

地上デジタル放送方式高度化に関わる
適用技術検討作業 中間報告

第一部:VVC 規格の主観評価実験(計画案)

2021年2月9日

デジタル放送システム開発部会
一般社団法人 電波産業会

まえがき

総務省からの諮問第 2044 号「放送システムに関する技術的条件」(2019 年 6 月 18 日)を受け、情報通信審議会放送システム委員会に地上デジタル放送方式高度化作業班が設置され、技術的条件の検討が始まった。本活動の一環として、2020 年 6 月 22 日に、ARIB に対して映像符号化及び音声符号化方式の高度化に必要な技術的検討の依頼があった。

本依頼を受け、映像符号化方式作業班は、映像符号化方式の検討を開始している。2020 年には、最新の映像符号化規格 VVC を有力候補とし、規格調査を進めると共に、VVC 規格を用いた際の所要ビットレートを求めるための主観評価実験の計画案を作成した。

本報告書は、VVC 規格の主観評価実験案、及び規格調査結果を中間報告としてまとめたものである。前者(本資料)は第一部に、後者は第二部に分けている。

VVC 規格の主観評価実験(計画案)

内容

| | |
|--|----|
| 1. 目的 | 5 |
| 2. 参考文献 | 5 |
| 3. 実験内容 | 6 |
| 3.1. 実施時期、場所 | 6 |
| 3.2. 評価映像 | 6 |
| 3.3. VVC エンコーダ | 6 |
| 3.3.1. エンコードするカテゴリと評価方法 | 7 |
| 3.4. VVC 符号化条件 | 7 |
| 3.5. 主観評価実験 | 8 |
| 3.5.1. 所要ビットレート推定の基準 | 8 |
| 3.6. 準備スケジュール | 8 |
| 付録 1 主観評価実験に用いるシーン | 9 |
| 1. カテゴリ 1 (1080/60/P, SDR) | 9 |
| 2. カテゴリ 2 (2160/60/P, SDR) | 9 |
| 3. カテゴリ 3 (2160/60/P, HDR) | 9 |
| 4. (補足)シーン選定基準 | 10 |
| 4.1. 概要 | 10 |
| 4.2. 映像情報メディア学会標準動画の VVC 符号化難易度 | 10 |
| 4.2.1. カテゴリ 1 (1080/60/P, SDR)の符号化難易度 | 11 |
| 4.2.2. カテゴリ 2 (2160/60/P, SDR) の符号化難易度 | 11 |
| 4.2.3. カテゴリ 3 (2160/60/P, HDR)の符号化難易度 | 12 |
| 付録 2 VVenC の性能とサポートツール | 13 |
| 付録 3 VVC 実用化動向・予測 | 15 |
| 1. 半導体の技術進化予測 | 15 |
| 2. VVC 実用化予測 | 15 |
| 2.1. HEVC 実用化の実績 | 15 |
| 2.2. VVC 実用化予想時期 | 16 |
| 2.3. 2024 – 2025 年の VVC 実用化予想 | 16 |

| | |
|---------------------------|----|
| 2.4. 技術課題..... | 17 |
| 3. 放送以外の VVC サービス予測 | 17 |
| 付録 4 新型コロナウイルス感染対策 | 18 |

1. 目的

VVC 規格[1]を適用した地上波デジタル放送の運用ガイドラインや技術基準の策定に寄与することを目的とし、UHDTV 及び HDTV の VVC 符号化映像の主観評価実験を実施して、所要ビットレートを明らかにする。

2. 参考文献

- [1] Recommendation ITU-T H.266 (08/2020)
“Versatile video coding”
- [2] Recommendation ITU-R BT.2245-8 (10/2020)
“HDTV and UHDTV including HDR-TV test materials for assessment of picture quality”
- [3] Recommendation ITU-R BT.500-14 (10/2019)
“Methodologies for the subjective assessment of the quality of television images“
- [4] ARIB TR-B43 1.2 版(2020)
「高ダイナミックレンジ映像を用いた番組制作の運用ガイドライン」
- [5] 映像情報メディア学会/ARIB
「ハイビジョン・システム評価用標準動画像第 2 版」(2010)
- [6] 映像情報メディア学会/ARIB
「超高精細・広色域標準動画像 － A シリーズ解説書」(2015)
- [7] 映像情報メディア学会/ARIB
「超高精細・広色域標準動画像 － B シリーズ解説書」(2017)
- [8] 映像情報メディア学会/ARIB
「超高精細・広色域 HDR 版標準動画像 － C シリーズ解説書」(2019)
- [9] https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/tags/VTM-10.0
- [10] <https://github.com/fraunhoferhhi/vvenc>
- [11] ARIB STD-B32 3.11 版(2018)
「デジタル放送における映像符号化、音声符号化及び多重化方式」

3. 実験内容

3.1. 実施時期、場所

2021年6月上旬(詳細日程は調整中、5月中の可能性も有り)に、NHK放送技術研究所にて実施する。

3.2. 評価映像

UHDTV(4K, SDR/HDR)及びHDTV(2K, SDR)形式の、映像情報メディア学会の標準動画像とする。使用するシーンは付録を参照のこと。

| カテゴリ | 映像形式 | 準拠規格 | シーン |
|------|------------------|--------------------|--|
| 1 | 1080/60/P SDR | Rec. ITU-R BT.709 | ハイビジョン・システム評価用標準動画像第二版 B シリーズ[5]の 8 シーン |
| 2 | 2160/60/P SDR | Rec. ITU-R BT.2020 | 超高精細・広色域標準動画像 A シリーズ[6]、及び B シリーズ[7]のそれぞれ 8 シーン |
| 3 | 2160/60/P HDR | Rec. ITU-R BT.2100 | 超高精細・広色域 HDR 版標準動画像 C シリーズ[8]の 8 シーン |
| 4 | 2160/60/P SDR | Rec. ITU-R BT.2020 | カテゴリ 3 の 8 シーンを、ARIB TR-B43 [4]の参考資料 5 に記載の方法で SDR にトーンマッピングしたもの |

注記: VVC エンコーダ A(後述)が未対応のため、1080/60/I や 4320/60/P は本実験の対象外とする。なお、これら映像形式を規格策定対象から外すことを意味しない。120/P も同様である。

注記: カテゴリ 4 は、カテゴリ 3 がカテゴリ 2 と同等の符号化難易度であることを検証するために用いる。

各カテゴリの色空間、画素値精度、シーン長はそれぞれ、Y'CbCr 4:2:2, 10-bit, 10 sec とする。シーン毎に、開始フレーム時刻を定める。

3.3. VVC エンコーダ

実放送の所要ビットレートを導出する目的上、放送開始時に実現可能性が高いエンコーダを用いることが望ましい。そこで、2025年度にハードウェアで実現可能なリアルタイムエンコーダの画質を実現するシミュレータ(下表のエンコーダ A)を主に使用する。併せて、本実験の妥当性検証のために、VVC 規格の標準化作業で開発された参照ソフトウェアの最適化版(下表のエンコーダ B)を補助的に用いる。

| エンコーダ | 説明 |
|-------|--|
| A | NEC 社 VVC エンコーダシミュレータ 総務省の技術試験事務にて開発した、2025年頃にハードウェアで実現可能なリアルタイムエンコーダの画質を実現するソフトウェアシミュレータ 生成ビットストリームは VVC 参照ソフトウェア VTM-10.0 [9]で復号可能 |
| B | Fraunhofer HHI 研究所 VVC ソフトウェアエンコーダ VVenC [10] VVC 規格の標準化作業で開発された、最高性能を達成する参照ソフトウェア VTM (VVC Test Model)の最適化版 リアルタイム動作よりも高性能化を主眼としており、将来の放送サービスにそのまま適用することは難しい 性能と処理速度とのバランスが取れた medium 設定を使用(MCTFは無効化) 生成ビットストリームは VVC 参照ソフトウェア VTM-10.0 で復号可能 BSD ライセンスの元で利用可能 |

3.3.1. エンコードするカテゴリと評価方法

| エンコーダ | カテゴリ 1 (2K/SDR) | カテゴリ 2 (4K/SDR) | カテゴリ 3 (4K/HDR) | カテゴリ 4 (4K/SDR) |
|---------|----------------------|-----------------------|-----------------|-----------------|
| A (NEC) | ○(主観評価) ○(客観評価*1) | ○(主観評価) ○(客観評価*1) | ○(主観評価) | × |
| B (HHI) | ○(客観評価*1) | ○(客観評価*1) (主観評価*3) | × | ○(客観評価*2) |

注記 1: 選定シーン以外も符号化し、付録 1 に示した PSNR 分布と同様な結果となることを確認する。

注記 2: エンコーダ B によるカテゴリ 2 の符号化と比較して、相対的な PSNR レンジを確認する。

注記 3: 符号化難易度が高い 4 シーンについて主観評価を行い、エンコーダ A の結果との差を確認する。

3.4. VVC 符号化条件

今回の実験で設定するビットレートは、高度広帯域衛星デジタル放送の映像符号化方式検討時の値を参考に決めた。これよりも低いビットレートでの評価実験は別途(時期は未定)とする。その際には、今回の実験で有効にしていないピクチャレベルツール(プレフィルタ、動的解像度変換等、使用するエンコーダがサポートしていない、主観画質向上ツール)の適用が想定される。

| 項目 | 設定 | |
|----------------------------|-------------------------|--|
| プロファイル | Main 10 (10-bit, 4:2:0) | |
| ビットレート [Mbps] (*1) (*2) | 1080/60/P 2160/60/P | 3, 5, 7, 10 10, 15, 20, 30 |
| 符号化パラメータ | GOP 構造 | 階層 B 参照 ARIB STD-B32 [11]に記載の L3 構造 |
| | IRAP 間隔 | 32/60 sec |
| | GOP 長 | 8 フレーム |
| | CPB サイズ | 1 秒分 |
| CU レベル無効化 ツール(*3) | エンコーダ A | スクリーンコンテンツ向けツール(IBC)は無効化 他のツールは非公開 |
| | エンコーダ B | 付録参照 |

注記 1: 高度広帯域衛星デジタル放送の映像符号化方式検討時(HEVC 方式)のビットレート幅は 1080/60/P, 2160/60/P でそれぞれ、10 Mbps – 15 Mbps, 30 Mbps – 40 Mbps であった。高ビットレートレンジでは VVC 方式適用により 30%のビットレート削減が確実に見通せることから、高ビットレート側の値をそれぞれ 10 Mbps ($\approx 15 \times 0.7$), 30 Mbps ($\approx 40 \times 0.7$)とした。低ビットレートレンジの値は、これらの値の 70%, 50%, 30%とした。

注記 2: エンコーダ B での客観評価向け符号化時には、1080/60/P, 2160/60/P それぞれ 8 Mbps, 20 Mbps のみとする。エンコーダ B での主観評価向け符号化時には、1080/60/P, 2160/60/P それぞれ、3 Mbps 及び 7 Mbps, 10 Mbps および 20 Mbps とする。

注記 3: 本実験ではピクチャレベルツールである、画面分割(Tile)、動的解像度変換(RPR)、プレフィルタ(MCTF)は無効にしている。運用規定の議論に向け、今後評価する。CU(Coding Unit, 符号化単位となるブロック)レベルの各ツールは一部無効化して実験するが、運用規定では全て利用可能とする方向で議論を進めている。

3.5. 主観評価実験

| 項目 | 内容 | |
|--------|---|----------------------------|
| 評価方法 | 二重刺激劣化尺度(DSIS)法、5段階(1-5)劣化尺度 基準映像—評価映像のペアを一回提示 | |
| 観視条件 | Rec. ITU-R BT.500-14 [3] | |
| 評価者 | 専門家 15人以上(スクリーニングあり) | |
| ディスプレイ | 1080/60/P | ソニー社製 55-inch LCD PVM-X550 |
| | 2160/60/P | ソニー社製 55-inch LCD PVM-X550 |
| 視距離 | 1080/60/P | 画面高の3倍(3H) |
| | 2160/60/P | 画面高の1.5倍(1.5H) |

3.5.1. 所要ビットレート推定の基準

放送品質を満足するサービスを提供するためには、画質の許容限界と見なされる平均評価値(Mean Opinion Score, MOS) 3.5以上がほぼ全ての映像で満足することが望まれ、また MOS 3未満は放送品質として許容しがたいと考えられる。そこで、平均値と分散を元に、MOS 3.5以上及び MOS 3未満に該当するか否かを、有意水準 5%で検定する。

3.6. 準備スケジュール

2021年1月: カテゴリ3映像(4K, HDR)のトーンマッピング

2021年3月: エンコーダ Bによる符号化処理

2021年4月: エンコーダ Aによる符号化映像入手、主観評価実験用映像編集

2021年6月: 主観評価実験実施、報告書作成

付録1 主観評価実験に用いるシーン

1. カテゴリ 1 (1080/60/P, SDR)

| | | | |
|---|---|--|---|
|  |  |  |  |
| No. 201 Ginkgo trees | No. 202 Truck train | No. 204 Red leaves (pan up) | No. 209 Fountain (dolly) |
|  |  |  |  |
| No. 210 Studio concert | No. 214 Basketball | No. 218 Horse racing (dirt) | No. 265 Fountain (chromakey) |

2. カテゴリ 2 (2160/60/P, SDR)

注記:A.7 “River”は難易度が非常に高いため、事前に符号化映像を見た上で採用を判断する。

| | | | |
|---|---|--|---|
|  |  |  |  |
| A.3 Trains C | A.6 Festival | A.7 River | B.6 Paddock |
|  |  |  |  |
| B.7 Marathon (start) | B.9 Marathon (panning) | B.11 Water polo (Scrolling) | B.13 Drama (coffee) |

3. カテゴリ 3 (2160/60/P, HDR)

| | | | |
|---|---|--|---|
|  |  |  |  |
| C.1 Fireworks (willow) | C.5 Fireworks (barrage) | C.6 Drama (standing up) | C.8 Drama (sunset) |
|  |  |  |  |
| C.11 Swim race (backstroke) | C.12 Volleyball (fixed) | C.15 Paddock (fixed) | C.17 Horse race (homestretch) |

4. (補足)シーン選定基準概要

高度広帯域衛星デジタル放送の映像符号化方式検討の際(以下、HEVC 時)と同様に、多様な符号化難易度や絵柄を含む映像セットを選定する。

符号化難易度は、VTM を用いて符号化した場合の輝度信号の PSNR とする。

下図は、HEVC 時の符号化難易値(及び選定シーン)である。符号化難易度が高いシーンを中心に、中間のシーンも含めている。

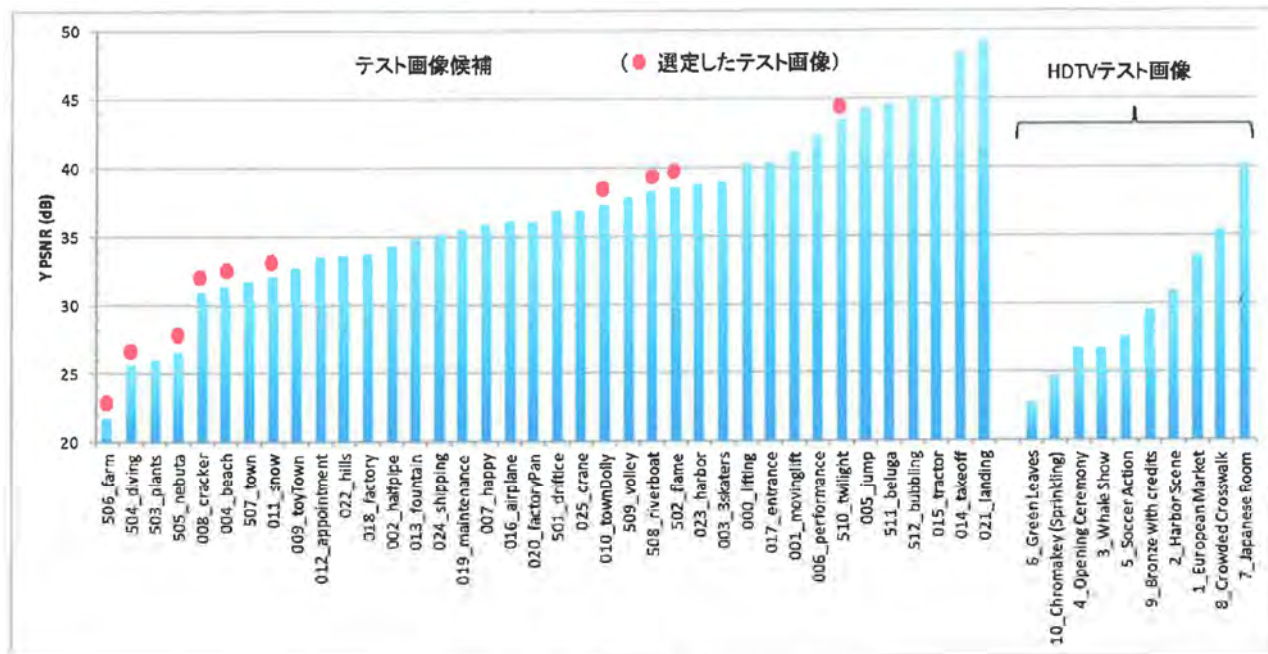


図 A3-1 テスト画像の符号化難易

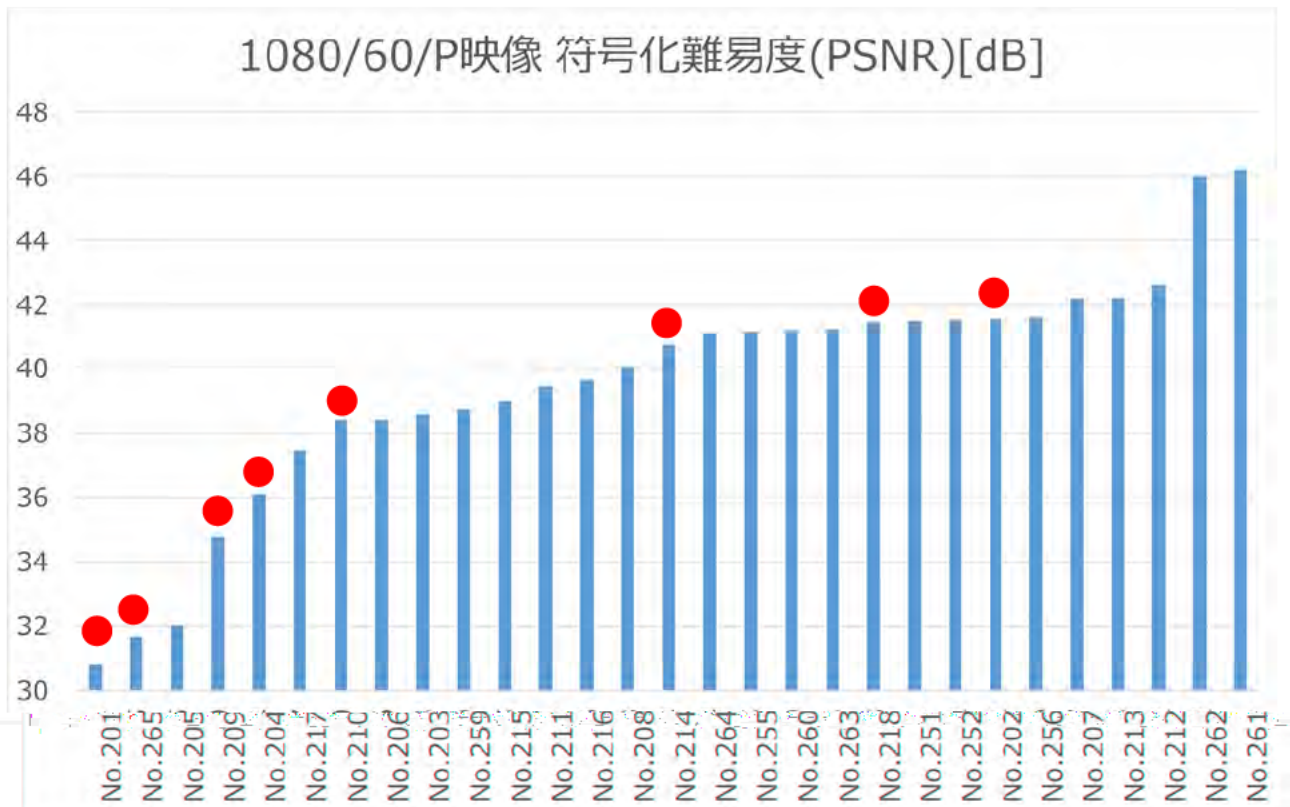
4.2. 映像情報メディア学会標準動画の VVC 符号化難易度

日本電気(株)より、VTM 符号化時の符号化難易度データを提供されており、この値を元にシーン選定を行った。以下、各カテゴリでの符号化難易度分布と選定シーン(赤丸)を示す。

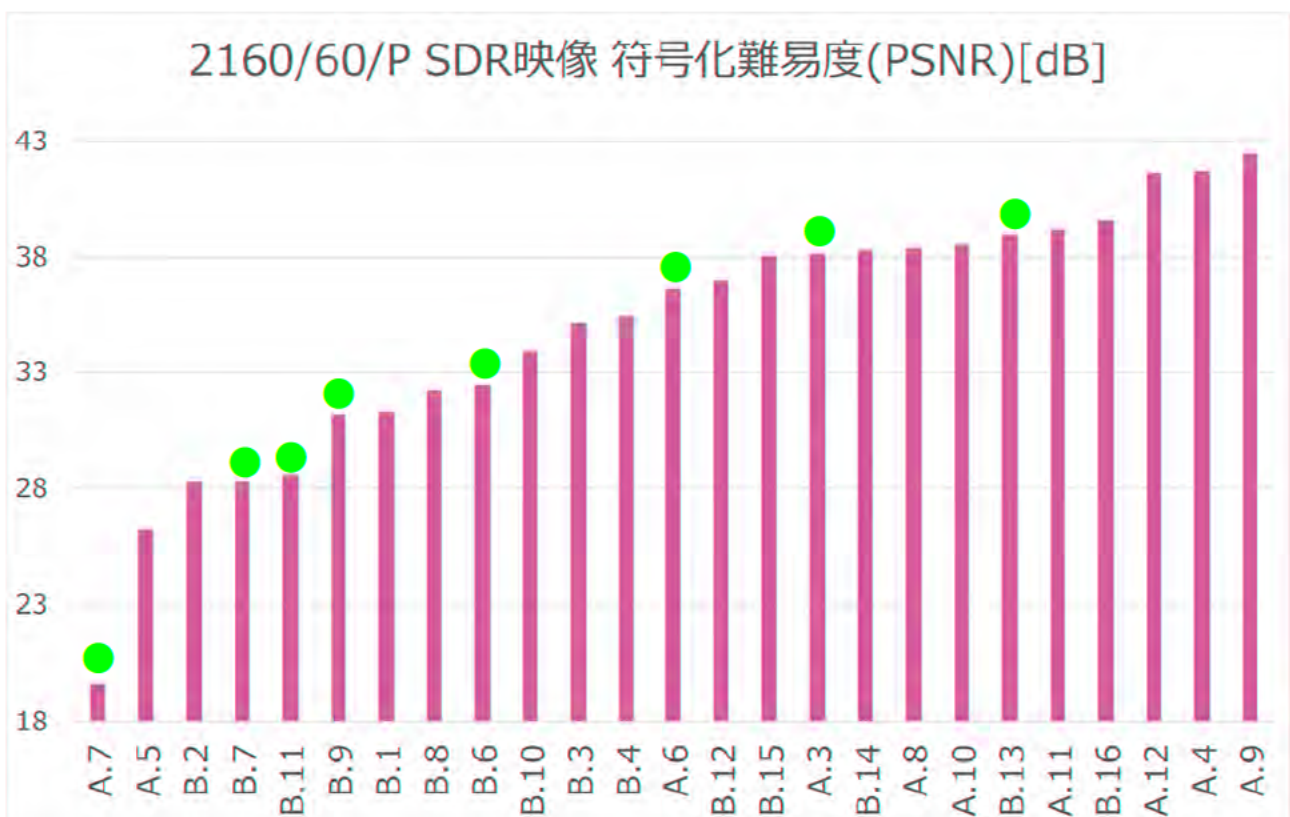
符号化条件は以下の通り。

| 項目 | 設定 |
|-----------|--|
| VTM バージョン | 6.0 |
| 符号化モード | Random access |
| ビットレート | 8 Mbps (1080/60/P), 20 Mbps (2160/60/P) 注記: レート制御を行う代わりに、複数の QP 値での固定 QP 符号化を行い、設定ビットレートでの PSNR を推定 |
| シーン長 | 65 フレーム 注記: 先頭 7 秒目から 8 秒目まで。ただしこの時間帯にシーンチェンジ等が含まれるシーンについては、10 秒目から 11 秒目まで |

