	300.0	300.0
SVGA	800	600	532 480	-	-	-	31.1	62.3	125.5	251.0	300.0	300.0	300.0	300.0	300.0	300.0
QHD	960	540	552 960	-	-	-	30.0	60.0	120.8	241.7	300.0	300.0	300.0	300.0	300.0	300.0
XGA	1024	768	786 432	-	-	-	-	42.1	85.0	170.0	300.0	300.0	300.0	300.0	300.0	300.0
720p HD	1280	720	983 040	-	-	-	-	33.7	68.0	136.0	272.0	300.0	300.0	300.0	300.0	300.0
4VGA	1280	960	1 228 800	-	-	-	-	-	54.4	108.8	217.6	300.0	300.0	300.0	300.0	300.0
SXGA	1280	1024	1 310 720	-	-	-	-	-	51.0	102.0	204.0	300.0	300.0	300.0	300.0	300.0
525 16SIF	1408	960	1 351 680	-	-	-	-	-	49.4	98.9	197.8	300.0	300.0	300.0	300.0	300.0
16CIF	1408	1152	1 622 016	-	-	-	-	-	41.2	82.4	164.8	300.0	300.0	300.0	300.0	300.0
4SVGA	1600	1200	1 945 600	-	-	-	-	-	34.3	68.7	137.4	274.8	300.0	300.0	300.0	300.0
1080 HD	1920	1080	2 088 960	-	-	-	-	-	32.0	64.0	128.0	256.0	300.0	300.0	300.0	300.0
2Kx1K	2048	1024	2 097 152	-	-	-	-	-	31.8	63.7	127.5	255.0	300.0	300.0	300.0	300.0
2Kx1080	2048	1080	2 228 224	-	-	-	-	-	30.0	60.0	120.0	240.0	300.0	300.0	300.0	300.0
4XGA	2048	1536	3 145 728	-	-	-	-	-	-	-	85.0	170.0	300.0	300.0	300.0	300.0

16VGA	2560	1920	4 915 200	-	-	-	-	-	-	-	54.4	108.8	217.6	217.6	300.0	300.0
3616x1536 (2.35:1)	3616	1536	5 603 328	-	-	-	-	-	-	-	47.7	95.4	190.8	190.8	300.0	300.0
3672x1536 (2.39:1)	3680	1536	5 701 632	-	-	-	-	-	-	-	46.8	93.7	187.5	187.5	300.0	300.0
3840x2160 (4*HD)	3840	2160	8 355 840	-	-	-	-	-	-	-	32.0	64.0	128.0	256.0	300.0	300.0
4Kx2K	4096	2048	8 388 608	-	-	-	-	-	-	-	31.8	63.7	127.5	127.5	255.0	300.0
4096x2160	4096	2160	8 912 896	-	-	-	-	-	-	-	30.0	60.0	120.0	120.0	240.0	300.0
4096x2304 (16:9)	4096	2304	9 437 184	-	-	-	-	-	-	-	-	-	-	113.3	226.6	300.0
4096x3072	4096	3072	12 582 912	-	-	-	-	-	-	-	-	-	-	85.0	170.0	300.0
7680x4320	7680	4320	33 423 360	-	-	-	-	-	-	-	-	-	-	32.0	64.0	128.0
8192x4096	8192	4096	33 554 432	-	-	-	-	-	-	-	-	-	-	31.8	63.7	127.5
8192x4320	8192	4320	35 651 584	-	-	-	-	-	-	-	-	-	-	30.0	60.0	120.0

参考文献

- [1] ITU-T Recommendation H.265: High Efficiency Video Coding (04/2013)
- [2] 村上篤道、浅井光太郎、関口俊一編『高効率映像符号化技術 HEVC/H.265 とその応用』、オーム社、2013.
- [3] G.J.Sullivan, et.al., “ Overview of the High Efficiency Video Coding (HEVC) Standard”, IEEE Transaction on Circuits and Systems for Video Technology, vol. 22, no. 12, Dec. 2012.
- [4] ITU-T Recommendation H.264: Advanced video coding for generic audiovisual services (03/2005). (H.264 (2005) Cor.1(09/2005)を含む)
- [5] 角野眞也ほか『H.264/AVC 教科書』、インプレス、2004.
- [6] 亀山渉、花村剛監修、『デジタル放送教科書 (上)』、IDG ジャパン、2003.
- [7] 放送技術開発協議会 TV 情報符号化委員会 高能率符号化作業班、『MPEG-2 規格の概要 (Video) : MPEG 符号化要素技術』 1994 年 9 月 26 日.