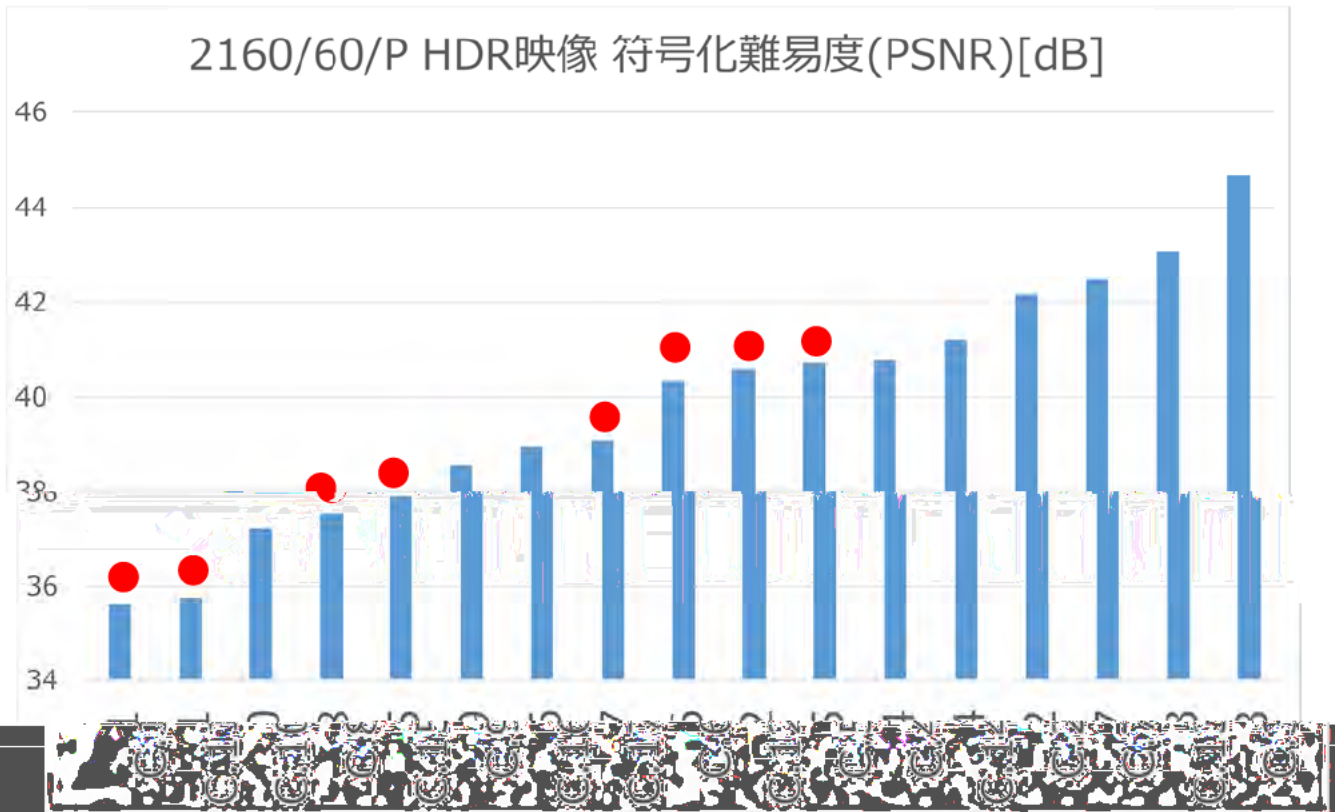
4.2.1. カテゴリ 1 (1080/60/P, SDR)の符号化難易度



4.2.2. カテゴリ 2 (2160/60/P, SDR) の符号化難易度



4.2.3. カテゴリ 3 (2160/60/P, HDR)の符号化難易度



付録2 VVenC の性能とサポートツール

VVenC (v0.2)には5つの動作モード”Slower”, ”Slow”, “Medium”(今回の実験で採用), “Fast”, “Faster”がある。

下表は、各動作モードの性能である。“ Δ PSNR”及び“Speed-up”はそれぞれ、HEVC (HM-16.22)からのBD-Rate 削減率、VVC (VTM-11.0)からの速度向上度である。

| 動作モード | HD(2K) | | UHD(4K) | |
|--------|---------------|----------|---------------|----------|
| | Δ PSNR | Speed-up | Δ PSNR | Speed-up |
| Faster | -6.7% | 240x | -15.8% | 380x |
| Fast | -26.8% | 160x | -29.4% | 230x |
| Medium | -37.7% | 70x | -39.3% | 110x |
| Slow | -40.8% | 27x | -42.1% | 43x |
| Slower | -44.0% | 5.2x | -45.4% | 8.8x |

下表は、有効・無効がビットストリームレベルで設定可能なツールと、各動作モードでの有効(Yes)、無効(No)の関係をまとめたものである。各ツールの概要は本報告の第二部を参照されたい。

| Tool/Setting | Slower | Slow | Medium | Fast | Faster |
|--------------|--------|------|--------|------|--------|
| Chroma TS | No | No | No | No | No |
| MTS | Yes | No | No | No | No |
| LFNST | Yes | Yes | Yes | No | No |
| Joint CbCr | Yes | Yes | Yes | No | No |
| SAO | Yes | Yes | Yes | Yes | Yes |
| ALF | Yes | Yes | Yes | Yes | No |
| CCALF | Yes | Yes | Yes | Yes | No |
| LMCS | Yes | Yes | Yes | No | No |
| TMVP | Yes | Yes | Yes | Yes | Yes |
| SbTMVP | Yes | Yes | Yes | No | No |
| AMVR | Yes | Yes | Yes | No | No |
| BDOF | Yes | Yes | Yes | Yes | No |
| SMVD | Yes | Yes | Yes | No | No |
| DMVR | Yes | Yes | Yes | Yes | No |
| MMVD | Yes | Yes | Yes | No | No |
| SBT | Yes | Yes | No | No | No |
| Affine | Yes | Yes | Yes | No | No |
| Affine AMVR | No | No | No | No | No |
| PROF | Yes | Yes | Yes | No | No |
| BCW | No | No | No | No | No |
| CIIP | Yes | Yes | No | No | No |
| GPM | Yes | Yes | Yes | No | No |
| ISP | No | No | No | No | No |
| MRL | Yes | Yes | Yes | No | No |

| Tool/Setting | Slower | Slow | Medium | Fast | Faster |
|--------------|--------|------|--------|------|--------|
| MIP | Yes | Yes | Yes | No | No |
| CCLM | Yes | Yes | Yes | Yes | Yes |
| IBC | No | No | No | No | No |
| DQ | Yes | Yes | Yes | No | No |
| Dual-Tree | Yes | Yes | Yes | Yes | Yes |
| MCTF (*) | Yes | Yes | Yes | Yes | No |

(*)動き補償付き時間方向ノイズ除去フィルタ。入力映像に適用するプレフィルタであり、標準には含まれていないものの、JVET 会合にて 3%程度の BD-Rate 削減が報告されている。

付録3 VVC 実用化動向・予測

映像符号化の高度化に必要な技術的検討にあたり、放送サービスが開始される時期(2025年を想定)の、VVCに準拠したエンコーダ及びデコーダの技術動向及び技術予測に関する以下の情報提供を、当作業班委員に依頼した。以下に提供された情報をまとめる。

1. 半導体の技術進化予測
2. VVC 実用化予測
3. 放送以外の VVC サービス予測

1. 半導体の技術進化予測

情報提供元:パナソニック社

| 項目 | 予測 |
|---------|---|
| メモリバンド幅 | 現在の最先端の高速メモリは DDR4-3200 (25.6Gbps)が主流であり、2025 年頃は DDR5 へのシフトにより 2 倍程度の高速化が予測される。一般的な民生用の組込向け SoC やコーデック LSI では一世代前のメモリや低消費電力メモリが主流であるが、同様に 2 倍程度の高速化が進むものと予想される。VVC のワーストメモリバンド幅は HEVC 相当が保障されているため、同じ画像サイズであれば現状の HEVC と同等のメモリ構成(例えば 4320/60/P で 4 チャンネル)で実現可能であり、さらに 2025 年以降にはメモリのチャンネル数を減らすことが可能となりチップ面積およびコストの削減が進む可能性が高い。 |
| プロセス | 現在の最先端プロセスは 7nm 程度であるが、2025 年頃は 4nm – 5nm 程度となり 2 – 3 倍の高密度化が予想される。一般的な民生用の組込向け SoC やコーデック LSI ではまだ 15nm – 25nm のプロセスがボリュームゾーンとしては主流であるが、同様に 3 – 4 倍程度の高密度化が進むものと予想される。VVC の回路規模は、デコーダで HEVC の 2 倍程度、エンコーダで 4 倍程度が予想され、チップサイズは解決可能な範囲である可能性が高い。チップコストに関しては、2025 年の時点では HEVC と比較して割高となるが、以降は普及価格帯に落ち着くと予想される。 |
| 動作周波数 | プロセス高密度化、消費電力・発熱の課題から、2025 年時点でも大幅な動作周波数の向上は見込めないと予想される。ブロック単位処理の並列化実装を進めても、1つのパイプラインステージに要求されるワーストのサイクル数の削減には限界があり、また CABAC 処理のスループットの向上も限定的であるものと予想される。従って、2025 年時点でも 4320/60/P 以上のエンコーダ・デコーダは、1チップ内でのマルチコア構成のチップが市場に出回る可能性が想定される。 |

2. VVC 実用化予測

2.1. HEVC 実用化の実績

情報提供元:パナソニック社

HEVC 初期の国内の民生用受信機向け LSI のスケジュール

- 2013 年 1 月: HEVC 第 1 版の最終国際標準ドラフト化(仕様確定)
- 2014 年後半: 受信機向けデコーダ 2160/60P, 4:2:0/10-bit(1 コア×1 チップ構成)
- 2016 年後半: 受信機向けデコーダ 4320/60/P, 4:2:0/10-bit(4 コア×1 チップ構成)

2.2. VVC 実用化予想時期

情報提供元:パナソニック社、池上通信機社

| 映像フォーマット | パナソニック社(*1)(*2) | 池上通信機社 |
|--------------------------|-----------------------------|---|
| 1080/P/60 | Main 10 デコーダ:2022 年度中 | デコーダ:2020 – 2021 年度 ソフトウェア、GPU での実現 デコーダ:2021 – 2022 年度 LSI での実現 エンコーダ:2023 – 2024 年度 LSI での実現 |
| 2160/P/60 | Main 10 デコーダ:2022 年度中 | エンコーダ・デコーダ:2021 – 2022 年度 複数 FPGA での実現 エンコーダ・デコーダ:2023 – 2024 年度 (複数)LSI での実現 |
| 4320/P/60 | Main 10 デコーダ:2022 – 2024 年度 | 市場要求に依存し不明 |
| 2160/P/120 4320/P/120 | 不明(*3) | 回答無し |

注記1: Multilayer プロファイルは、初期世代では対応が見送られることが想定されるが、マルチコア構成の活用により、限定された条件の下で将来的に対応が進んでいく可能性がある。ただし、検証および制御層の開発工数を鑑み、市場ニーズに合わせながら実対応が進んでいくものと予想する。

注記2: 12-bit 対応は、完全にハードが別設計となるため、HEVC 世代と同様に 10-bit 対応とは別の時間軸で特定用途向けとして遅れて開発が進んでいくものと予想する。

注記3: コア数とメモリチャンネル数を積み上げれば技術的には実現可能であるが、コストが見合わないため民生用受信機として量産化可能となるまでには時間がかかると予想する。

2.3. 2024 – 2025 年の VVC 実用化予想

情報提供元:ソニー社

| 項目 | 予測 |
|-----------|---|
| 製品 | 受信機用 SoC (ASIC による実現) が市場に登場する |
| フレームレート | 120/P まで |
| 画素精度 | 放送伝送用は 10-bit、業務用 (番組交換フォーマット等) は 12-bit |
| ダイナミックレンジ | 放送伝送用、業務用共 SDR と HDR をサポート |
| 走査方式 | エンコーダ実装時の検証コストが下がることから、Progressive のみ対応のエンコーダが想定される |
| プロファイル | 放送伝送用は Main10、業務用 (番組交換フォーマット等) は VVC version 2 |
| 階層符号化 | 120/P – 60/P 時間方向階層符号化 |

2.4. 技術課題

情報提供元:パナソニック社、池上通信機社、ソニー社

| 会社名 | 課題 |
|---------|---|
| パナソニック社 | メモリバンド幅およびチップサイズに起因するコスト課題はある程度解決されると予想するが、動作周波数限界による処理可能ブロック数がボトルネックとなり、4320/60/P 以上のエンコーダ・デコーダでは、引き続き1チップ内でのマルチコア構成のチップが市場に出回る可能性あり。同様に、CABAC スループットの限界による、ビットレート制限およびバッファリング遅延の余裕度にも注意が必要。 |
| 池上通信機社 | エンコーダの各種モードセレクトやレート制御パラメータ導出の高効率なリアルタイム判定アルゴリズム開発 |
| ソニー社 | 市場からの要請と実装・サービスコストのトレードオフから、技術課題が特定される |

3. 放送以外の VVC サービス予測

情報提供元:パナソニック社、池上通信機社、ソニー社、フジテレビジョン社

| 会社名 | 予想 |
|-----------|---|
| パナソニック社 | VVC 第 1 版は、遠隔コミュニケーション、ネット配信サービス、監視カメラ、バーチャルリアリティー等に展開する。更に VVC 第 2 版により、業務用カメラ、医療用途等へ拡張されていく。 |
| 池上通信機社 | 放送の素材伝送 映像監視市場(但し低コストの LSI が必要) ネット動画配信(スマートフォン用 SoC にデコーダが実装できる時期) |
| ソニー社 | 録画メディアあるいは見逃しへの対応 ストリーミング パッケージメディア(進化の可能性?) |
| フジテレビジョン社 | 素材伝送、アーカイブ及び番組パッケージ販売、効率的な番組交換フォーマット 放送通信連携サービスによる映像配信や、VOD 等の動画配信サービス デジタルシネマ・パブリックビューイング等への配信、クラウドゲームの画面転送等 |

付録4 新型コロナウイルス感染対策

複数人が実験室(密閉空間)に入って実施する主観評価実験を行うにあたり、新型コロナウイルスへの適切な感染対策が必要である。

「感染拡大の防止と研究活動の両立に向けたガイドライン(改訂)」(令和2年10月6日 文部科学省)の1章「研究室・執務室等での活動(学生等の研究室活動を含む。)」について」の記載事項を参考に、以下の対策を講じる。

| 項目 | 施策 |
|--|--|
| いわゆる 3 密(換気の悪い密閉空間、多数が集まる密集場所、近距離での会話や発話が生じる密接場面)の回避 | 主観評価実験を実施する実験室、控室での空調稼働 各セッションでの参加者、実験管理者の人数の制限(実験管理者2名、実験参加者4ないし6名) マイク(もしくは録音装置)及びスピーカーを用いたガイダンス実施 |
| 十分な対人距離(1m以上)の確保 | 各モニタ間の十分な間隔を確保 |
| 水と石鹸による手洗いの徹底 | 施設内のトイレでの手洗い実施 |
| 入口及び施設内の手指の消毒設備の設置 | 実験室及び控室入口に手指消毒液を設置 |
| マスクの着用 | 参加者及び実験管理者への徹底周知 |
| 施設の換気 | 空調稼働 各セッションの終了後、扉を開けて換気を実施 |
| アクリル板・透明ビニールカーテン等の設置 | 各モニタ間、同一モニタ視聴者間(*)、実験管理者の前にアクリル板を設置 (*)モニタ視聴時の妨げにならないように工夫 |
| 施設(ドアノブ・エレベータボタン等)の消毒 | 実験実施前に実施 |
| 症状(発熱や風邪症状等)のある方等の入場制限 | 実験参加者への念書に、入場制限の旨を記載 非接触型体温測定器を用いた体温測定の実施 |
| 行動記録や健康記録の取得 | 人の流れを記録(氏名、時間、場所) |
| 研究活動に専念できる環境の整備 | 複数の控室の確保 データ解析はリモートで実施 |

地上デジタル放送方式高度化に関わる
適用技術検討作業 中間報告

第二部:情報源符号化部 H.266 | MPEG-I VVC 規格
の概要

2021年2月9日

デジタル放送システム開発部会
一般社団法人 電波産業会

まえがき

総務省からの諮問第 2044 号「放送システムに関する技術的条件」(2019 年 6 月 18 日)を受け、情報通信審議会放送システム委員会に地上デジタル放送方式高度化作業班が設置され、技術的条件の検討が始まった。本活動の一環として、2020 年 6 月 22 日に、ARIB に対して映像符号化及び音声符号化方式の高度化に必要な技術的検討の依頼があった。

本依頼を受け、映像符号化方式作業班は、映像符号化方式の検討を開始している。2020 年には、最新の映像符号化規格 VVC を有力候補とし、規格調査を進めると共に、VVC 規格を用いた際の所要ビットレートを求めるための主観評価実験の計画案を作成した。

本報告書は、VVC 規格の主観評価実験案、及び規格調査結果を中間報告としてまとめたものである。前者は第一部に、後者(本資料)は第二部に分けている。

目次

| | |
|--|----|
| 1. 標準化経緯..... | 6 |
| 2. 方式の概要..... | 7 |
| 2.1. 共通符号化ツール..... | 7 |
| 2.2. スクリーンコンテンツ向け符号化ツール..... | 12 |
| 2.3. HDR/WCG コンテンツ向けの符号化特徴..... | 13 |
| 2.4. ハイレベル機能..... | 14 |
| 3. 符号化データ構造..... | 15 |
| 3.1. ビットストリーム構造..... | 15 |
| 3.1.1. ビットストリーム、シーケンス、アクセスユニット、ピクチャユニット..... | 16 |
| 3.1.2. レイヤ..... | 16 |
| 3.1.3. NAL ユニット..... | 17 |
| 3.1.4. ピクチャ種別..... | 19 |
| 3.1.5. 参照ピクチャリサンプリング..... | 19 |
| 3.2. ピクチャ分割..... | 20 |
| 3.2.1. タイル、スライス、サブピクチャ..... | 20 |
| 3.2.2. ウェーブフロント並列処理..... | 21 |
| 3.2.3. 仮想境界..... | 22 |
| 3.3. ブロック分割..... | 22 |
| 3.3.1. CU 分割..... | 22 |
| 3.3.2. Dual Tree..... | 24 |
| 4. イントラ予測..... | 24 |
| 4.1. 輝度イントラ予測..... | 24 |
| 4.2. 色差イントラ予測..... | 26 |
| 4.3. MRL (Multiple Reference Line)..... | 26 |
| 4.4. CCLM (Cross-component Linear Model)..... | 27 |
| 4.5. PDPC (Position Dependent intra Prediction Combination)..... | 28 |
| 4.6. MIP (Matrix-based Intra Prediction)..... | 29 |
| 4.7. ISP (Intra Sub-Partitions)..... | 30 |
| 5. インター予測..... | 31 |
| 5.1. AMVP モード..... | 31 |
| 5.1.1. AMVP (Adaptive Motion Vector Prediction)..... | 31 |
| 5.1.2. AMVR (Adaptive Motion Vector Resolution)..... | 32 |

| | | |
|--------|--|----|
| 5.1.3. | SMVD (Symmetric Motion Vector Difference)..... | 32 |
| 5.2. | マージモード | 33 |
| 5.2.1. | Regular Merge | 33 |
| 5.2.2. | MMVD (Merge mode with Motion Vector Difference)..... | 33 |
| 5.2.3. | CIIP (Combined Inter-picture merge and Intra-picture Prediction) | 34 |
| 5.2.4. | GPM (Geometric Partitioning Mode) | 35 |
| 5.2.5. | DMVR (Decoder-side Motion Vector Refinement)..... | 36 |
| 5.3. | サブブロックモード | 37 |
| 5.3.1. | アフィンモード..... | 37 |
| 5.3.2. | PROF (Prediction Refinement with Optical Flow)..... | 38 |
| 5.3.3. | SbTMVP (Subblock-based Temporal MVP) | 39 |
| 5.4. | 動き補償 | 40 |
| 5.4.1. | MC (Motion Compensation)..... | 40 |
| 5.4.2. | SIF (Switched Interpolation Filter)..... | 40 |
| 5.4.3. | Wrap around MC (Wrap around Motion Compensation) | 41 |
| 5.4.4. | RPR (Reference Picture Resampling)..... | 41 |
| 5.5. | 予測画像生成..... | 42 |
| 5.5.1. | BDOF (Bi-Directional Optical Flow) | 42 |
| 5.5.2. | BCW (Bi-prediction with CU-level Weights)..... | 43 |
| 5.5.3. | WP (Weighted Prediction) | 43 |
| 6. | 変換・量子化..... | 43 |
| 6.1. | 変換..... | 43 |
| 6.1.1. | MTS (Multiple Transform Selection) | 43 |
| 6.1.2. | SBT (Subblock Transform)..... | 45 |
| 6.1.3. | LFNST (Low Frequency Non-Separable Transform)..... | 46 |
| 6.2. | 量子化..... | 47 |
| 6.2.1. | QP (Quantization parameter) | 47 |
| 6.2.2. | スケーリング行列 | 48 |
| 6.2.3. | DQ (Dependent quantization)..... | 48 |
| 6.3. | 変換スキップおよび可逆符号化 | 49 |
| 7. | ループ内フィルタ | 50 |
| 7.1. | LMCS (Luma Mapping with Chroma Scaling)..... | 50 |
| 7.1.1. | 輝度マッピング | 51 |
| 7.1.2. | 色差スケーリング | 52 |
| 7.2. | DBF (Deblocking Filter) | 53 |

| | | |
|--------|--|----|
| 7.2.1. | LADF (Luma Adaptive Deblocking Filter) | 54 |
| 7.2.2. | LTDF (Long Tap Deblocking Filter) | 55 |
| 7.3. | SAO (Sample Adaptive Offset) | 55 |
| 7.4. | ALF (Adaptive Loop Filter)・CCALF (Cross-Component ALF) | 55 |
| 7.4.1. | ALF におけるフィルタ形状 | 55 |
| 7.4.2. | 輝度 ALF におけるクラス分け | 56 |
| 7.4.3. | フィルタの変形 | 56 |
| 7.4.4. | フィルタ処理 | 57 |
| 7.4.5. | CC-ALF | 57 |
| 7.4.6. | CTU 境界における処理 | 58 |
| 8. | エントロピー符号化 | 59 |
| 8.1. | CABAC 符号化 | 59 |
| 8.2. | 残差係数符号化 | 60 |
| 8.2.1. | RRC (Regular Residual Coding) | 60 |
| 8.2.2. | TSRC (Transform Skip Residual Coding) | 62 |
| 8.2.3. | JCCR (Joint Coding of Chroma Residual) | 64 |
| 9. | スクリーンコンテンツ符号化ツール | 64 |
| 9.1. | IBC (Intra Block Copy) | 64 |
| 9.2. | BDPCM (Block Differential Pulse Coded Modulation) | 65 |
| 9.3. | Palette mode | 65 |
| 9.4. | ACT (Adaptive Color Transform) | 66 |
| 10. | プロファイルとレベル | 66 |
| 10.1. | プロファイル | 66 |
| 10.2. | レベル | 69 |
| 11. | 参考文献 | 71 |

1. 標準化経緯

H.266 / MPEG-I VVC (Versatile Video Coding、以下 VVC) は、H.264 / MPEG-4 AVC (以下 AVC) や H.265 / MPEG-H HEVC (以下 HEVC) よりもさらに高効率な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) 傘下の SG16 / WP3 / Q.6 の呼称である VCEG (Video Coding Experts Group) 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 JCT1 (Joint Technical Committee 1: 第1合同技術委員会) 傘下の SC29 / WG11 の呼称である MPEG (Moving Picture Experts Group) の合同組織として 2015 年 10 月に設立された JVET (Joint Video Exploration Team) 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.266」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-I VVC」は ISO / IEC JCT1 のマルチメディア符号化規格の作業グループの通称である 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-I シリーズの Part3 (ISO/IEC 23090-3) として位置付けられている。

VCEG と MPEG との合同組織である JCT-VC (Joint Collaborative Team on Video Coding) により 2013 年 1 月に HEVC の第 1 版の標準化を完了した後、JCT-VC は通信・ネットワークにおける動画トラフィックの急激な増加や VR 等の動画コーデックのユースケースの拡大などから、HEVC を超えた圧縮効率への市場要求が高まることを確認し、KTA (Key Technical Area) と呼ばれる新しい符号化技術の検証を開始した。そして、2015 年 10 月に VCEG と MPEG は、KTA を継承した合同組織である JVET (Joint Video Exploration Team) を設立し、新しい符号化方式の規格化を目指した活動を本格化した。JVET の活動を通じて HEVC に対して有意な性能差を実現できることを確認したことを受けて、2017 年 10 月に技術公募 (Call for Proposals) を発行し、VVC の正式な標準化活動を開始した。

JVET では VTM (VVC Test Model) と呼ばれる参照ソフトウェアを会合毎に更新しながら開発を進め、同じく HEVC の参照ソフトウェアである HM をアンカーとして、符号量削減率、符号化処理時間増加率、および復号処理時間増加率を統合的に評価しながら技術開発を行うことで高圧縮効率と低処理量の両立を目指した。最終的に 4K UHD 素材において、HM 比で符号量削減率約 41%、符号化処理時間約 9.4 倍、復号処理時間約 1.9 倍となっている。そして、2019 年 7 月の標準化委員会草案 (Committee Draft) を経て、2020 年 7 月に最終規格草案 (Final Draft International Standard) を発行して、VVC 第 1 版の標準化を完了した。

第 1 版では、幅広いアプリケーションに適用可能な 4:2:0 かつ 10-bit 以下の映像信号に対応する Main 10 プロファイル、より高品質な 4:2:2 および 4:4:4 の映像信号に対応する Main 10 4:4:4 プロファイル、静止画像信号の圧縮に限定した Main 10 Still Picture プロファイルと Main 10 4:4:4 Still Picture プロファイル、空間解像度や多視点映像等のスケーラビリティに対応した Multilayer Main 10 プロファイルと Multilayer Main 10 4:4:4 プロファイルの 6 つのプロファイルが策定されている。

さらに、この第 1 版発行に続いて、より一層の高画質化や機能拡張を目的とした追加検討を行うことが議論されている。具体的には、業務用途の制作・伝送を目的とした 12-bit 以上のサンプリングや超高ビットレートおよび超高フレームレートに対応するプロファイルの策定、Neural Network を用いた高画質化や機能拡張の検討を進めていくことが想定されている。

一方、技術面での標準化とは別に、VVC のプロモーションやライセンスを含めたエコシステムを議論する業界団体である MC-IF (Media Coding Industry Forum) が組織されている。MC-IF では、IP ライセンスのエコシステム形成を目指した検討、業界からの要求に基づいたサブプロファイル設定に関する議論、プロモーション活動、相互運用性の促進等の取り組みが進められている。

2. 方式の概要

2.1. 共通符号化ツール

本頁では、VVC 符号化において重要なエッセンスであるプロファイルの拡張・適応化に関する主要技術について、AVC や H.264 などの従来の符号化方式と比較しながら簡単に説明する。なお、各ツールの詳細については 3 章以降に記述する。

図 2-1 に VVC エンコーダの概略構成を示す。基本的なブロック構成は AVC や H.264 と類似しているが、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。入力の映像信号に関しては、Main 10 プロファイルは 8-bit と 10-bit のモノクロおよび YUV4:2:0 フォーマットに、Main 10 4:4:4 プロファイルは 8-bit と 10-bit のモノクロ、YUV4:2:0、YUV4:2:2、YUV4:4:4 (RGB 等も含む) フォーマットに対応している。なお、HEVC 同様にプログレッシブ素材に加えてインターレース素材もサポートしている。

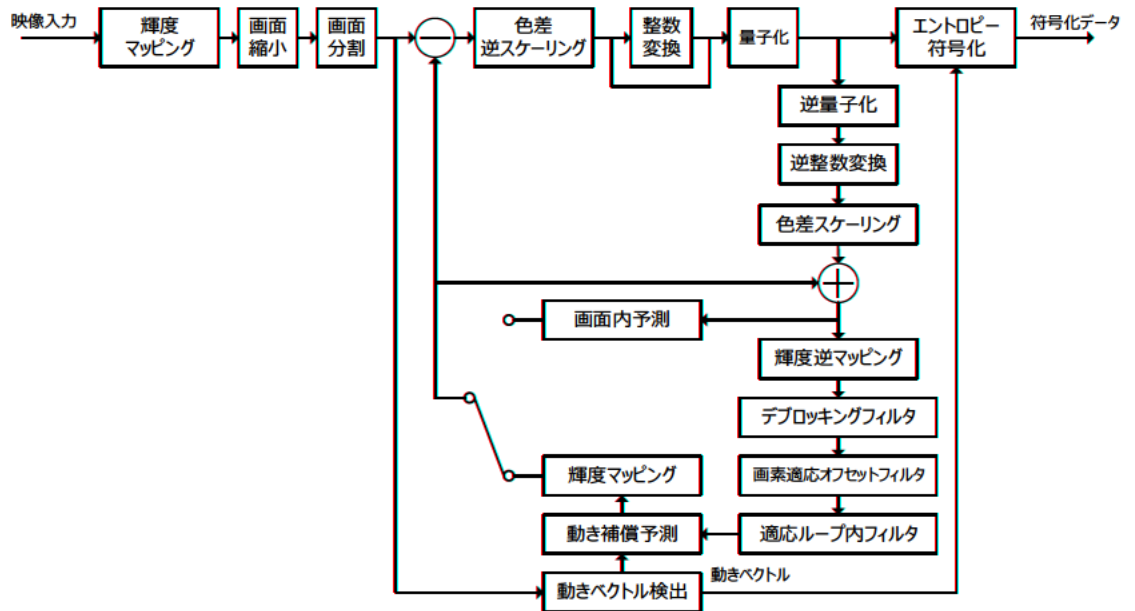


図 2-1 VVC エンコーダの概略構成

ブロック分割

VVC では各ピクチャの符号化処理を、HEVC と同様に CTU (Coding Tree Unit) と呼ばれる正方面素ブロックの単位で行う。CTU は MPEG-2 や AVC における MB (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 HEVC では最大 64x64 画素ブロックであったのに対し、VVC では最大 128x128 画素ブロックに拡張されている。

イントラ・インター予測、変換・量子化、エンタロピー符号化などの一連の符号化処理は、CTU を再帰的に木構造分割することによって得られる CU (Coding Unit) の単位で行う。各ノードでの分割パターンは、HEVC では 4 分木のみであったのに対し、VVC ではさらに 3 分木および 2 分木 (水平もしくは垂直方向に分割) が使用可能となっている。映像信号の局所的な特性に応じて CTU を多様な大きさ・形状の CU に分割することができるため、符号化効率の更なる向上が可能となった。また、イントラスライスでは輝度成分と色差成分とで個別の木構造を用いる Dual tree も選択可能となっている。一方、CU サイズのバリエーションが増えたことによって符号化処理の演算量が増加する可能性があるため実装の工夫が必要となる。

イントラ予測およびインター予測を行う単位である予測ブロックは基本的に CU と共通であり、また、予測差分信号の変換・量子化処理を行う TU(Transform Unit)も特殊なモードを除いて CU と共通となっている。なお、CTU、CU、TU を構成する輝度・色差成分信号の各ブロックをそれぞれ CTB(Coding Tree Block)、CB(Coding Block)、TB(Transform Block)と呼ぶ。

イントラ予測

AVC や HEVC と同様のコンセプトで、予測ブロック毎に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って予測差分信号を変換符号化する。HEVC では 35 種類の予測モードであったのに対し、VVC では 67 種類に拡張され、さらに長方形の予測ブロックでは短辺と長辺の比率に応じて予測方向の割り当てをシフトする Wide angle 予測が導入されている。一部の予測モードに対しては、通常のイントラ予測によって生成された予測画像に対して、さらに画素位置毎に選択される周辺の予測参照画素を用いて予測画像の更新を行う PDPC (Position Dependent intra Prediction Combination)が適用される。また、従来の隣接 1 画素列のみではなく、隣接 3 画素列から任意の 1 画素列を選択して予測参照画素として使用する MRL (Multiple Reference Line)予測も可能となっている。

色差成分の予測に関しては、従来の方式に加えて、輝度成分の復号済み画素から線形予測を用いて色差成分の予測画像を生成する CCLM (Cross-component Linear Model)予測が新たに採用されている。これにより、輝度信号と色差信号の相関が高いブロックにおいてより効率の良い予測が可能となっている。

また、新しいアプローチによるイントラ予測モードとして、MIP (Matrix-based Intra Prediction)と ISP (Intra Sub-partitions)とがある。MIP は、予め学習によって得られた重み行列を用いて、予測ブロックの左および上に隣接する予測参照画素を入力として掛け合わせることで、処理対象ブロックの予測画像を導出する方法である。規格で定義された重み行列に適した画素構成を持つブロックがあった場合に、従来の方式では生成できない特徴を持った、より適切な予測画像を生成できる可能性がある。また、ISP は、予測ブロックを縦方向もしくは横方向に 2 つもしくは 4 つの同じサイズのサブブロックに分割し、サブブロック毎に同一モードのイントラ予測、予測差分信号の変換処理、復号画像の再構成を繰り返して行う方法である。処理対象画素と予測参照画素との距離を短くすることでより効率の良い予測が可能となる。

インター予測

インター予測も、AVC や HEVC と同様のコンセプトで、予測ブロック毎にフレームメモリに格納されている複数枚の参照画像を用いて動き補償を行うことで予測画像を生成して予測差分信号を変換符号化する。動き補償は、HEVC 同様に輝度成分は 8 タップフィルタ、色差成分は 4 タップフィルタを採用しているが、動きベクトルの精度は HEVC が輝度、色差それぞれ 1/4、1/8 画素精度であったのに対して、VVC では輝度、色差それぞれ 1/16、1/32 画素精度に高精度化されている。しかし、予測ブロック単位での動き補償におけるブロックサイズおよび予測方向に関する制約は HEVC 同様であり、復号処理におけるワーストケースでのメモリバンド幅は HEVC 相当となっている。

インター予測のモードに関しては、HEVC 同様に、近傍の復号済みブロックから導出された予測動きベクトルに対して差分動きベクトルを符号化する AMVP モードと、差分動きベクトルを符号化せずに復号側で動きベクトルを特定するマージモードとが基本となっているが、数多くの機能拡張が行われている。

まず、AMVP モードおよびマージモードのそれぞれに対して、1 つの予測ブロックを複数のサブブロックに分解し、サブブロック単位で動き補償を行うサブブロックモードが新たに導入されている。サブブロックモードはアフィン予測と SbTMVP (Subblock-based Temporal Motion Vector Prediction)から構成されている。アフィン予測は、予測ブロックの左上、右上、左下の位置に該当する 3 つの予測動きベクトル

からアフィン変換式を用いてサブブロック毎の動きベクトルに展開して動き補償を行う方法である。これにより回転・拡大・縮小を伴う被写体に対しても少ない符号量で適切な予測が可能となる。また、SbTMVP は、過去の復号済みピクチャで用いられた動きベクトルを時間予測動きベクトルとして参照する際に、8x8 画素のサブブロックに分割して予測動きベクトルを参照する方法である。

AMVP モードでは、差分動きベクトルの精度を 1/4、1/2、1/1、4/1 等の複数の画素精度から選択可能とする AMVR (Adaptive Motion Vector Resolution)や、双方向予測において片方の差分動きベクトルのみを符号化し、もう一方の差分動きベクトルを点対称になるように導出する SMVD (Symmetric Motion Vector Difference)などの機能が追加されている。

マージモードでは、通常のマージモードに加えて、MMVD (Merge Mode with Motion Vector Difference)、CIIP (Combined Inter and Intra Prediction)、GPM (Geometric Partitioning Mode)が追加されている。MMVD は、通常のマージモードで導出された予測動きベクトルに対して、上下左右のいずれかの方向と中心からの離散的な距離の情報を示すインデックス値を用いて、値をシフトさせるモードである。CIIP は、通常のマージモードで生成されたインター予測画像と、平面予測モードを用いて生成されたイントラ予測画像とを重み付け平均によって合成するモードであり、1つの予測ブロック内にイントラ予測に適した領域とインター予測に適した領域とが混在している場合などに有用なモードとなっている。GPM は、矩形の予測ブロックを指定された位置・角度で 2 分割し、2つの台形もしくは三角形の領域として動き補償を行うモードである。2つの領域の境界線の歪みが発生しないように、互いにオーバーラップして重み付け平均化される。なお、メモリバンド幅の増加を回避するために片方向予測のみに限定されている。また各モードにおいて、残差係数を全てゼロとするスキップモードが使用可能である。

さらに、導出された動きベクトルおよび予測画像を復号側で補正して予測精度の向上を図るツールとして、DMVR (Decoder side Motion Vector Refinement)と BDOF (Bi-directional Optical Flow)がある。DMVR は、双方向予測で指定された 2 枚の参照ブロック同士で、5x5 画素範囲でシフトさせて最も SAD 値が小さくなる位置を探索し、得られたシフト値を用いて動きベクトルを補正するツールである。BDOF は、双方向予測で生成された 2 枚の予測画像を用いて、画素毎の輝度値の時間変異と空間勾配値からオプティカルフローを導出し、処理対象ブロックにおける補正值に変換して予測画像を補正するツールである。このように局所的な被写体の変化を復号側で予測して補正することで、ストリームに信号を追加することなく予測精度の向上を実現する仕組みが導入されている。

また、HEVC 同様に、フェードシーンでの動き予測精度を向上する仕組みとして、スライスヘッダに符号化されたパラメータを用いて予測画像の重み付け補正を行う WP (Weighted Prediction)がサポートされている。さらに CU 単位で 5 種類の重みを切り替えながら双方向予測の重み付け補正を行う BCW (Bi-prediction with CU-level Weight)が追加されている。

変換・量子化

VVC では、予測差分信号に対して水平・垂直方向に分離して適用される一次変換に加え、一次変換後の係数に対して非分離で適用される二次変換も採用されている。また、HEVC と同様にブロック単位で変換処理をスキップする TS (Transform Skip)もサポートしている。

一次変換は、HEVC が最大 32x32 画素ブロックまでだったのに対し、VVC では最大 64x64 画素ブロックにまで拡張されている。ただし行列演算による演算回数の増加を防ぐために、大きなブロックサイズでは高周波数域成分の係数値を強制的にゼロとする仕組みが導入されている。また、DCT2、DST7、DCT8 の 3 つの変換基底から任意のものを選択可能とする MTS (Multiple Transform Selection)、1 つの変換ブロックを縦・横方向に 2 つもしくは 4 つのサブブロックに分割し 1 つのサブブロックのみを選択して変換処理を行う SBT (Subblock Transform)などのツールが採用されている。

二次変換は、一次変換後の係数に対して非分離で適用される変換処理であり、LFNST (Low Frequency Non-Separable Transform)と呼ばれる。一次変換後の係数に残存している係数分布の偏りを、

再変換を行うことでさらに低周波数域に集中させて情報量を圧縮させる効果がある。LFNST では、演算量の増加を回避するため、一次変換後の低周波数域の最大 48 係数までを変換対象とし、再変換後の係数も最大 16 係数までに限定して符号化する仕様となっている。

量子化については、HEVC 同様の固定量子化ステップによる量子化器に加え、状態遷移テーブルに従って量子化ステップ位置が 1/2 シフトした量子化器を係数毎に切り替えて適用する DQ (Dependent Quantization) が導入されている。エンコーダで最適な量子化器が選択されるように状態遷移のパスを制御することで非常に大きな符号量削減効果を実現している。また、HEVC 同様に量子化行列 (Quantization Matrix(QM), もしくは Quantization Scaling List) もサポートしている。

なお、HEVC に採用されていた、各画素の画素値をそのままビットストリームに記述する IPCM モード、および変換・量子化処理をスキップする Transquant bypass モードについては、他のツールで代替可能もしくは明確なユースケースがないとして VVC では削除されている。

ループ内フィルタ

ループ内フィルタは、DBF (Deblocking Filter) および SAO (Sample Adaptive Offset) に加え、新たに ALF (Adaptive Loop Filter) および LMCS (Luma Mapping with Chroma Scaling) が採用されている。

DBF は、HEVC 同様に、再構成画像に対してブロック境界歪みを低減するために適用されるフィルタ処理である。HEVC が隣接するブロック境界との間隔に依存せずに常に片側 3 画素に適用されるフィルタであったのに対し、VVC では隣接するブロック境界との間隔および境界の種類に応じて、片側 1 画素から 7 画素の範囲で適用される画素数が自動的に切り替わる仕様となっている。これにより、HEVC では適用除外となっていた 4 画素幅のブロック境界にも DBF が適用可能となり、それぞれのブロックサイズに対して最適なフィルタ処理が可能となっている。また、SAO は基本的に HEVC と共通であり、Band Offset と Edge Offset の 2 種類から構成され、DBF 処理後の画像に対してリングング歪みやランダムな離散ノイズの除去を行う。

ALF は、SAO 処理後の画像に対して適用するフィルタであり、輝度成分は 7x7 画素のダイヤモンド形状、色差成分は 5x5 画素のダイヤモンド形状を持ち、処理対象画素の特徴量に応じて複数のフィルタ係数セットの中から適応的に選択されて用いられる。フィルタ係数はエンコーダで任意のものを設計可能であり、例えばウィナーフィルタを用いることで SN 値を大幅に改善することが可能となる。さらに、輝度成分にフィルタをかけて導出した値をオフセット値とし、対応する色差成分の画素値を補正する CC-ALF (Cross Component Adaptive Loop Filter) と呼ばれるツールも採用されている。

なお、CTU をラスタスキャン順で処理する際に必要となる上側隣接境界画素のラインメモリに関しては、DBF も ALF も 4 画素列の範囲のみで処理できるような仕組みが導入されており、SAO を含めても HEVC と同様の 5 画素列 (実装方法によっては 4 画素列) のラインメモリで実装可能となっている。

LMCS は、入力画像の輝度成分の画素値のステップ幅を情報の重要度に応じて偏りを持たせる変換処理である。例えば、重要な情報が集中している輝度レベルの範囲ではより細かなステップ幅に変換し、そうでない範囲はより粗いステップ幅に変換する。変換・量子化、予測画像生成、再構成画像生成は変換後のドメインで実施し、ループ内フィルタ、動き補償の参照画像、表示用画像は変換前のドメインに戻して処理を適用する。さらに、輝度成分の変換後のステップ幅の偏りに反比例するように、色差成分の予測差分信号に対してスケーリングを行うことで符号量の増加を抑制する仕組みが用いられている。

残差係数・エントロピー符号化

変換・量子化後の残差係数は、HEVC と同様に 1 つの変換ブロックを複数の 4x4 係数ブロックに分割して符号化する。残差係数の符号化方法は、変換処理を適用したブロックに使用される RRC

(Regular Residual Coding)と、変換処理をスキップしたブロックに使用される TSRC (Transform Skip Residual Coding)の 2 種類がある。どちらも HEVC のコンセプトを継承し、小さな係数値は複数のフラグに変換して符号化し、大きな係数値はゴロム・ライス符号を用いて 1 つの信号として符号化するが、周辺係数値との相関関係の違いからそれぞれに最適化された 2 種類の独立した符号化方式となっている。また、色差の 2 つの成分のうち、片方の成分の残差係数のみを符号化し、もう一方の成分の残差係数は他方を変換して流用する JCCR (Joint Coding of Chroma Residual)と呼ばれるツールもある。

エントロピー符号化については、HEVC と同様に全てのプロファイルにおいて、CTU 以下の符号化データに対して CABAC (Context-based Adaptive Binary Arithmetic Coding)が適用される。実装コスト削減のために確率遷移を制御するルックアップテーブルが省略された一方で、確率予測の精度向上のために、短期的変動に反応する確率予測器と長期的変動に反応する確率予測器の 2 つの予測器を用いたモデルが導入されている。なお、CABAC の処理能力を示すスループットは、HEVC よりも処理が複雑になったため若干低下している。CTU より上位のヘッダ層については、従来通り CABAC ではなく固定長符号やゴロム符号等が用いられている。

表 2-1 に、本節で概要を説明した共通符号化ツールについて、HEVC の Main プロファイルと比較したツール比較表を示す。

表 2-1 共通符号化ツール一覧

| 要素技術 | VVC (Main10) | HEVC (Main) |
|--------|---|---|
| ブロック分割 | CU : 4x4~128x128 4 分木、3 分木、2 分木 Dual tree | CU : 8x8~64x64 4 分木 |
| イントラ予測 | 67 モード予測 Wide angle 予測 PDPC (Position Dependent Prediction Combination) MRL (Multiple Reference Line) CCLM (Cross-component Linear Model) MIP (Matrix-based Intra Prediction) ISP (Intra Sub-partitions) | 35 モード予測 |
| インター予測 | MV : 輝度 1/16, 色差 1/32 MC : 輝度 8-tap, 色差 4-tap AMVP モード スキップモード マージモード アフィンモード SbTMVP (Subblock-based Temporal MVP) AMVR (Adaptive MV Resolution) SMVD (Symmetric MV Difference) MMVD (Merge Mode with MV Difference) DMVR (Decoder side MV Refinement) BDOF (Bi-directional Optical Flow) GPM (Geometric Partitioning Mode) CIIP (Combined Inter and Intra Prediction) BCW (Bi-prediction with Cu-level Weight) WP (Weighted Prediction) | MV : 輝度 1/4, 色差 1/8 MC : 輝度 8-tap, 色差 4-tap AMVP モード スキップモード マージモード WP |

| | | |
|---------------|---|--|
| 変換・量子化 | TB : 2x2~64x64, 正方形, 長方形 DCT2, DST7, DCT8 TS (Transform Skip) MTS (Multiple Transform Selection) SBT (Subblock Transform) LFNST (Low Frequency Non-Separable Transform) QM (Quantization Matrix) DQ (Dependent Quantization) | TB : 4x4~32x32, 正方形 DCT2, DST7 TS IPCM Transquant bypass QM |
| ループ内 フィルタ | DBF SAO (Sample Adaptive Offset) ALF (Adaptive Loop Filter) CC-ALF (Cross Component ALF) LMCS (Luma Mapping with Chroma Scaling) | DBF SAO |
| エントロピー 符号化 | CABAC RRC (Regular Residual Coding) TSRC (Transform Skip Residual Coding) JCCR (Joint Coding of Chroma Residual) | CABAC RRC |

2.2. スクリーンコンテンツ向け符号化ツール

パソコンのデスクトップ画面やゲームやアニメの CG 映像などのスクリーンコンテンツ向けの符号化ツールは、HEVC では第 2 版の SCC 拡張プロファイルとして導入されていたが、VVC では第 1 版の各プロファイルに採用されている。表 2-2 に、スクリーンコンテンツ向け符号化ツールの一覧を記載する。

イントラ予測に関するツールとしては、同じピクチャの復号済み周辺領域から予測ブロック単位でコピーして予測画像を取得する IBC (Intra Block Copy)、水平もしくは垂直方向のイントラ予測を行った予測差分画像に対し、変換処理を行うことなく隣接画素の差分値を直接符号化する BDPCM (Block Differential Pulse Coded Modulation)、表現可能な画素値の中から幾つかの代表値のみをパレットとして持たせ、その中から画素毎に 1 つを選択して復号画像とする Palette mode がある。いずれも、HEVC の SCC 拡張プロファイルで規定されていたツールを展開したものであるが、VVC の処理方式に合わせて処理の簡略化および最適化が行われている。

変換に関するツールとしては、HEVC の SCC 拡張プロファイルと同様に、4:4:4 フォーマットにおいて予測差分信号を RGB カラースペースから YCgCo カラースペースに変換して処理を行う ACT (Adaptive Color Transform) がある。また、共通符号化ツールとして説明した TS (Transform Skip) もスクリーンコンテンツ向けとして有効なツールである。

なお、Palette と ACT については Main 10 4:4:4 のプロファイルでのみ使用可能なツールとなっているが、それ以外のツールは全てのプロファイルにおいて他のツールと区別なく使用可能である。また、ここでは ACT をスクリーンコンテンツ向け符号化ツールとして説明したが、YCgCo カラースペースに適した画像であれば通常コンテンツにおいても使用されている。

表 2-2 スクリーンコンテンツ向け符号化ツール一覧

| 要素技術 | ツール名 | 備考 |
|--------|---|------------------|
| イントラ予測 | IBC (Intra Block Copy) | |
| | BDPCM (Block Differential Pulse Coded Modulation) | |
| | Palette | Main10 4:4:4 系のみ |

| | | |
|----|--------------------------------|------------------|
| 変換 | ACT (Adaptive Color Transform) | Main10 4:4:4 系のみ |
| | TS (Transform Skip) | 共通符号化ツール |

2.3. HDR/WCG コンテンツ向けの符号化特徴

VVC では、従来のダイナミックレンジ(SDR: Standard Dynamic Range)の映像だけでなく、Rec. ITU-R BT.2100 に規定される HLG 方式、PQ 方式などの高ダイナミックレンジ(HDR: High Dynamic Range)・高色域(WCG: Wide Color Gamut)の映像に対する符号化にも適したツールが採用されている。また、HDR 方式の映像(UHD の映像では SDR も同様)では、色差のサンプリング位置(図 2-2)として、type-2 のみが規定されていることから、色差サンプリング位置に応じた符号化ツールの制御が導入されている。HDR/WCG コンテンツ向けの符号化ツールの一覧を表 2-3 に示す。

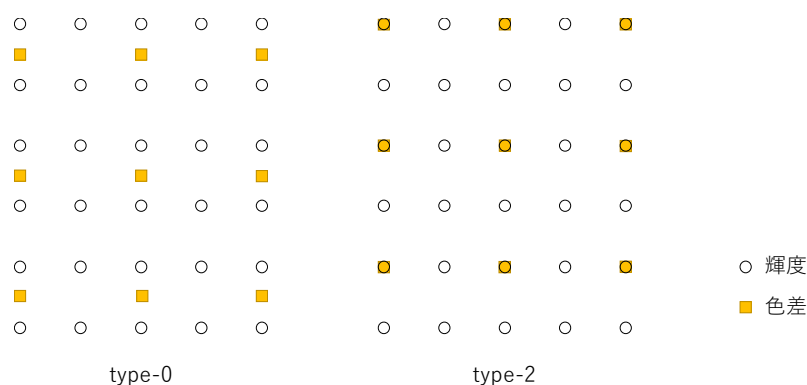


図 2-2 4:2:0 における色差のサンプリング位置

表 2-3 HDR/WCG 向け符号化ツール一覧

| 要素技術 | ツール名 |
|--------------|---|
| イントラ予測 | CCLM (Cross-component Linear Model) |
| 量子化 | Chroma QP mapping |
| ループ内 フィルタ | LMCS (Luma Mapping with Chroma Scaling) |
| | LADF (Luma Adaptive Deblocking Filter) |

CCLM は輝度と色差の成分間相関を用いた予測であることから、輝度画素と色差画素との位置関係が重要となる。CCLM では、ビットストリームに格納された符号化対象映像の色差サンプリング位置に関する情報に基づいて輝度信号をサブサンプリングし、色差信号の予測に用いるよう規定されている。

HDR/WCG コンテンツでは、従来の SDR 映像と比較して、輝度信号のヒストグラムに特に偏りが大きくなる傾向が強いため、ダイナミックレンジを符号化ループ内で変換する LMCS による客観・主観画質の向上が大きい。また、HEVC 以前では、色差の量子化パラメータは輝度の量子化パラメータをもとに、固定のテーブルにより変換されるよう規定されていたが、HDR における輝度のダイナミックレンジ拡大、WCG における色の表現範囲の拡大にともなって、コンテンツに応じた柔軟な量子化パラメータの変換テーブルが利用可能となっている。ループ内フィルタにおいても、輝度のヒストグラムの偏りに着目して、DBF のパラメータを制御する LADF (Luma Adaptive Deblocking Filter)が導入されている。HDR における輝度の高いハイライトの領域と輝度の低い暗い領域とでは、画素値レベルでの同程度のノイズであっても主観画質に与える影響が異なる。そこで、LADF により輝度信号に応じて DBF のフィルタ強

度を調整することにより主観画質を向上させることが可能となる。なお、SDR 映像においてもコンテンツごとのヒストグラムの偏りがあることから、LMCS および LADF は SDR 映像に対しても有効であり、主観画質の改善があることが確認されている。

2.4. ハイレベル機能

VVC では、動画像符号化の基本となるツールに加えて、多様なユースケースに即した様々なハイレベル機能を実現するための特殊なツールが導入されている。表 2-4 に、想定されている主なハイレベル機能と関連するツールの一覧を記載する。

360° 画像符号化機能は、360° を撮影した画像を 1 枚のピクチャにマッピングして通常の動画像として符号化する機能である。ピクチャの一部分のみを独立して取り出して復号および表示することを可能とする Subpicture、複数の画像のつなぎ目に不適切にループ内フィルタが適用されてしまうことを回避する VB (Virtual Boundary)、ピクチャ端を周回して動き補償を行うことを可能とする Wrap around MC といったツールで実現されている。

空間/SN/多視点スケーラビリティは、HEVC では第 2 版の Multilayer 拡張規格によって対応していたが、VVC では第 1 版で規格化されたツールである Multilayer によって実現可能となっている。具体的には、1 つのアクセスユニット内に複数のレイヤに属するピクチャを符号化し、特定のレイヤのピクチャのみを抽出して復号および表示することを可能としている。レイヤ間のインター予測参照については、HEVC の Multilayer 拡張規格と同様に、同一表示時刻(即ち同一アクセスユニット内)のピクチャのみ参照可能となっている。その際に、異なる解像度のピクチャを参照可能とするため、動き補償と同時に解像度変換を行って予測画像を生成する参照ピクチャリサンプリング(Reference Picture Resampling, RPR)が使用される。なお、Multilayer に関しては基本プロファイルとは異なる Multilayer 専用のプロファイルのみでサポートされている。

時間階層スケーラビリティは、HEVC と同様に、空間/SN/多視点スケーラビリティで使用される Multilayer とは異なり、時間階層構造を規定するサブレイヤによって実現され、全てのプロファイルでサポートされている。

インターレース符号化は、HEVC 同様に、表示制御以外の処理は全てプログレッシブ素材の符号化と共通となっており、1 フレームを 1 ピクチャで符号化するか、もしくは 1 フィールドを 1 ピクチャで符号化するかのどちらかを SPS で選択可能となっている。なお、HEVC では、フィールド構造を用いた場合にリーディングピクチャの使用制限があったが、VVC ではそれが緩和されている。

シーケンス内解像度変更は、同一レイヤ・シーケンス内でピクチャ毎に符号化解像度を変更する機能であり、RPR によって実現される。ストリーミング等の帯域変動が大きい状況で本機能を適用することで画質破綻の回避が可能となる。RPR は他に、符号化解像度を固定したまま、参照ピクチャ内の注目部分領域の参照にも使用できる。ズームシーンにて、被写体の拡大・縮小に対応した動き補償が可能になり、符号化効率が向上する。なお、ピクチャ内の部分領域は Scaling window として PPS で指定される。

低遅延符号化は、ランダムアクセスポイント(ビットストリーム途中からの復号開始ポイント)において、符号量の多いイントラピクチャを使用せず、各インターピクチャにイントラブロックの位置をずらしつつ挿入することで符号量を分散させ、必要なバッファ容量を削減して低遅延を実現する機能であり、GDR (Gradual Decoding Refresh)と呼ばれる。従来規格でも符号化時に様々な制約を用いることで部分的に実現可能であったが、VVC では、GDR ピクチャであることを示す GDR NAL ユニットの、およびリフレッシュ領域と非リフレッシュ領域の境界に不適切にループ内フィルタが適用されてしまうことを回避する VB (Virtual Boundary)によって、より直接的にこの機能がサポートされている。

並列符号化を実現するツールとしては、HEVC と同様に、Slice や Tile を用いたピクチャの分割、お

よび、CTUライン単位で並列してCABAC処理を可能とするWPP (Wavefront Parallel Processing)がサポートされている。なお、HEVCで導入されていたDependent sliceはVVCでは廃止されている。

ロスレス符号化は、HEVCではIPCMモードやTransquant bypassモードによって実現されていたが、VVCではそれらのツールが廃止され、TS (Transform Skip)と量子化値制御およびループ内フィルタ等をブロック単位で明示的にオフにすることで実現される。ただし、符号化効率の向上を図って、TSを用いたブロックにおいて2種類の残差係数の符号化方法(RRCとTSRC)のうち最適な一方を選択して使用することが可能となっている。

表 2-4 ハイレベル機能一覧

| 機能 | 関連ツール | 備考 |
|-----------------------|---|-------------------------------|
| 360° 画像符号化 | Subpicture VB (Virtual Boundary) Wrap around MC | |
| 空間/SN/多視点 スケーラビリティ | Multilayer RPR (Reference Picture Resampling) | Multilayer Profileのみ |
| 時間階層 スケーラビリティ | Syntaxのみで対応 | HEVCと同様 |
| インターレース符号化 | Syntaxのみで対応 | HEVCと同様 リーディングピクチャも使用可 |
| シーケンス内 解像度変更 | Scaling window RPR (Reference Picture Resampling) | |
| 低遅延符号化 | GDR NALU VB (Virtual Boundary) | |
| 並列符号化 | Slice Tile WPP (Wavefront Parallel Processing) | HEVCと同様 Dependent Sliceは廃止 |
| ロスレス符号化 | TS (Transform Skip) RRC/TSRC切替 | IPCM、Transquant bypassは廃止 |

3. 符号化データ構造

3.1. ビットストリーム構造

VVCのビットストリーム構造は、基本的にHEVCを踏襲し、Multilayerに対応する形に拡張されている。

3.1.1. ビットストリーム、シーケンス、アクセスユニット、ピクチャユニット

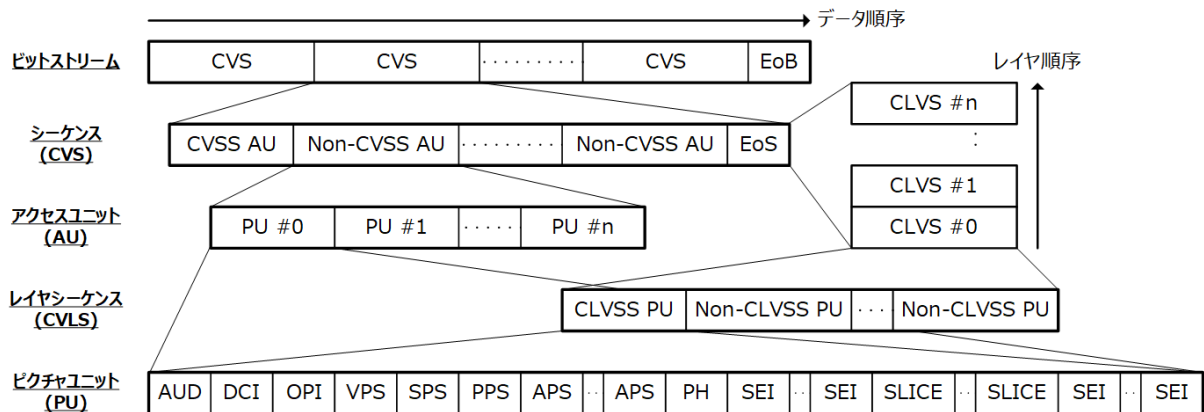


図 3-1 VVC ビットストリームの構造

ビットストリームは複数のシーケンス(Coded Video Sequence, CVS)と End of Bitstream NAL unit (EoB)から構成される。

CVS は複数のアクセスユニット(Access Unit, AU)と End of Sequence NAL unit (EoS)から構成される。CVS 先頭の AU は Coded Video Sequence Start (CVSS) AU と呼ばれる。CVS をレイヤ(エラー! 参照元が見つかりません。節)毎に分割したものを Coded Layer Video Sequence (CLVS)と呼ぶ。

AU は、同一出力時刻の、一つ(非 Multilayer ストリーム時)もしくは複数(Multilayer ストリーム時)のピクチャユニット(Picture Unit, PU)から構成される。Multilayer ストリーム時の場合、レイヤ番号が小さいレイヤ(下位レイヤ)の PU から順に格納される。

CLVS は、同一レイヤの PU から構成される。CLVS 先頭の PU は、Coded Layer Video Sequence Start (CLVSS) PU と呼ぶ。CLVSS PU は、ピクチャ種別が IRAP もしくは GDR(エラー! 参照元が見つかりません。節)となる PU に限定される。

PU は、複数の NAL ユニット(Network Abstraction Layer Unit, NALU)(エラー! 参照元が見つかりません。節)から構成される。

3.1.2. レイヤ

VVC 規格は、HEVC 拡張規格と同じく、複数レイヤ(空間解像度、品質、視点)からなる Multilayer 機能をサポートしている。なお、階層 B ピクチャ参照構造により実現される時間方向の階層符号化では、各階層をサブレイヤと呼び、Multilayer 機能とは別に扱う。

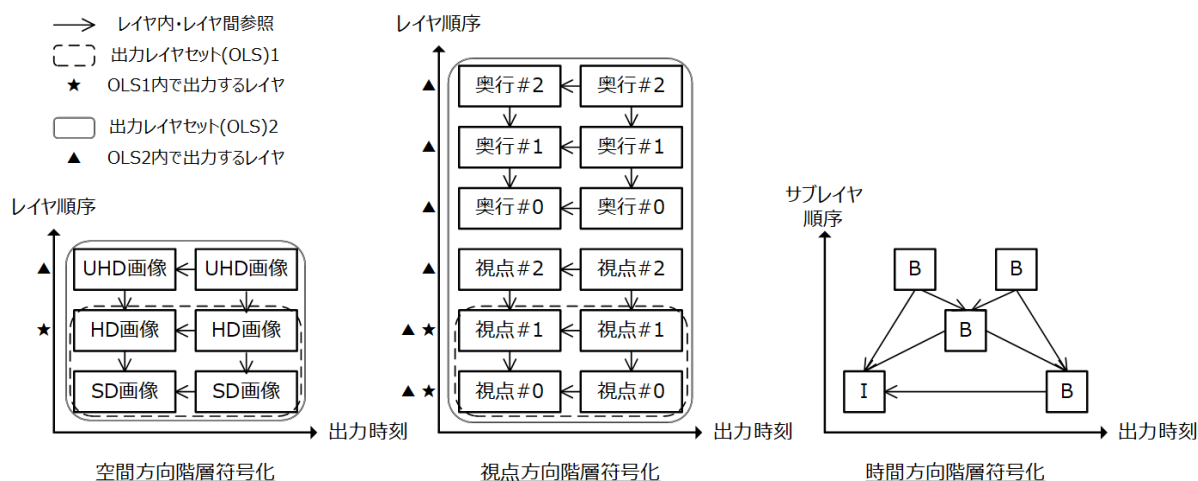


図 3-2 階層符号化におけるレイヤ構造及び OLS、サブレイヤ構造

Multilayer ストリームの場合、あるレイヤを復号出力する際に必要なレイヤの集合を **Output Layer Set (OLS)**と呼ぶ。OLS は複数設定可能であり、また、各 OLS は複数の出力レイヤを持つことができる。これにより、視点方向階層符号化において仮想視点画像の合成に必要な複数視点画像と奥行画像の出力を選択可能になる。

3.1.3. NAL ユニット

NAL ユニット(NALU)は、ビットストリームの基本アクセス単位である。NALU には、画素値の符号化データである **Video Coding Layer (VCL) NALU** と、各種ヘッダ情報である **Non-VCL NALU** の二種類がある。VVC で規定される NALU をエラー! 参照元が見つかりません。にまとめる。VCL NALU について、HEVC では参照ピクチャか非参照ピクチャの区別を名称レベルで行っていたが、VVC ではピクチャヘッダ内のフラグ `ph_non_ref_pic_flag` で区別する。

表 3-1 NAL ユニットの種類(Reserved を除く)

| 名称(nal_unit_type) | 種別 | 説明 |
|-------------------|-----|---|
| TRAIL_NUT | VCL | Trailing Picture (TP)を構成するスライス。 TPは、IRAPピクチャよりも復号順序及び出力順序が後となるピクチャ。 |
| STSA_NUT | VCL | Stepwise Temporal Sub-layer Access (STSA)ピクチャを構成するスライス。 STSA ピクチャは、時間方向階層符号化ビットストリームにおける、復号サブレイヤの切り替えが可能なピクチャである。HEVC にあった Temporal Sub-layer Access (TSA)ピクチャは無くなり、STSA ピクチャのみが規定される。 |
| RADL_NUT | VCL | Random Access Decodable Leading (RADL)ピクチャを構成するスライス。 RADL ピクチャは、IRAP ピクチャよりも復号順序が後で出力順序が先となるピクチャ Leading Picture(LP)の一つであり、直前の IRAPAU から復号を開始した場合にも正常復号が可能なピクチャである。 |

| | | |
|----------------------------------|---------|---|
| RASL_NUT | VCL | Random Access Skipped Leading (RASL)ピクチャを構成するスライス。 RADL ピクチャは、LP の一つであり、直前の IRAP AU から復号を開始した場合は正常復号できないピクチャである。 |
| IDR_W_RADL IDR_N_LP | VCL | Instantaneous Decoding Refresh (IDR)ピクチャを構成するスライス。 IDR ピクチャは Intra Random Access Point (IRAP)ピクチャの一つであり、LP として RASL ピクチャを持たない。 |
| CRA_NUT | VCL | Clean Random Access (CRA)ピクチャを構成するスライス。 CRA ピクチャは IRAP ピクチャの一つ。HEVC にあった Broken Link Access (BLA)ピクチャ(ビットストリーム編集により CRA ピクチャから名称が変更されるピクチャ)は無くなった。 |
| GDR_NUT | VCL | Gradual Decoding Refresh (GDR)ピクチャを構成するスライス。 GRA ピクチャは IRAP ピクチャと同じく復号開始ピクチャであるが、全ての CU をイントラ符号化する必要はない。 |
| OPI_NUT | Non-VCL | Operating Point Information (OPI) VVC で新たに導入されたヘッダ情報。 ビットストリームに含まれる OLS のインデックスを記述する。 |
| DCI_NUT | Non-VCL | Decoding Capability Information (DCI) VVC で新たに導入されたヘッダ情報。 ビットストリームに含まれる各 OLS のプロファイル・レベル情報を記述する。 |
| VPS_NUT | Non-VCL | Video Parameter Set (VPS) ビットストリームに含まれる OLS 及びサブレイヤの情報を記述する。 |
| SPS_NUT | Non-VCL | Sequence Parameter Set (SPS) 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) 複数ピクチャで共通なパラメータを記述する。 |
| PREFIX_APS_NUT SUFFIX_APS_NUT | Non-VCL | Adaptation Parameter Set (APS) VVC で新たに導入されたヘッダ情報。 ALF パラメータ、LMCS パラメータ、スケーリングリストパラメータを記述する。参照するスライスの前、もしくは後に出現する。 |
| PH_NUT | Non-VCL | Picture Header (PH) VVC で新たに導入されたヘッダ情報。 ピクチャ内の各スライスで共通なパラメータを記述する。 |
| AUD_NUT | Non-VCL | AU Delimiter (AUD) AU の境界を示す。 |
| EOS_NUT | Non-VCL | End of Sequence (EoS) シーケンスの境界を示す。 |

| | | |
|----------------------------------|---------|--|
| EOB_NUT | Non-VCL | End of Bitstream (EoB) ビットストリームの境界を示す。 |
| 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) HEVC と同じく、スライスの前、もしくは後に出現する。 |

3.1.4. ピクチャ種別

ピクチャは HEVC と同じく、復号処理の観点で復号開始ピクチャ、リーディング(Leading)ピクチャ、トレーリング(Trailing)ピクチャに分類される。

復号開始ピクチャは、従来規格の Intra Random Access Point(IRAP)ピクチャ(いわゆるイントラピクチャ)と、新たに追加された Gradual Decoding Refresh(GDR)ピクチャの二種類がある。GDR は超低遅延動作時に用いられる、いわゆるイントラスライスリフレッシュであり、GDR ピクチャはリフレッシュ周期先頭ピクチャに相当する。従来規格では GDR 動作時でもビットストリーム先頭ピクチャはイントラピクチャである必要があったが、この制約が撤廃された。GDR ピクチャでは、PH で GDR ピクチャであることを示すフラグ(ph_gdr_pic_flag)を立て、GDR ピクチャ間隔(ph_recovery_poc_cnt)を送る。

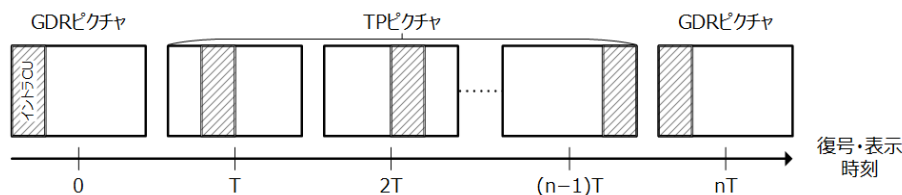


図 3-3 GDR ピクチャ

HEVC と同じく、VVC は主にプログレッシブ映像を対象とするが、インターレース映像も符号化可能である。フィールド構造に対応した符号化ツールは無いが、適切に各フィールドを出力表示するためのパラメータが定義されている。なお、HEVC では、インターレース映像の場合(sps_field_seq_flag = 1)には、LP ピクチャを使用する場合に、トップフィールドを IRAP ピクチャ、対応するボトムフィールドを TP ピクチャにできない、という制約があったが、VVC では緩和されている。

3.1.5. 参照ピクチャリサンプリング

VVC は、処理ピクチャと参照ピクチャとの解像度が異なる場合にも適切なインター予測を可能にする参照ピクチャリサンプリング(Reference Picture Resampling, RPR)の機能を備えている。HEVC でも、空間方向階層符号化において RPR に相当する機能を有していた。VVC の RPR は、スケーリング単位を参照ピクチャ全体ではなく、インターCU の参照範囲とすることで、以下の符号化形態が可能となった。

空間方向階層符号化: レイヤ毎に異なった空間解像度で符号化する。なお、RPR は色差形式や画素値精度のスケーリングには対応していないため、レイヤ間参照は、同一の色差形式及び画素値精度の場合に限定される。

シーケンス内解像度変更: データ量削減のためにピクチャの符号化解像度を落とす。従来規格には画素予測誤差の空間解像度を落とす機能があったが、VVC ではピクチャ自体の空間解像度を落として符号化する。RPR により、空間解像度の変更がピクチャ単位となり、また縮小率をきめ細かく調整

可能となった。ストリーミング等の帯域変動が大きい状況での画質破綻の回避に利用可能である。

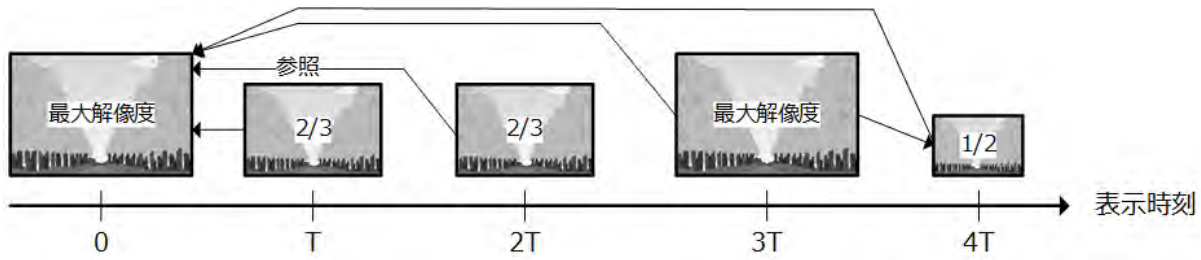


図 3-4 シーケンス内解像度変更

スケーリングウィンドウ: ピクチャの符号化解像度を一定にしたまま、ピクチャのインター予測参照範囲を変更する。本範囲はスケーリングウィンドウ(Scaling Window)と呼ばれ、画面の上下左右オフセットによって指定する。ズーミングシーンの動き予測効率向上が可能になる。

図 3-5 スケーリングウィンドウによるズーミングシーンへの対応

3.2. ピクチャ分割

3.2.1. タイル、スライス、サブピクチャ

ピクチャは、HEVC と同様に、一ないし複数のタイル(Tile)に分割される。タイルは、基本ブロックである Codec Tree Unit (CTU)を単位とする矩形領域である。タイル境界を跨ぐ参照が制約されるため、ピクチャの空間分割処理が可能になる。

スライスは、従来規格のスライスとは若干異なり、構成単位が CTU から、タイル(もしくはタイル内の CTU ライン)に変更されている。これは、ストリーミング時のパケットサイズ均一化(CTU 単位でスライスのデータ量をチェックし、閾値を超えた時点でスライス境界を入れる)の要件はもはや不要と判断されたためである。また HEVC の Dependent Slice は廃止された。更に、360° 映像に対応するため、Raster-Scan Slice モードと Rectangular Slice モードの二モードが導入された。前者は従来と同じく、ラスタスキャン順にスライスを配置するモードである。後者はスライスがカバーする領域が小画面(タイルもしくはタイル内の CTU ラインを単位とする矩形領域)となるモードであり、画面内のスライス順番が緩和される。Rectangular Slice モードの場合、各スライスの順番は画面内ラスタスキャン順でなくてもよい。ただし、スライスの上側及び左側に接するスライスは先に処理されている、という制約が付く。

サブピクチャは、VVC で新しく導入された分割手法である。サブピクチャは一ないし複数のスライス(Rectangular slice モード限定)から構成される。HEVC の Motion Constrained Tile Set (MCTS)と同様

に、360° 映像の符号化に用いられる。動き補償において、サブピクチャの境界をピクチャ境界と見なすことができるため、動きベクトルがサブピクチャ境界を跨ぐ場合にも、他サブピクチャ内の画素を参照することなく正しく復号できる。VVC は、サブピクチャデータが、VVC 準拠のビットストリームとなるようにビットストリームから抽出する処理(Subpicture sub-bitstream extraction process)を、レイヤ・サブレイヤ抽出処理と合わせて規定している。

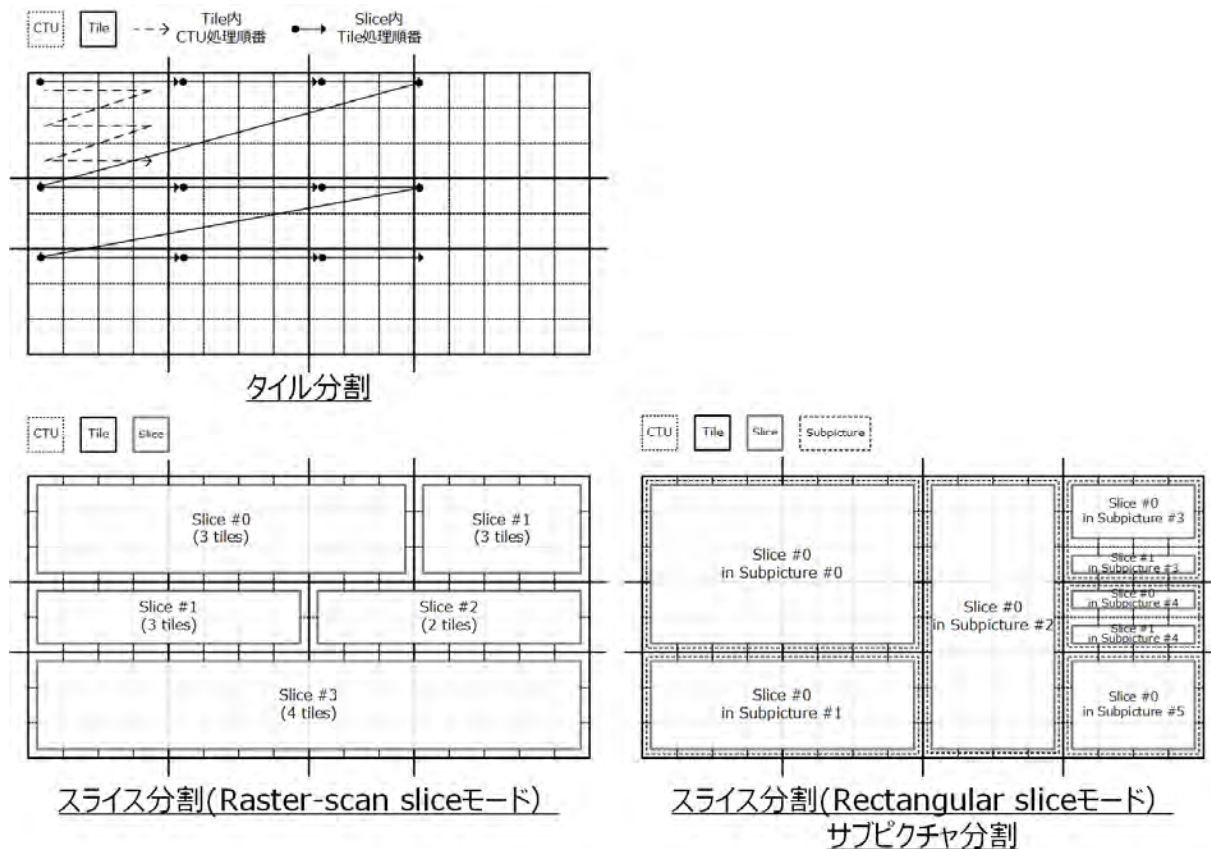


図 3-6 タイル、スライス、サブピクチャによるピクチャ分割

3.2.2. ウェーブフロント並列処理

HEVC と同じウェーブフロント並列処理(Wavefront Parallel Processing, WPP)がサポートされている。HEVC と異なり、タイルと併用可能である。

WPP は、CTU ラインを単位とした並列処理を可能にする。この際、各 CTU ラインの処理の開始をずらして CTU ライン間の参照を可能にし、また各 CTU ライン開始時に CABAC 状態を上隣接 CTU の最終 CABAC 状態に初期化することで、符号化効率の低下を抑えている。

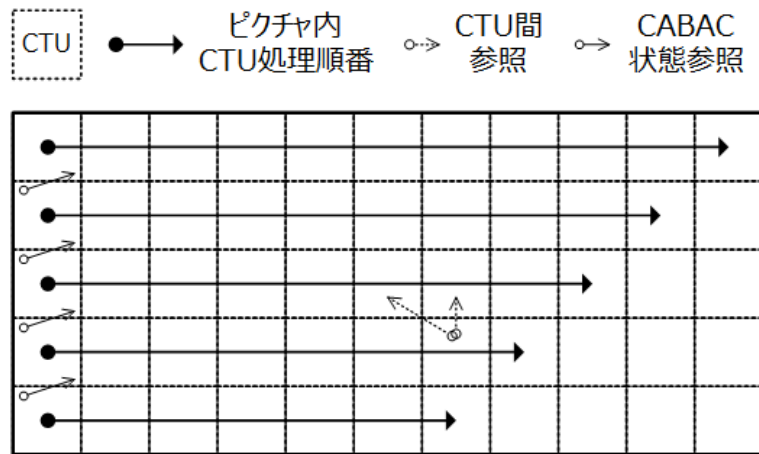


図 3-7 ウェーブフロント並列処理

3.2.3. 仮想境界

360° 映像における各カメラ映像のつなぎ目での画質劣化、GDR 動作時におけるリフレッシュ境界での正常復号を助ける仕組みとして、仮想境界(Virtual Boundary, VB)が導入された。VB を跨ぐループ内フィルタは無効となる。VB はピクチャ縦方向及び横方向にそれぞれ 3 本まで挿入でき、位置の指定は 8 画素単位となる。

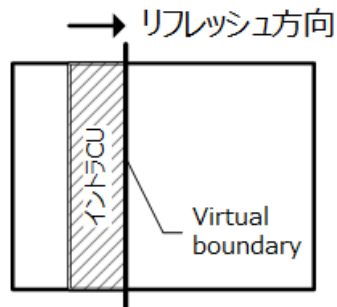


図 3-8 仮想境界

3.3. ブロック分割

3.3.1. CU 分割

VVC では各ピクチャの符号化処理を、HEVC と同様に CTU (Coding Tree Unit)と呼ばれる正方面素ブロックの単位で行う。各々の CTU は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 HEVC では最大 64x64 画素ブロックであったのに対し、VVC では最大 128x128 画素ブロックに拡張されている。

イントラ・インター予測、変換・量子化、エントロピー符号化などの一連の符号化処理は、CTU を再帰的に木構造分割することによって得られる CU (Coding Unit)の単位で行う。各ノードでの分割パターンは、HEVC では 4 分木のみであったのに対し、VVC では、図 3-9 のようにさらに 2 分木および 3 分木 (水平もしくは垂直方向に分割)も使用可能となっている。多様な大きさ・形状の CU に分割することで符号化効率の更なる向上が可能となるが、最適な分割方法を判定するための符号化処理の演算量が増加する可能性があるため実装の工夫が必要となる。なお、VVC ではピクチャサイズは 8 の倍数の任意の値を取ることが可能であるため、CTU のサイズで割り切れない場合にピクチャの右端および下端において長方形の CTU が出現し、シンタックスで指定することなく強制的に長方形の CTU に収まるサイ

ズの CU に分割する制御が行われる。なお、CTU(Coding Tree Unit)、CU(Coding Unit)、TU(Transform Unit)を構成する輝度・色差成分信号の各ブロックをそれぞれ CTB(Coding Tree Block)、CB(Coding Block)、TB(Transform Block)と呼ぶ。

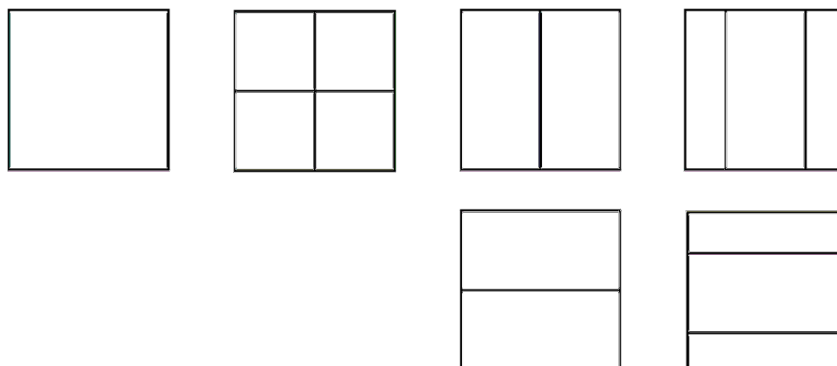


図 3-9 ブロック分割パターン

表 3-2 に各ブロックの取りうるサイズの範囲を示す。HEVC では予測ブロックおよび変換ブロックが CU とは別の分割で指定されていたが、VVC では CU の分割形状のバリエーションが増えたことから、特殊なモードおよび最大・最小サイズを除いて基本的に予測ブロックおよび変換ブロックは CU と共通化されている。

最大サイズは、輝度 CB は 128x128、輝度 TB は 64x64 であり、色差 CB および TB は 4:2:0 フォーマットの場合それぞれ縦横 1/2 となる。

イントラ予測では、ISP モード(4.7 節参照)の場合に、輝度 CB が 2 つもしくは 4 つのサブブロックに分割される可能性があるため、それに伴い輝度 TB の最小サイズが 4x4 に加えて 1x16/16x1/2x8/8x2 となっている。また、SCIPU(3.3.2 節参照)によって色差 CB が 16 画素未満となることが制限されるため、色差 CB および TB の最小サイズが 2x8/8x2/4x4 となっている。なお、イントラ予測処理は TB の単位で行われる。

インター予測では、HEVC と同様に CB の最小サイズが輝度は 4x8/8x4、色差は 2x4/4x2 となるが、TB は SBT(6.1.2 節参照)による分割が行われる可能性があるため、最小サイズが輝度は 4x4、色差は 2x2 となる。なお、インター予測処理は CB の単位で行われる。

表 3-2 ブロックサイズの範囲

| 分類 | | ブロックサイズ | |
|--------|-----------|---------|-----------------------------|
| イントラ予測 | 輝度 | CB | 4x4~128x128 |
| | | TB | 1x16/16x1/2x8/8x2/4x4~64x64 |
| | 色差(4:2:0) | CB | 2x8/8x2/4x4~64x64 |
| | | TB | 2x8/8x2/4x4~32x32 |
| インター予測 | 輝度 | CB | 4x8/8x4~128x128 |
| | | TB | 4x4~64x64 |
| | 色差(4:2:0) | CB | 2x4/4x2~64x64 |
| | | TB | 2x2~32x32 |

VVC では、ハードウェア実装におけるパイプライン構成の 1 ステージの処理単位を、TB の最大サイズである 64×64 (4:2:0 の色差は 32×32) に区切って実装可能とするために、VPDU (Virtual Pipeline Data Unit) と呼ばれる概念が導入されている。具体的には、CU サイズが 64×64 より大きい場合は、CU 内にそれぞれの 64×64 領域 (VPDU) が完全に包含される形状となるように制限が行われ、逆に CU サイズが 64×64 以下の場合は、 64×64 領域 (VPDU) 内にそれぞれの CU が完全に包含される形状となるように制限が行われる。また、同様にインター予測の各処理モジュールを 16×16 単位のパイプライン構成に区切って実装可能とするために、CU サイズが 16×16 より大きい場合でも、MV 導出以外の全ての処理が 16×16 単位に分割して実装しても結果が完全に一致するようになっている。

3.3.2. Dual Tree

CU 分割は、基本的には輝度と色差とで共通の分割が適用されるが、I スライスに限り、輝度と色差とで独立した CU 分割を可能とする Dual Tree (Chroma Separate Tree, CST) が選択可能となっている。なお、Dual Tree を選択した場合は、対象 I スライス内の全ての CTU で Dual Tree による CU 分割が行われる。例えば、輝度成分に比べて色差成分の変化が小さい場合などに、Dual Tree を用いて色差成分のブロックサイズを大きくすることで符号化効率を向上させることが可能となる。

Dual Tree では、 64×64 までは強制的に分割され、 64×64 以下の CU 分割について輝度と色差とで独立したシンタックスを用いてそれぞれ分割が指定される。ブロック分割が異なるため、輝度に依存した色差の処理において特別な制御が必要となっている。まず、輝度ブロックのパラメータを色差ブロックで参照する処理 (イントラ予測モード、QP 等) については、対象色差ブロックの中心位置に対応する輝度ブロックのパラメータを参照する。また、CCLM (4.4 節参照) では、輝度ブロックのイントラ予測処理が完了するのを待ってから色差ブロックのイントラ予測を開始する必要があるため、待ち時間のワーストケースが 32×32 輝度ブロックの処理時間以下となるように、輝度と色差のブロック分割の組合せによって CCLM の使用が制限されている。

また、イントラ予測において、パイプライン処理における周辺参照画素へのアクセス頻度を抑制するために、色差ブロックが極端に小さく分割されないように制限する Local Dual Tree (Smallest Chroma Intra Prediction Unit, SCIPU) と呼ばれる仕組みも導入されている。具体的には、イントラ予測の場合に色差ブロックが $2 \times 2 / 4 \times 2 / 2 \times 4 / 2 \times N$ となることを禁止し、そのようなブロックが発生するような CU 分割が指定された場合は、輝度ブロックのみ分割を行い色差ブロックに関しては分割を途中で停止する。これにより局所的に輝度と色差とでブロック分割が異なることになるため、Dual Tree と同様の制御が行われる。

4. イントラ予測

4.1. 輝度イントラ予測

輝度イントラ予測は符号化ブロックに対して、当該符号化ブロックのイントラ予測モードにしたがって予測信号生成単位である予測ブロック単位に予測処理を実施する。予測ブロックは符号化ブロックまたは**エラー! 参照元が見つかりません**。節記載の符号化ブロックを分割したサブブロックを選択できる。ただし、当該予測ブロックが**エラー! 参照元が見つかりません**。節の MIP を示す場合、MIP による予測信号生成を行う。

図 4-1 は $N \times N$ 画素の予測ブロックと隣接参照画素との関係を示す。生成した隣接参照画素より、図 4-2 に示すイントラ予測モードに従って予測信号を生成する。モード 0 は Planar 予測、モード 1 は DC 予測、残りのモードが Angle 予測を示す。モード総数は、HEVC が 35 モードであったのに対し、VVC では 67 モードに拡張されている。

Planar 予測は図 4-3 に示す予測ブロックの左及び上の参照画素から内挿的に予測する。黒画素は参照画素、グレー画素は予測対象画素を示す。DC 予測は参照画素の平均値を一様に予測ブロック

内にコピーする予測である。Angle 予測は予測方向に従って参照画素をコピーする予測モードであり、HEVC と比較して予測方向の分解能が約 2 倍となっている。

モード-1~-14、67~80 は Wide angle 予測と呼ばれる予測モードであり、非正方の予測ブロックに対してのみ選択可能な、Angle 予測の最大角を超える角度の予測を行う。本予測は、非正方ブロックの短辺側の参照画素を用いる Angle 予測の一部と置き換えることで、符号化に要するモード数の増加なしで実現している。

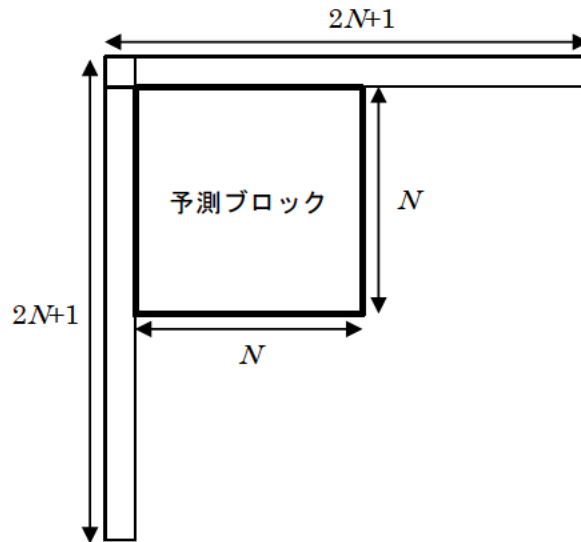


図 4-1 予測ブロックと隣接参照画素

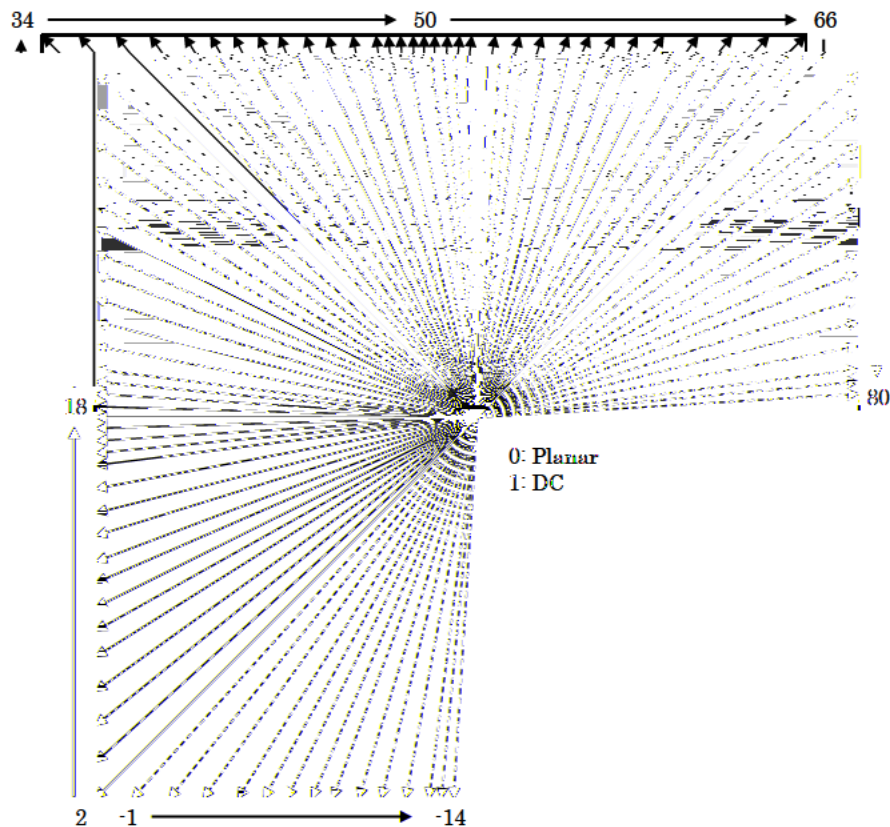


図 4-2 イントラ予測モード

図 4-3 Planar 予測

4.2. 色差イントラ予測

色差信号についても輝度信号と同じく予測ブロックに対して、隣接参照画素を算出し、これを用いた予測処理を実施する。ただし、選択可能な予測モードはモード 0 の Planar 予測、モード 1 の垂直方向予測、モード 2 の水平方向予測、モード 3 の DC 予測、モード 4 の DM 予測である。DM 予測は当該 CU の座標($cbWidth / 2$, $cbHeight / 2$)の位置の輝度信号の予測ブロックと同じ予測を行う予測モードである。ここで、 $cbWidth$ 、 $cbHeight$ は当該 CU の水平方向と垂直方向の画素数をそれぞれ示す。

ただし、当該 CU のシンタックス $cclm_mode_flag$ が 1 を示す場合はエラー! 参照元が見つかりません。節で説明する CCLM を実施する。

4.3. MRL (Multiple Reference Line)

MRL は、輝度イントラ予測において、隣接参照画素を、予測ブロックに隣接するラインだけでなく、1 画素及び 2 画素離れたラインの計 3 つの参照ラインから CU 単位に明示的に選択して予測する方式である。ただし、Planar 予測については常に予測ブロックに隣接するラインとする。また、色差イントラ予測についても常に予測ブロックに隣接するラインに固定となる。

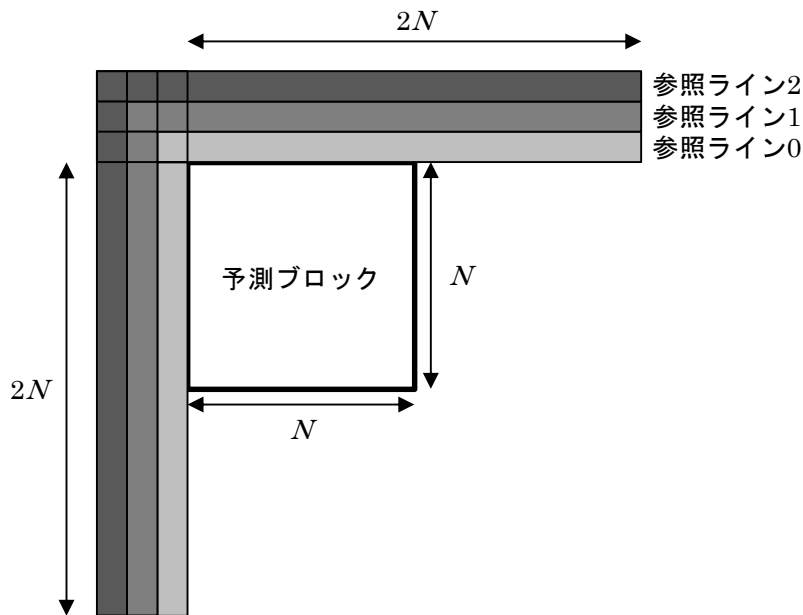


図 4-4 MRL

4.4. CCLM (Cross-component Linear Model)

CCLM 予測は符号化済みの輝度信号から色差信号を下記式にしたがって線形予測するものである。ここで、 pred_c は予測された色差信号、 rec_L' はデブロックフィルタ前の復号済み輝度信号を色差信号の解像度にダウンサンプルしたもの、 α 、 β は予測ブロックの隣接参照画素の輝度信号と色差信号から算出されるパラメータである。

$$\text{pred}_c(i,j) = \alpha \cdot \text{rec}_L'(i,j) + \beta$$

さらに、CCLM 予測は上記 α 、 β を算出する際の隣接参照画素の配置に図 4-5～図 4-7 に示す LM モード、LM-A モード、LM-L モードの 3 つを選択可能としている。なお、グレーで示した輝度信号の隣接参照画素は、色差信号の隣接参照画素に対応するように、色差フォーマットおよび位相に応じたフィルタによってダウンサンプルされる。

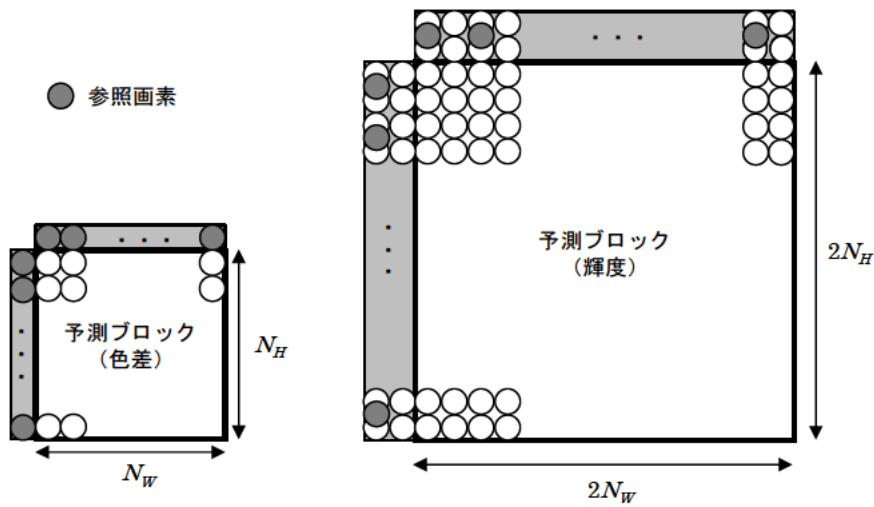


図 4-5 LM モードの参照画素配置

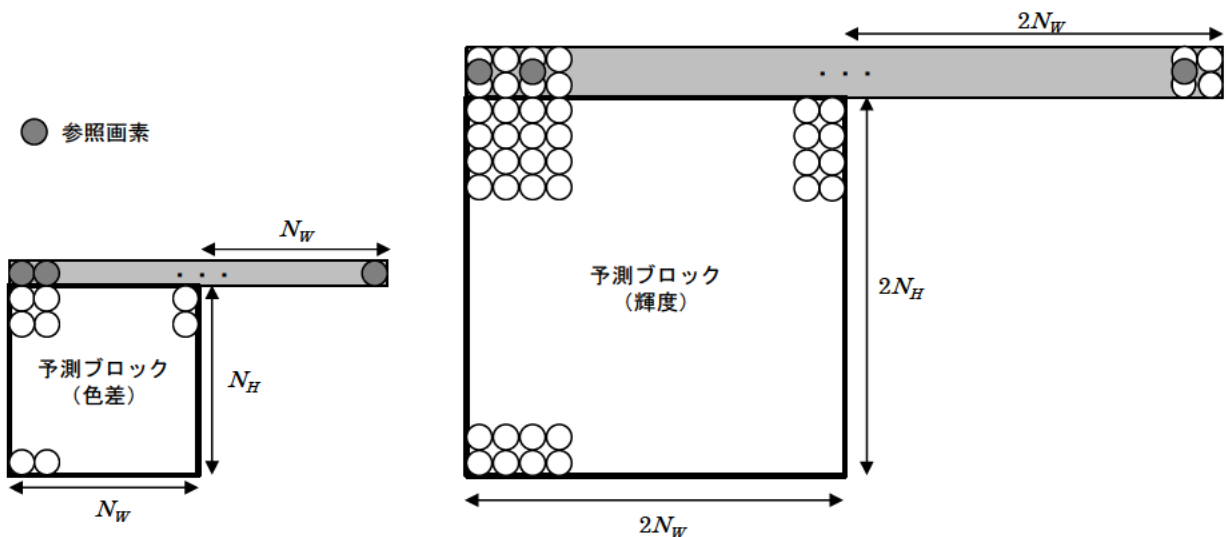


図 4-6 LM-A モードの参照画素配置

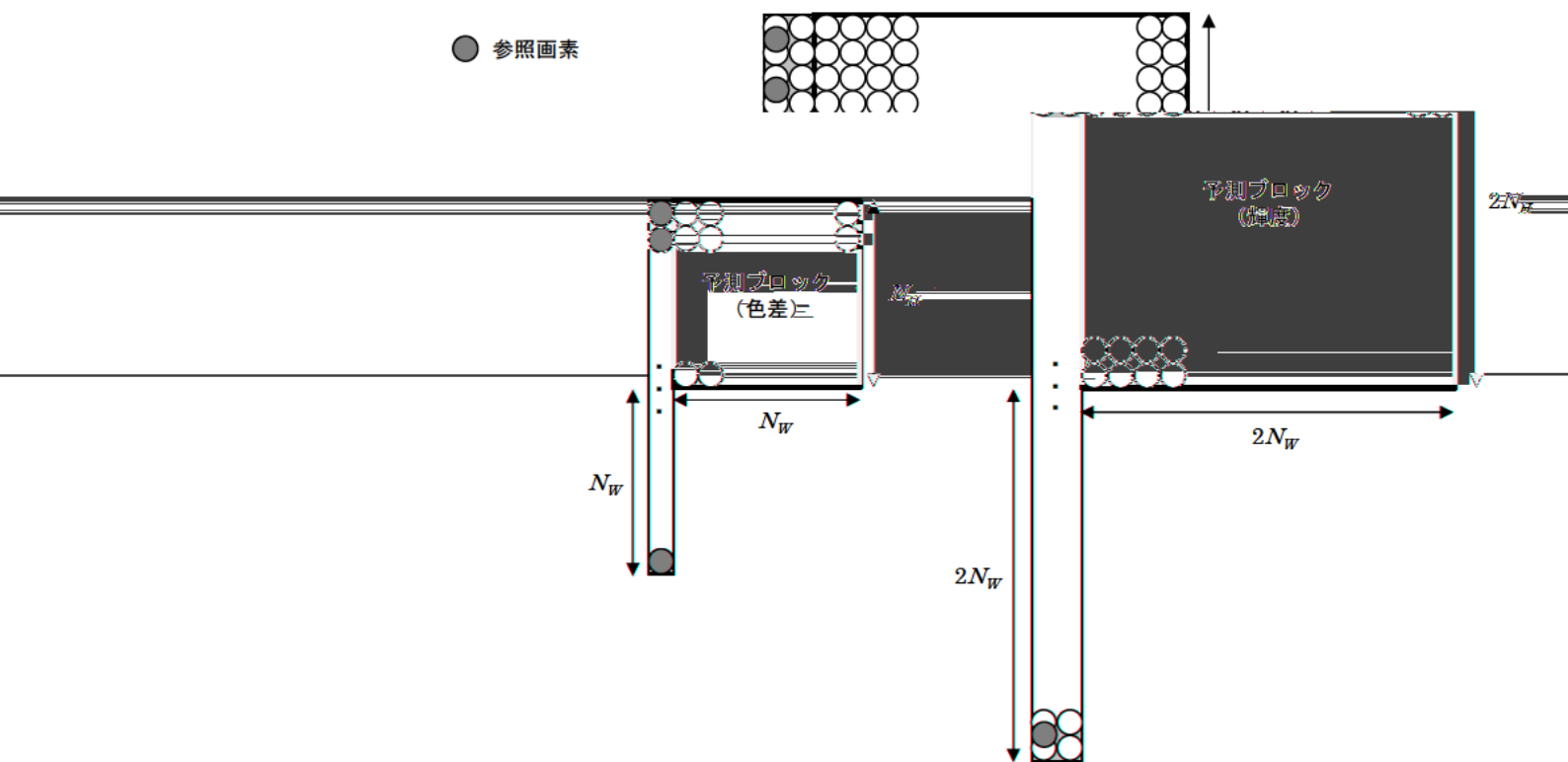


図 4-7 LM-L モードの参照画素配置

4.5. PDPC (Position Dependent intra Prediction Combination)

PDPC (位置依存イントラ予測コンビネーション) は、輝度および色差のイントラ予測後の予測隣接する参照画素を用いて更新する処理であり、0、1、18、50、2~10、および、58~66 の角度において、予測画像の各画素 $\text{pred}(x',y')$ と隣接する参照画素 $R_{i,j}$ を以下の式で線形結合する時、重み付け係数である wL 、 wT 、 wTL は、角度モードと予測画像の位置 (x',y') によって決まる。

$$\text{pred}(x',y') = (wL \times R_{-1,y'} + wT \times R_{x',-1} - wTL \times R_{-1,-1} + (64 - wL - wT + wTL) \times \text{pred}(x',y') + 32) \gg 6$$

図 4-8 に、0(Planar)と1(DC)、18(水平)、50(垂直)、2~10(右上斜め)、および、58~66(左下斜め)それぞれの場合での、予測画像の各画素の位置と利用する隣接参照画素の位置関係を示す。PDPC は、幅と高さの両方が 4 以上のブロックに適用される。

PDPC の適用により、原画像との差分が少ない予測画像を生成できる。

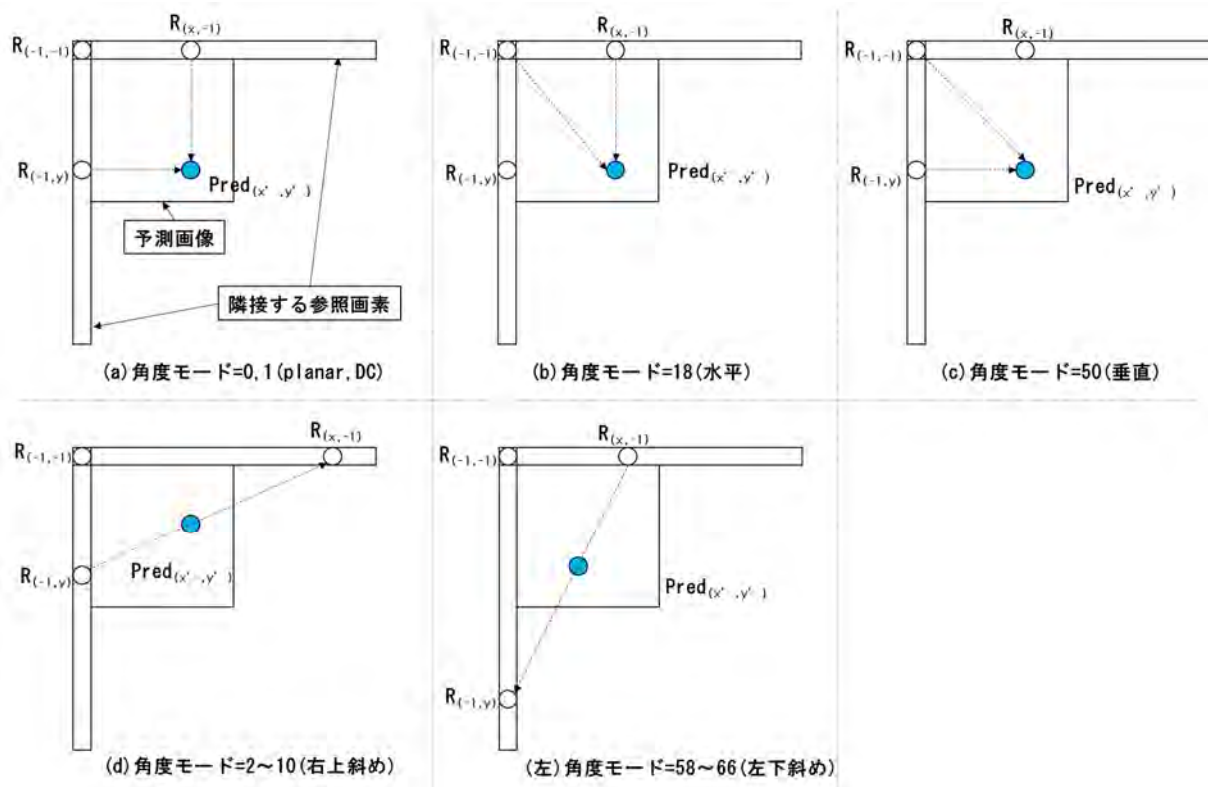


図 4-8 PDPC の各予測モードにおける隣接参照画素

4.6. MIP (Matrix-based Intra Prediction)

MIP(行列イントラ予測)は、輝度および色差成分のブロックに対し、図 4-9 のように、対象とする矩形ブロックの左側と上側に隣接する参照画素に対する平均処理、行列演算、および補間処理に基づいて予測画像を生成する。

平均処理は、左側と上側に隣接する参照画素($\text{bdry}_{\text{left}}$ 、 bdry_{top})を平均した画素 $\text{bdry}_{\text{red}}^{\text{left}}$ 、 $\text{bdry}_{\text{red}}^{\text{top}}$ を生成する。この画素の数はブロックサイズによって異なり、 $H=W=4$ の場合はそれぞれ 2 画素となり、他の場合はそれぞれ 4 画素となる。行列演算は、 $\text{bdry}_{\text{red}}^{\text{left}}$ と $\text{bdry}_{\text{red}}^{\text{top}}$ からなるベクトル bdry_{red} に対し、ブロックサイズに応じて規格で規定されている変換行列 A とオフセット b を用いた演算を行い、予測画像の一部を生成する。補間処理は、隣接した参照画素、および、行列処理にて生成した予測画像を用いて、水平、垂直の順で近傍の 2 つの有効画素から重み付け平均を実施し、最終的な予測画像を生成する。

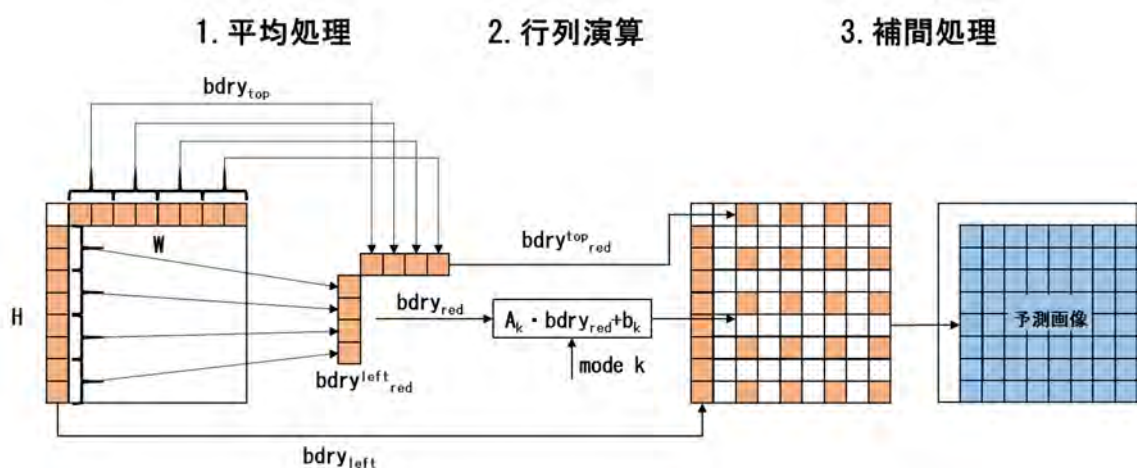


図 4-9 8x8 ブロック時の MIP 処理の概要

4.7. ISP (Intra Sub-Partitions)

ISP (イントラサブパーティション分割) は、輝度成分のブロックのみに適用され、処理対象画素と予測参照画素との距離を短くすることでより効率の良いイントラ予測を可能とするために、輝度成分ブロックを垂直または水平に等分割する。分割したサブブロック単位で予測画像の生成および残差信号の変換処理を行う。

分割の数は、ブロックサイズが 4x8 もしくは 8x4 の場合に 2、その他の場合に 4 となり、1 つのサブブロックあたりの画素数は 16 画素以上となる (16 画素以上であれば 1 画素幅のサブブロックまで分割可能)。分割後の輝度成分ブロックのイントラ予測の水平の幅を 4 以上と制約する。このため、分割後の輝度成分ブロックの水平の幅が 4 未満の時、同輝度成分に対する予測処理と変換処理の単位が異なる。なお、ISP は 16 画素以下の輝度成分ブロックでは利用できない。

図 4-10 および図 4-11 に、ブロックサイズが 16x16 の場合に垂直分割し予測処理と変換処理単位が一致する例と、ブロックサイズが 4x8 の場合に垂直分割し予測処理と変換処理単位が異なる例を示す。

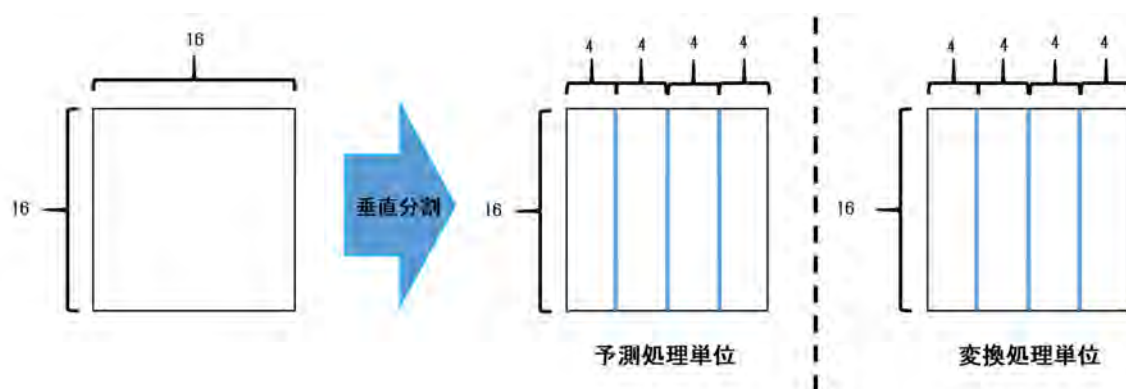


図 4-10 16x16 ブロック時の ISP 分割の例

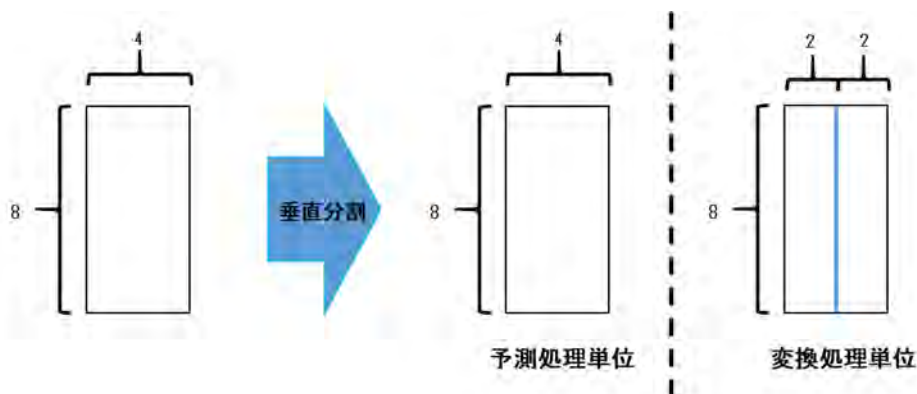


図 4-11 4x8 ブロック時の ISP 分割の例

5. インター予測

5.1. AMVP モード

5.1.1. AMVP (Adaptive Motion Vector Prediction)

動きベクトルの最も基本的な復号方法は、復号側で算出した予測動きベクトルに、エンコーダ側から伝送されてくる差分動きベクトル(MVD)を加算することである。この予測ベクトルを算出する処理をAMVPと呼ぶ。

AMVP では、HEVC と同様に複数の予測ベクトル候補を算出し、その中からエンコーダ側で選択したものを予測ベクトルとする。予測ベクトル候補の算出法は、HEVC と同様の Spatial MVP、Temporal MVP、Zero MVに加え、History-based MVP が新たに導入された。それぞれ下記の順で、予測ベクトル候補の数が2に到達するまで予測ベクトルを導出する。

1. Spatial MVP
2. Temporal MVP
3. History-based MVP
4. Zero MV

Spatial MVPとTemporal MVPを導出する際に参照するブロックの配置は図 5-1 に示す通りHEVCと同様である。

History-based MVP は、HMVP リストと呼ばれるリストに保管されている動きベクトルを参照して予測する。HMVPリストには、CU 毎に復号された動きベクトルが最大6つFIFO形式で格納される。すなわち、既に6つの動きベクトルがリストに格納されている場合、新しい動きベクトルをリストに追加する際、最も古い動きベクトルはリストから削除される。HMVPリストは、CTU行の先頭で初期化される。

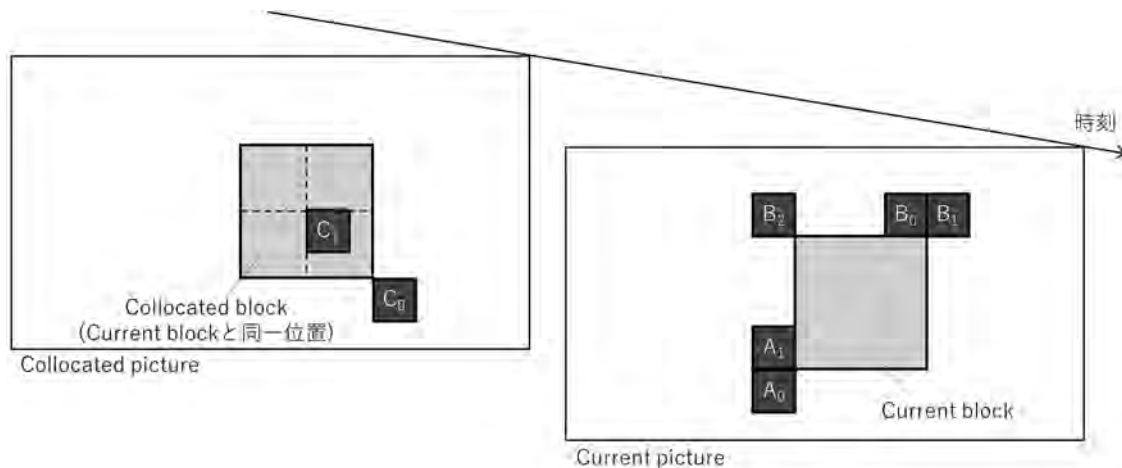


図 5-1 Spatial MVP と Temporal MVP の候補

5.1.2. AMVR (Adaptive Motion Vector Resolution)

HEVC の SCC 拡張プロファイルではスライス毎に動きベクトルの精度を選択することができた。一方、VVC では CU 毎に動きベクトルの精度を選択する、AMVR が採用された。具体的には、アフィン動き補償予測(エラー! 参照元が見つかりません。節参照)を行う場合には 1/16、1/4、整数画素精度の中からいずれか、IBC(エラー! 参照元が見つかりません。節参照)を行う場合は整数、4 画素精度のいずれか、アフィン及び IBC 以外の通常の動き補償予測の場合には 1/4、1/2、整数、4 画素精度の中からいずれかの精度を、それぞれ CU 毎に選択することができる。ここで、1/2 画素精度が選択された場合は、エラー! 参照元が見つかりません。節に記載の通り通常の 8 タップではなく 6 タップの内挿フィルタを用いる。当該 CU において非零の差分動きベクトルが復号された場合に、動きベクトル精度の選択情報が追加で復号される。復号後の動きベクトル(= 予測動きベクトル + 差分動きベクトル)がエンコーダ側で選択した精度になることを担保するため、予測動きベクトルについても当該 CU の動きベクトル精度で量子化した後に、差分動きベクトルを加算する。

5.1.3. SMVD (Symmetric Motion Vector Difference)

SMVD は、いわゆる双方向予測を行う CU において、差分動きベクトルが過去方向と未来方向の参照フレームについて上下左右対称になることを仮定し、動きベクトルの符号量を削減するツールである。

SMVD は、まずスライス単位で L0 側の時間的に最も近い参照フレームと、L1 側の時間的に最も近い参照フレームが、それぞれ当該フレームに対して未来と過去、または過去と未来といった組み合わせとなる場合に適用できる。

上記の条件を満たすスライスにおいて、CU 毎に SMVD を使用するかどうかのフラグを伝送する。当該 CU に SMVD を使用する場合、L0 の差分動きベクトルのみが復号され、L0 の参照インデックス、L1 の参照インデックス、L1 の差分動きベクトルの復号が省略される。参照インデックスは L0/L1 それぞれについて当該フレームに時間的に最も近いフレームとみなす。L1 の差分動きベクトルは、L0 の差分動きベクトルの符号を x 方向、y 方向それぞれについて反転させたものとみなす(図 5-2 参照)。上記のように算出した差分動きベクトルと、予測動きベクトル(MVP)を加算することで、当該ブロックの動きベクトルを復号する。

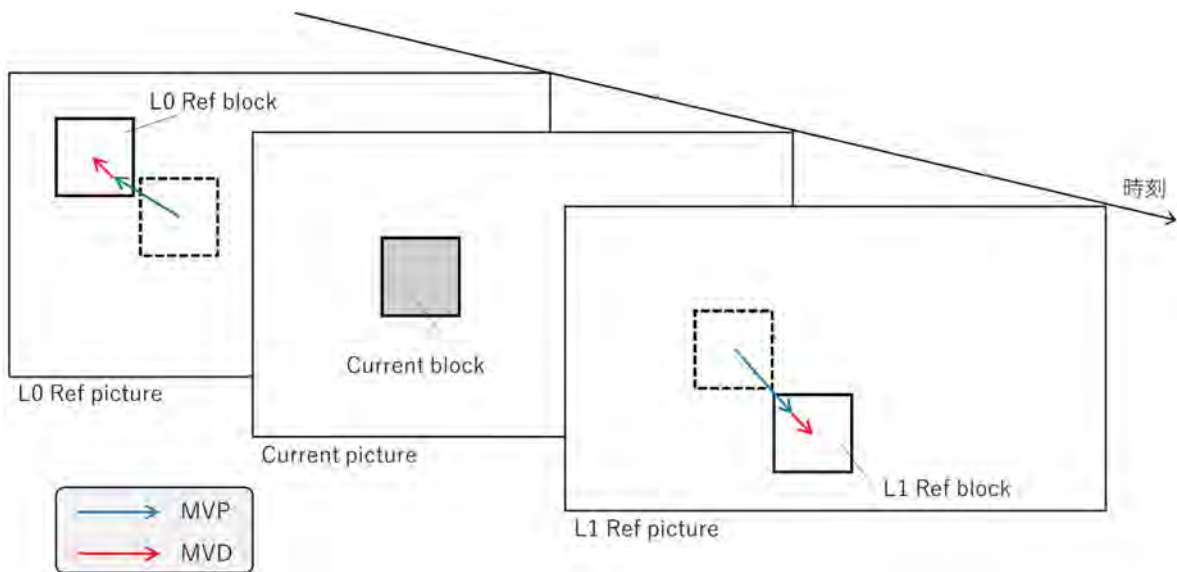


図 5-2 SMVD

5.2. マージモード

5.2.1. Regular Merge

VVC では、HEVC で採用されたマージモードが拡張された。HEVC のマージモードに対応する通常のマージモードは、後述する MMVD、CIIP、GPM などと区別するために **Regular merge** と呼ばれる。HEVC との主な差分は以下の通りである。

- 最大候補数の拡張 (5→6)

HEVC のマージモードでは、CU 毎に最大 5 つのマージ候補を導出し、その中から選択することができた。VVC では、最大のマージ候補数が 6 へ拡張された。

- マージ候補の算出方法

VVC では、以下の 5 つの方法を用いてマージ候補を算出する。それぞれの算出方法から、複数のマージ候補が算出される場合もある。HEVC と比較して、結合双予測が廃止され、その代わりに History-based MVP と Pairwise average MVP が追加された。

1. Spatial MVP
2. Temporal MVP
3. History-based MVP
4. Pairwise average MVP
5. Zero MV

Pairwise average MVP は、マージ候補リストから二つの動きベクトルを抽出し、それらの平均値を算出することで新たなマージ候補を算出する。

5.2.2. MMVD (Merge mode with Motion Vector Difference)

Regular merge では、差分動きベクトルを一切復号せず、マージリストから選択した動きベクトルをそのまま用いることしかできない。これに対して、MMVD では、マージモードにおいても差分動きベクトルを復号することを可能としている。ただし、通常の差分動きベクトルの復号とは異なり、差分動きベクトル

を復号するのは x 方向か y 方向のいずれか一方のみで、取り得る差分値も離散的である。

具体的には、上記の Regular merge を用いる CU において、まず MMVD を使用するかどうかを示すフラグを伝送する。MMVD を使用する場合、追加で Direction IDX と Distance IDX を伝送する。Direction IDX を用いて、表 5-1 に示す通り差分動きベクトルの符号を復号する。更に、Distance IDX を基に、

表 5-2 に示すように差分動きベクトルの絶対値を復号する。なお、差分動きベクトルの精度については、ピクチャ単位で整数精度とするか小数精度とするかを選択することができる。最終的に、上記の符号と絶対値の積をとることで差分動きベクトルを得る。よって、符号の値が 0 の場合は、当該軸方向の差分動きベクトルが 0 であることを意味する。

なお、双予測を行うブロックにおいても、差分動きベクトルは L0 と L1 で合わせて一つのみ復号する。この時、L0 では復号した差分動きベクトルをそのまま使用する。L1 においては、L1 参照フレームが L0 参照フレームと時間的に同じ方向にある場合は、前記の復号した差分動きベクトルをそのまま用い、L1 参照フレームが L0 参照フレームと時間的に逆の方向にある場合は、復号した差分動きベクトルの符号を反転してから使用する。

表 5-1 Direction IDX と符号との対応

| Direction IDX | x 方向の符号 | y 方向の符号 |
|---------------|---------|---------|
| 0 | +1 | 0 |
| 1 | -1 | 0 |
| 2 | 0 | +1 |
| 3 | 0 | -1 |

表 5-2 Distance IDX と差分動きベクトルの絶対値との対応

| Distance IDX | 差分動きベクトルの絶対値 | |
|--------------|--------------|---------|
| | 小数精度の場合 | 整数精度の場合 |
| 0 | 1/4 | 4/4 |
| 1 | 2/4 | 8/4 |
| 2 | 4/4 | 16/4 |
| 3 | 8/4 | 32/4 |
| 4 | 16/4 | 64/4 |
| 5 | 32/4 | 128/4 |
| 6 | 64/4 | 256/4 |
| 7 | 128/4 | 512/4 |

5.2.3. CIIP (Combined Inter-picture merge and Intra-picture Prediction)

CIIP は、マージモードの一つで、インター予測画像とイントラ予測画像の重み付き平均によって予測画像を生成する。

具体的には、CIIP モードの予測画像 predSamplesComb を、マージモードで導出したインター予測画像 predSamplesInter と、Planar モードで導出したイントラ予測画像 predSamplesIntra との重み付き平均

によって算出する。なお、同一の重み係数を輝度と色差の両方に適用する。

$$\text{predSamplesComb} = (w * \text{predSamplesIntra} + (4 - w) * \text{predSamplesInter} + 2) \gg 2$$

上式の重み係数 w は、左と上に隣接する符号化ブロックの予測モードによって決まる。

図 5-3 の左隣の A と上隣の B が両方ともイントラモードで符号化されていた場合、 w は 3 とする。もし、 A と B が両方ともイントラモードでなかった場合、 w は 1 とする。それ以外の場合、 w は 2 とする。

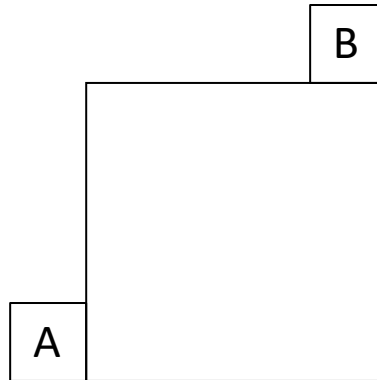


図 5-3 CIIP の重み係数 w の導出

CIIP は、スキップモード以外、CU の幅と高さの両方が 8 以上かつ 128 未満、かつ幅と高さとの比が $1/8$ より大きく 8 未満の場合に適用することができる。なお、BCW や GPM とは排他的に選択される。CIIP ではインター予測とイントラ予測の両方が必要となるが、デコーダのパイプライン構成において、各々の予測処理が並列で実施されかつインター予測が先行して完了する前提の下で、これらの制約を含め該当ステージに要求される追加サイクル数が最小となるように工夫されている。

5.2.4. GPM (Geometric Partitioning Mode)

GPM は、マージモードの一つで、通常のブロック分割では対応できない斜め線での分割された動き補償を可能とする。分割の形状を表す角度と距離の組み合わせを示すインデックスと、2 つの領域 A と B の動きベクトルを導出する 2 つのマージインデックスを指定する。この時の 2 つのマージインデックスは、最悪のメモリバンド幅を通常の双方向予測と同等を保証するために、それぞれ単方向の予測を導出する。

分割パターン数は 64 通り存在する。図 5-4 は、 8×8 ブロックの場合の 64 通りの重み係数のテーブルである。重み係数が 8 の場合、領域 A が選択され、重み係数が 0 の場合、領域 B が選択される。それ以外の場合は、重み係数に応じて 2 つの領域 A の動き補償予測画像と領域 B の動き補償予測画像の重み付き平均をして最終的な予測画像を生成する。

(0, 0)、つまり補正しない場合に対応する SAD は 3 / 4 として、初期位置から動きにくくし安定性を向上させる。また、この時点で、SAD 値が閾値より小さい場合は、処理を打ち切り、DMVR は適用されない。

続いて、上記 5x5 の内、中央の 3x3 の範囲(つまり±1 範囲)に補正ベクトルがある場合は、更に上記の最小 SAD となる点と左右に隣接する SAD(もしくは上下に隣接する SAD)の3点を通る 2 次曲線の極小値から X 方向(もしくは Y 方向)の小数精度の動きベクトル補正を推定し、最終的な補正動きベクトル *mvRefine* を導出する。

以上の処理から得られた補正後の動きベクトルによる L0 予測画像と L1 予測画像から、通常の動き補償での補間フィルタを用いて最終的な予測画像を生成する。

なお、上記の処理で求められた SAD 値は、後段で BDOF (Bi-Directional Optical Flow)を用いる場合の適用条件にも使われる。

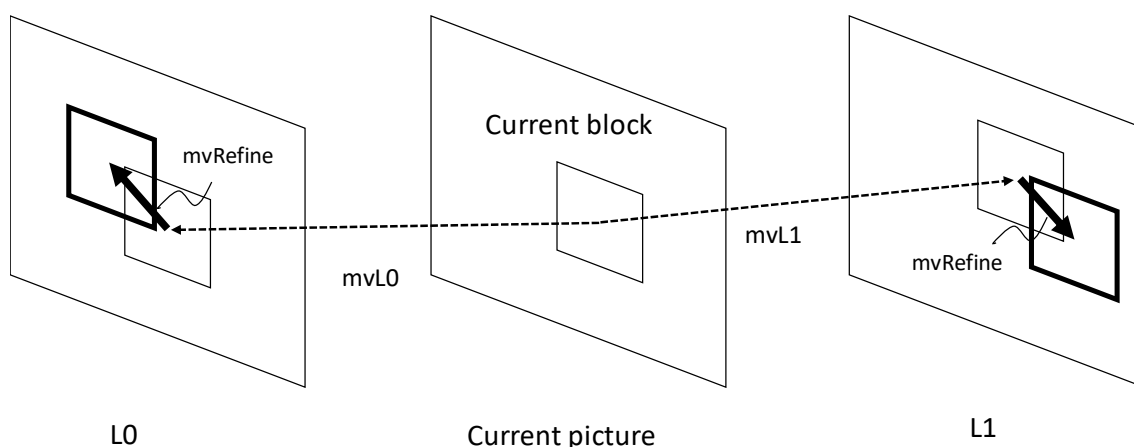


図 5-5 DMVR の概要図

DMVR は、幅、高さが 8 以上、かつ幅と高さとの積が 128 以上となるマージモードの CU に対して適用され、幅、高さが 16 以上の場合は、16 画素単位に分割されて実行される。また、双方向予測で、かつ、対象画像からの 2 つの参照画像までの距離が等しく、2 つの参照画像が対象画像に対して時間的に反対方向に位置することが適用条件である。なお、MMVD マージモード、WP、BCW、CIIP の時には適用されない。

5.3. サブブロックモード

5.3.1. アフィンモード

VVC では、アフィン動きモデルに基づいたサブブロックベースでのアフィン動き補償予測を行うことができる。アフィン動き補償予測では、図 5-6 のように CU 毎にコントロールポイントベクトル (*cpmv*) と呼ばれる動きベクトルを 2 つないし 3 つ伝送し、コントロールポイントベクトルに基づいて 4x4 画素サイズのサブブロック毎の動きベクトルを復号側で導出する。コントロールポイントが 2 つの場合を 4 パラメータ、3 つの場合を 6 パラメータとも呼び、4 パラメータとするか 6 パラメータとするかは CU 毎に選択することができる。サブブロック毎の動きベクトル (*mvsb*) は、下記の式に基づいて、各サブブロックの中央 (*x, y*) に位置する画素の動きベクトルとする。

$$\begin{cases} mvsb_x = Cx + Dy + cpmv_{0x} \\ mvsb_y = Ex + Fy + cpmv_{0y} \end{cases}$$

4パラメータの場合は、

$$\begin{cases} C = F = \frac{cpmv_{1x} - cpmv_{0x}}{w} \\ E = -D = \frac{cpmv_{1y} - cpmv_{0y}}{w} \end{cases}$$

6パラメータの場合は、

$$\begin{cases} C = \frac{cpmv_{1x} - cpmv_{0x}}{w} \\ D = \frac{cpmv_{2x} - cpmv_{0x}}{h} \\ E = \frac{cpmv_{1y} - cpmv_{0y}}{w} \\ F = \frac{cpmv_{2y} - cpmv_{0y}}{h} \end{cases}$$

とそれぞれ定義される。なお、ここで w はCUの幅、 h はCUの高さである。

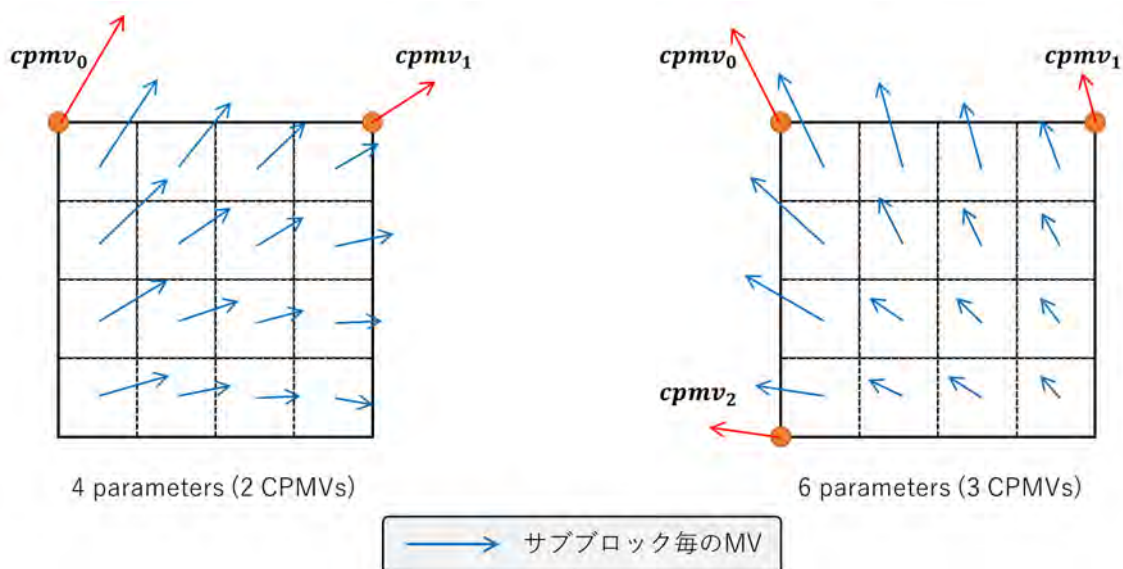


図 5-6 4パラメータ及び6パラメータアフィン

以上のように、VVCのアフィン動き補償予測は、いわゆるアフィン変換による画素単位での動き補償(回転、拡大を含む)を厳密には実現しておらず、あくまでサブブロック単位での平行移動モデルに基づく動き補償である。そのため、画素単位の変換については後述するPROF(オプティカルフローによる予測値洗練化)によって動き補償後の予測画像を補正することで疑似的に再現している。

5.3.2. PROF (Prediction Refinement with Optical Flow)

前述のアフィン動き補償予測が適用されたCUにおいて、適用条件を満たす場合に、輝度信号に対してPROFが自動的に適用される。主な適用条件は、当該CUの2つもしくは3つのコントロールポイントベクトルが同一ではないこと、である。PROFは、ピクチャヘッダにて有効/無効を制御することができる。PROFを適用するCUについては、以下の処理にて予測値の補正量を算出し、アフィン動き補償によって得られた予測値に前記補正量を加算することで最終的な輝度信号の予測値を算出する。

1. 輝度勾配の算出

サブブロック内の各画素位置 $\{(i, j) | i = 0 \dots w - 1, j = 0 \dots h - 1\}$ について、 $[-1, 0, 1]$ の係数を持つ 3-tap フィルタをも用いて、x 方向と y 方向の輝度勾配 $g_x(i, j), g_y(i, j)$ をそれぞれ算出する。

2. 補正量の算出

サブブロック内の各画素と、サブブロック中央に位置する画素との相対位置をベースに、以下の式で補正量 $\Delta I(i, j)$ を算出する。

$$\Delta I(i, j) = g_x(i, j) * \Delta v_x(i, j) + g_y(i, j) * \Delta v_y(i, j)$$

ここで、 $\Delta v_x(i, j), \Delta v_y(i, j)$ はそれぞれ以下の通り算出する。

$$\begin{cases} \Delta v_x(i, j) = C * dx(i, j) + D * dy(i, j) \\ \Delta v_y(i, j) = E * dx(i, j) + F * dy(i, j) \end{cases}$$

$$\begin{cases} dx(i, j) = i - x_{sb} \\ dy(i, j) = j - y_{sb} \end{cases}$$

なお、 C, D, E, F の値は、エラー！参照元が見つかりません。節で説明した値と同一である。また、 x_{sb}, y_{sb} は各サブブロックの中央の x 座標及び y 座標であり、以下の通り算出する。

$$\begin{cases} x_{sb} = \frac{w - 1}{2} \\ y_{sb} = \frac{h - 1}{2} \end{cases}$$

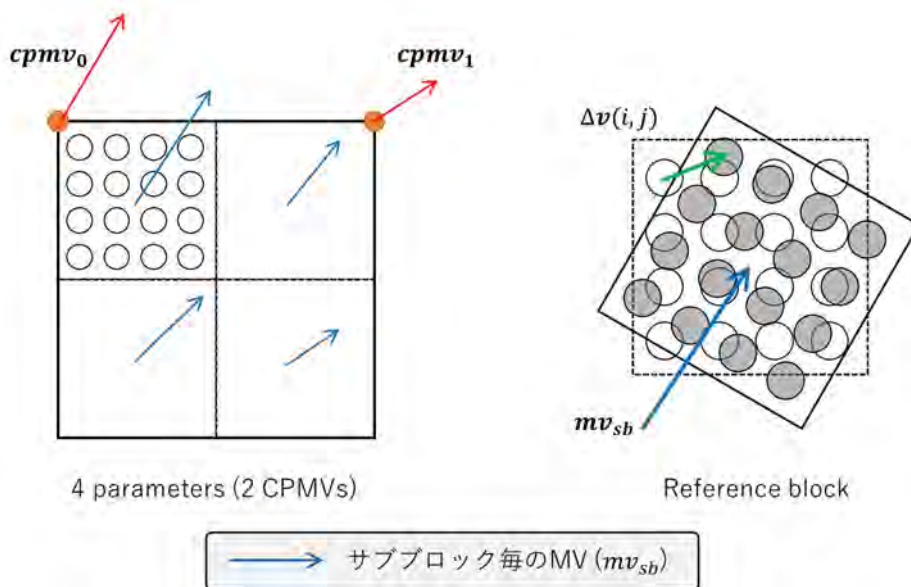


図 5-7 PROF

5.3.3. SbTMVP (Subblock-based Temporal MVP)

SbTMVP は、時間方向予測を用いてサブブロック単位での動きベクトルを導出するツールである。当該 CU に SbTMVP を適用する場合に、時間方向予測によって 8x8 画素サイズのサブブロック毎にベクトルを算出する。時間方向予測の方法は、通常の動きベクトル予測の時間方向予測と異なり、空間方

向から予測した動き量分だけシフトした領域を参照し、動きベクトルを予測する。具体的には、図 5-8 の A1 ブロックの動きベクトルが Collocated picture を指している場合、当該の動きベクトルの値をシフト量に設定する。A1 ブロックの動きベクトルが Collocated picture 指していない場合は、シフト量を 0 にする。その後、Collocated picture において、当該 CU と同一位置からシフト量分だけシフトした位置の動きベクトルを参照して、サブブロック毎のベクトルを算出する。

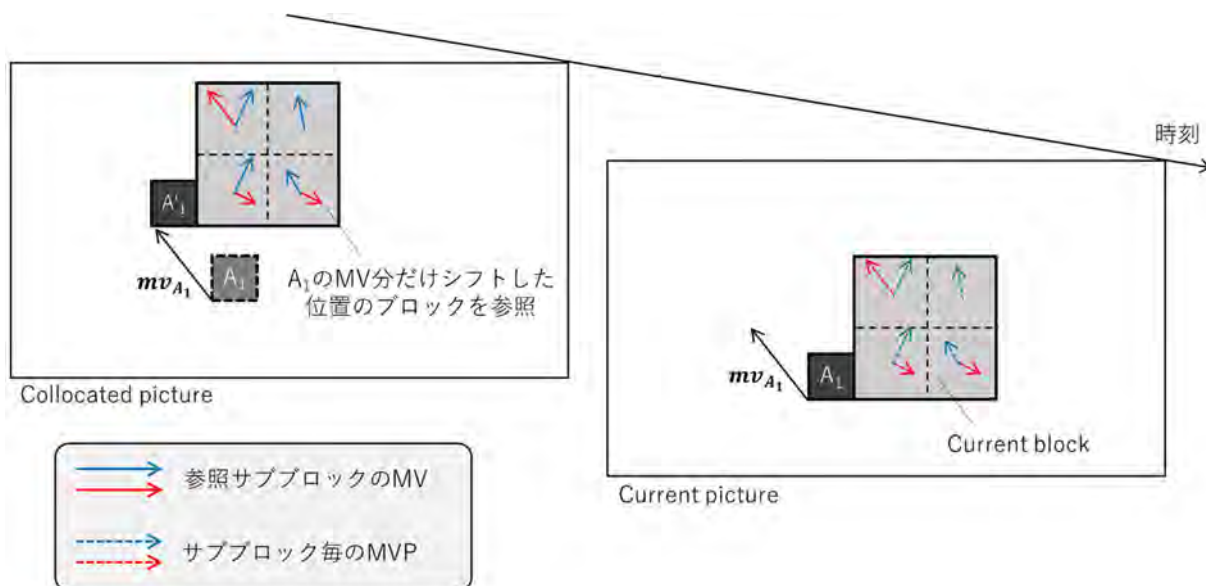


図 5-8 SbTMVP

5.4. 動き補償

5.4.1. MC (Motion Compensation)

MC のための補間フィルタは、アフィンモード以外の場合は、HEVC のフィルタの画素精度を拡張したものをを用いる。輝度は最大 8 タップ、色差は、最大 4 タップのフィルタが用意され、輝度は、1/4 から 1/16 画素精度に、色差 1/8 から 1/32 画素精度に拡張している。アフィンモードの場合は、新規の輝度用の 6 タップフィルタ、色差用の 4 タップフィルタを用いる。輝度ブロックでの最小単位は、HEVC と同様に、片方向予測の場合は 4x8 または 8x4 で、双方向予測は 8x8 に制限される。アフィンモードの場合は、4x4 単位の MC ブロックが発生するが、補間フィルタのタップ数が 6 タップと短くし、かつ、隣接する 4x4 ブロック間での動きベクトルのばらつきを制限することで、最悪のメモリバンド幅は HEVC と同等になることを保証している。

MC の補間画像生成の拡張技術として SIF、Wrap around MC、RPR があり、MC 後の予測画像生成技術として、BDOF、GPM、BCW、WP、CIIP がある。

5.4.2. SIF (Switched Interpolation Filter)

輝度の 1/2 画素位置の補間フィルタ係数は、2 種類あり、切り替えて使われる。

表 5-3 の $hpellIdx=0$ のフィルタ係数は、HEVC と同じ 8 タップであるが、 $hpellIdx=1$ は、異なる周波数特性を持つ 6 タップフィルタが採用されている。

マージモードの場合、 $hpellIdx$ は、マージ候補インデックスに基づいて隣接ブロックから推定される。マージモードでない場合は、アフィンモードでなく、かつ、イントラブロックコピー(IBC)でない場合で AMVR (Adaptive Motion Vector Resolution) モードの 1/2 画素精度の時に $hpellIdx$ は 1 を選択する。

表 5-3 1/2 画素位置の輝度の補間フィルタ係数

| hpelIdx | interpolation filter coefficients | | | | | | | |
|---------|-----------------------------------|--------|--------|--------|--------|--------|--------|--------|
| | f[0] | f[1] | f[2] | f[3] | f[4] | f[5] | f[6] | f[7] |
| 0 | -1 | 4 | -11 | 40 | 40 | -11 | 4 | -1 |
| 1 | 0 | 3 | 9 | 20 | 20 | 9 | 3 | 0 |

5.4.3. Wrap around MC (Wrap around Motion Compensation)

Wrap around MC は、360° 映像のための符号化ツールであり、水平方向のピクチャの外側から動き補償予測画像を作る場合のパディングの方法を変更する。

通常の動き補償では、動きベクトルが参照ピクチャの外側の画素を指し示す場合、ピクチャ境界の画素と同じ画素が外側に存在すると仮定して、動き補償画像が作成される。

正距円筒投影(ERP: Equi-Rectangular Projection)で表現された 360° 映像では、水平方向の画素は、円筒状に連続している。従って、水平方向については、参照ピクチャの外側を示す動きベクトルに対して参照ピクチャの反対側の画素を用いることで連続した画素を用いることができる。Wrap around MC では、水平方向ピクチャのパディング方法を変更し、図 5-9 のように、左側でピクチャの外側になる部分としてピクチャの右側の画素値を参照する。ERP では、右端と左端で同じ画像を重複して配置する場合を考慮して、重複量のオフセット値(pps_pic_width_minus_wraparound_offset)を設定している。なお、上下方向には、通常の動き補償と同様のパディング方法を適用する。

Wrap around MC は、符号化効率の改善のみならず、従来、投影の不連続な部分で生じていた視覚的な歪を改善することができる。

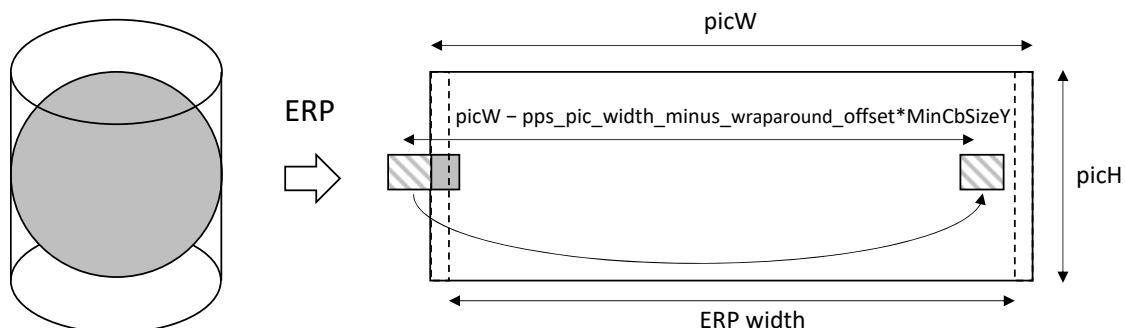


図 5-9 ERP と Wrap around motion compensation

5.4.4. RPR (Reference Picture Resampling)

RPR は参照ピクチャと対象ピクチャのサイズが異なる場合に、参照ピクチャのスケージングをしながら動き補償画像を生成する技術であり、画素単位で補間フィルタの適用対象となるサンプリング位置を決定することで対応する。さらに、符号化対象のピクチャの幅に対する参照ピクチャの幅の比、及び高さに対する比をそれぞれ計算し、参照ピクチャが符号化対象のピクチャよりも大きい場合に、その比率に応じて、補間フィルタを切り替える。

アフィンモードの場合で、比が 7/4 より大きい場合に 2 倍用の補間フィルタを用い、比が 5/4 より大きい場合に 1.5 倍用の補間フィルタを用いる。上記以外の場合は、通常のアフィンモード用のフィルタを用いる。

アフィンモードでない場合も、同様に、比が 7/4 より大きい場合と、比が 5/4 より大きい場合、それぞれ

別途補間フィルタ係数を切り替える。それ以外の場合は、通常の補間フィルタを用いる。なお、比が 1 以外の場合は、SIF による 1/2 画素精度の補間フィルタの切り替えは行わない。

5.5. 予測画像生成

5.5.1. BDOF (Bi-Directional Optical Flow)

BDOF は、双方向予測モードにおいて、2 つの動き補償予測画像の空間方向と時間方向の勾配情報を用いて 4x4 ブロック単位にオプティカルフローを求め、画素毎の予測画像の修正を行うことで予測効率を改善する手法である。

まず、処理対象のブロックの 2 つの動き補償予測画像 L0 と L1 のそれぞれに対して、4x4 のサブブロック毎に画素単位に水平方向と垂直方向の勾配を計算する。勾配は、対象画素に対して、水平方向ならば、左隣の画素と右隣の画素の差分、垂直方向ならば、対象画素に対して、上隣の画素と下隣の画素との差分である。この時、4x4 サブブロックに対して、周囲 1 画素を参照するため、6x6 ブロックの情報が必要となるが、処理対象のブロックの外側の画素に対しては、周囲の画素をコピーしてパディングして求める。また、時間方向の勾配として 2 つの動き補償予測画像 L0 と L1 の差分を求める。

次に、空間方向と時間方向の勾配情報から、オプティカルフローを求める計算を行い、4x4 サブブロック単位での補正オフセット(v_x, v_y)を求める。なお、オプティカルフローの導出においては、乗算を用いることなくシフト演算と Sign 処理のみで簡易的に求めることで、演算量の増加を抑制している。

最後に、2 つの動き補償予測画像 L0 と L1 の空間勾配情報と、補正オフセット(v_x, v_y)から、画素単位に動き補償予測画像を修正する値を求め、双方向予測の式の中で丸めオフセットと同時に加えることで、動き補償予測画像を補正する。

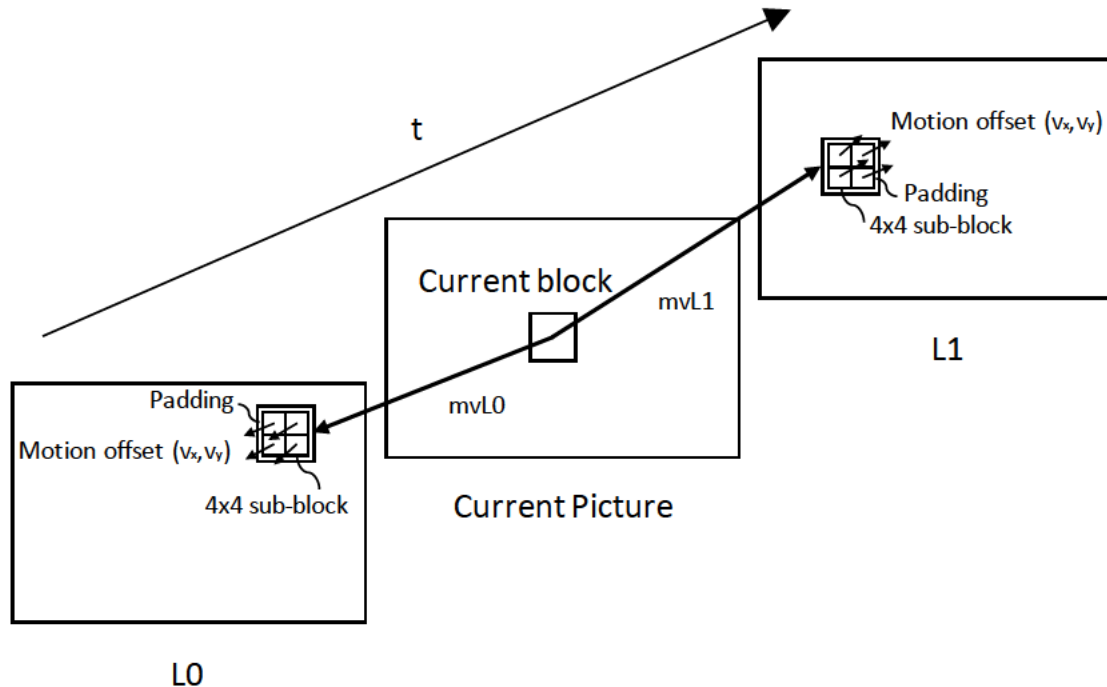


図 5-10 BDOF の概念図

BDOF は、幅、高さが 8 以上、かつ幅と高さとの積が 128 以上の輝度の CU に対して適用され、幅、高さが 16 以上の場合は、16 画素単位の処理ブロックに分割されて実行される。また、双方向予測で、

かつ、対象画像からの 2 つの参照画像までの距離が等しく、2 つの参照画像が対象画像に対して時間的に反対方向に位置することが条件である。また、アフィンモードなどのサブブロックマージモード、SMVD、WP、BCW、CIIP の時には適用されない。

なお、前段で、DMVR が適用された場合には、DMVR での SAD 値が閾値より小さい場合は、BDOF はスキップされる。

5.5.2. BCW (Bi-prediction with CU-level Weights)

BCW は、Coding Unit 単位に予め決められた複数の重み係数の中から選択して重み付き平均によって予測画像を生成する双方向予測モードである。{4, 5, 3, 10, -2} の 5 つの重み係数 w が用意されていて、L0 の予測画像と、L1 の予測画像に対して、 $(8 - w) : w$ の比で重み付き平均を行い、予測画像を生成する。なお、同一の重み係数を輝度と色差の両方に適用する。

重み係数は、インデックスによって選択できる。マージモードの場合、BCW のインデックスは、マージ候補インデックスに基づいて隣接ブロックから推定される。マージモードでない場合、双方向予測で、選択された L0 と L1 の参照ピクチャに Weighted Prediction の重み係数が存在しなくて、CU サイズの幅と高さを乗じた値が 256 以上の場合に適用され、BCW のインデックスは CU 単位に明示的に符号化、復号される。

また、参照可能な参照ピクチャが 1 ピクチャも未来にない場合、負の重み係数を含む 5 つすべての重み係数が使うことができるが、参照可能な参照ピクチャのいずれかが未来にある場合、負の重み係数を含まない 3 つのみを使うことができる。

5.5.3. WP (Weighted Prediction)

WP は、明るさが変動するフェードシーン等に有効な予測方式である。VVC では、予測画像の生成については、AVC の Explicit Weighted Prediction と HEVC の Weighted Prediction 仕様を踏襲している。BCW とは排他的に選択される。

6. 変換・量子化

この章では VVC における変換および量子化に関する技術について述べた後、可逆符号化の実現方法について述べる。

6.1. 変換

VVC における変換に関する技術について述べる。

6.1.1. MTS (Multiple Transform Selection)

HEVC では、変換基底として、DCT2 と DST7 が採用されているが、水平・垂直方向で任意に変換基底を選択することはできない。具体的には、輝度イントラ 4x4 画素ブロックの予測残差信号の場合、水平・垂直方向に一律で DST7 が適用され、それ以外の NxN 画素ブロック(N=4,8,16,32)の予測残差信号の場合、水平・垂直方向に一律で DCT2 が適用される。

一方、VVC では一次変換の変換基底として、エラー! ブックマークが自己参照を行っています。で定義される DCT2、DST7、DCT8 の 3 種類から水平・垂直方向毎に任意に選択可能とする MTS という手法が採用されている。ここで、エラー! ブックマークが自己参照を行っています。において、N は入力信号の標本点数を表し、VVC では最大 64 まで拡張されている。また、HEVC と同様に、各変換基底の係数は、符号付 8 ビット整数精度で近似されている。なお、輝度信号では、DCT2、DST7、DCT8 の 3

種が利用でき、色差信号では DCT2 のみが利用できる。

表 6-1 DCT2/DST7/DCT8 の基底関数の定義

| 変換タイプ | 基底関数 $T_i(j)$, $i, j = 0, 1, \dots, N-1$ |
|-------|---|
| DCT2 | $T_i(j) = \omega_0 \cdot \sqrt{\frac{2}{N}} \cdot \cos\left(\frac{\pi \cdot i \cdot (2j + 1)}{2N}\right)$ $, \text{where } \omega_0 = \begin{cases} \sqrt{\frac{1}{2}} & i = 0 \\ 1 & i \neq 0 \end{cases}$ |
| DCT8 | $T_i(j) = \sqrt{\frac{4}{2N + 1}} \cdot \cos\left(\frac{\pi \cdot (2i + 1) \cdot (2j + 1)}{4N + 2}\right)$ |
| DST7 | $T_i(j) = \sqrt{\frac{4}{2N + 1}} \cdot \sin\left(\frac{\pi \cdot (2i + 1) \cdot (j + 1)}{2N + 1}\right)$ |

DCT2

VVC における DCT2 は、HEVC 同様に、Chen の高速バタフライ演算を利用できるように設計されている。また、あるサイズの変換行列は、一つ大きいサイズの変換行列に内包されているため、変換処理・逆変換処理において、小さいサイズの変換回路を再利用して、大きいサイズの変換回路を構成することができる。HEVC では、DCT2 で利用できる標本点数 N は 4, 8, 16, 32 の 4 つであったが、VVC では、2, 4, 8, 16, 32, 64 の 6 つに拡張されている。また、標本点数 $N=4, 8, 16, 32$ に関しては、HEVC と VVC の DCT2 の変換行列の係数は同じであるため、DCT2 の変換回路の一部を HEVC と VVC とで共通化することができる。

DST7/DCT8

一般的に、イントラ予測の予測残差信号の係数の絶対値は参照画素から離れるほど大きな値になるという傾向がある。HEVC では、特にその傾向が顕著となる輝度 4×4 画素ブロックにおいて、DCT2 より効率的に低周波数域に変換係数を集中させることができる DST7 が変換基底として導入された。VVC では予測残差信号の係数の偏りに適した変換基底をより積極的に選べるように、標本点数 $N=4, 8, 16, 32$ の DST7 と DCT8 が導入されている。

ゼロアウト(Zero-out)

VVC では、一次変換の行列演算による演算回数を低減するため、大きいサイズの変換では高周波数領域の係数値を強制的にゼロ化する仕組みが導入されている。具体的には、標本点数 $N=64$ の DCT2、 $N=32$ の DST7 および DCT8 を用いて一次変換を行う場合、高周波領域側の半分の領域にゼロアウトが適用される。規格上では、 $N=64$ の DCT2、および $N=32$ の DST7/DCT8 の逆変換の変換行列のうち、高次側の基底ベクトルが不要となる。

Explicit MTS/Implicit MTS

MTS の制御方式には、Explicit MTS モードと Implicit MTS モードの 2 つのモードがある。Explicit MTS モードは、CU 単位でシグナルされるシンタックス `mts_idx` に基づいて、各水平・垂直方向の変換基底を指定することができる。**エラー! ブックマークが自己参照を行っています。**に Explicit MTS モードで選択可能な変換基底を示す。Explicit MTS モードを符号化ツールとして有効化するかどうかは、SPS において、イントラ予測・インター予測毎に指定可能である。

表 6-2 Explicit MTS モードにおける `mts_idx` に対応する変換基底

| <code>mts_idx</code> | 0 | 1 | 2 | 3 | 4 |
|----------------------|------|------|------|------|------|
| 水平方向の変換基底 | DCT2 | DST7 | DCT8 | DST7 | DCT8 |
| 垂直方向の変換基底 | DCT2 | DST7 | DST7 | DCT8 | DCT8 |

Implicit MTS モードは、`mts_idx` をシグナルしない代わりに、イントラ予測における変換ブロックのサイズや、インター予測において SBT(サブブロック変換)を適用する CU 内のサブブロックの位置に応じて、各水平・垂直方向の変換基底を自動的に指定するモードである。Explicit MTS モードと比較して符号化効率が低下するデメリットがあるが、エンコーダにおける変換基底を探索する処理コストを軽減するメリットがある。**エラー! ブックマークが自己参照を行っています。**にイントラ予測における変換ブロックサイズに対応する変換基底を示す。なお、SBT におけるサブブロック位置に対応する変換基底については、次節で説明する。

表 6-3 変換ブロックサイズに対応する変換基底

| | | |
|------------------|-----------------------|-------------|
| 変換ブロックの横幅 $nTbW$ | $4 \leq nTbW \leq 16$ | $16 < nTbW$ |
| 水平方向の変換基底 | DST7 | DCT2 |
| 変換ブロックの縦長 $nTbH$ | $4 \leq nTbH \leq 16$ | $16 < nTbH$ |
| 垂直方向の変換基底 | DST7 | DCT2 |

6.1.2. SBT (Subblock Transform)

SBT は、インターCU を縦または横方向に、2 つのサブブロックに分割し、一方のサブブロックの予測残差信号をゼロ化し、もう一方のサブブロックの予測残差信号に対してのみ変換処理を行う符号化ツールである。SBT はブロック分割、符号化ブロックフラグ、水平・垂直方向の変換基底を一括して指定することができ、効率よく符号量を削減するメリットがある。

SBT の制御に関する下記(a)~(e)のシンタックスは、インターCU でシグナルされる。

- a) `cu_sbt_flag`: SBT を適用するか否かを示すフラグ
- b) `cu_sbt_horizontal_flag`: 分割方向が縦または横方向かを示すフラグ
- c) `cu_sbt_quad_flag`: 変換処理を適用するサブブロックのサイズが、CU の 1/2 サイズまたは 1/4 サイズかを示すフラグ (2:2 分割、1:3 分割)
- d) `cu_sbt_pos_flag`: 変換処理を適用するサブブロックの位置を示すフラグ

図 6-に SBT の分割方向と変換処理を適用するサブブロックの位置の一例を示す。図 6-では、変換処理を適用するサブブロックのサイズは、CU の 1/2 サイズである(`cu_sbt_quad_flag = 0`)。なお、`cu_sbt_quad_flag = 1` の場合は、変換処理を適用するサブブロックのサイズは CU の 1/4 サイズである。

図 6- (a)は、ブロックを左右に 2 分割し、左側のサブブロックを変換対象とし、右側のサブブロックをゼロ化する例である。変換対象となるサブブロックには、水平方向に DCT8、垂直方向に DST7 が適用

される。

図 6- (b)は、ブロックを左右に 2 分割し、右側のサブブロックを変換対象とし、左側のサブブロックをゼロ化する例である。変換対象となるサブブロックには、水平・垂直方向の両方に DST7 が適用される。

図 6- (c)は、ブロックを上下に 2 分割し、上側のサブブロックを変換対象とし、下側のサブブロックをゼロ化する例である。変換対象となるサブブロックには、水平方向に DST7、垂直方向に DCT8 が適用される。

図 6- (d)は、ブロックを上下に 2 分割し、下側のサブブロックを変換対象とし、上側のサブブロックをゼロ化する例である。変換対象となるサブブロックには、水平・垂直方向の両方に DST7 が適用される。

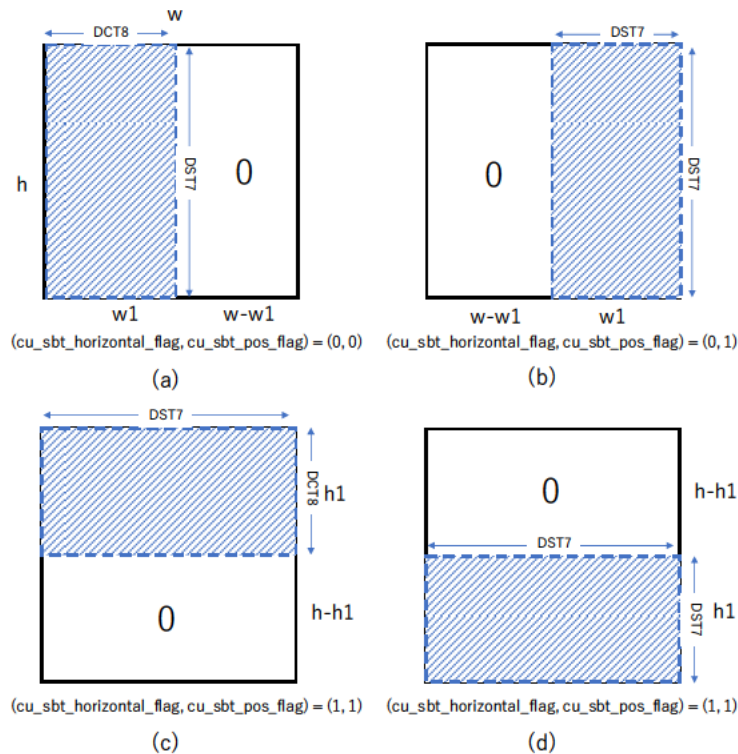


図 6-1 SBT のブロック分割方向と変換対象とするサブブロックの位置、及び対応する変換基底

6.1.3. LFNST (Low Frequency Non-Separable Transform)

VVC では、イントラ CU において、一次変換後の係数を非分離型の直交変換により再変換し、残存する係数分布の偏りをさらに低周波数領域に集中させ、冗長性を軽減させる二次変換が新たに導入されている。この二次変換は LFNST と呼ばれる。

LFNST では、演算量の増加を回避するため、一次変換後の低周波数領域の最大 48 係数までを変換対象とし、再変換後の係数を最大 16 係数まで限定して符号化する。具体的には、短辺のサイズが 8 未満の変換ブロックの場合、一次変換後の低周波数領域にある 16 個の係数を変換対象とし、短辺のサイズが 8 以上の変換ブロックの場合、一次変換後の低周波数領域にある 48 個の係数を変換対象とする。また、横幅 $nTbW$ と縦幅 $nTbH$ の両方が 4 または 8 である変換ブロックでは、再変換後の係数は最大 8 係数に限定され、そのほかの変換ブロックでは、再変換後の係数は最大 16 係数に限定される。

LFNST には、16 行 48 列(16x48)と 16 行 16 列(16x16)の 2 種類の変換行列がある。各変換行列には、学習により予め求められた 8 種の変換行列がそれぞれ定義されている。

VVC では、LFNST は一次変換の水平・垂直方向の変換基底が DCT2 のときのみ利用が制限されている。つまり、MTS と LFNST を同時に使うことを禁止させることで、変換処理の演算量を減らしている。

LFNST は、Dual Tree の場合、輝度信号と色差信号の両方において利用可能である。処理遅延の低減のため、Single Tree の場合、輝度信号にのみ制限している。

図 6-に LFNST による変換処理の概要を示す。変換対象となる係数は、イントラ予測モード番号に応じて定まるスキャン方法により一次元配列化される。次に、一次元配列化された変換対象の係数は、イントラ予測モード番号、変換ブロックサイズおよびシンタックス `lfnst_idx` に応じて定まる変換行列により二次変換され、一次元配列上の再変換後の係数がえられる。一次元配列上の再変換後の係数は、斜め方向スキャン `diagScan` に基づいて、二次元配列上に再配置される。なお、再変換後の係数が再配置される領域外の係数はゼロ化される。

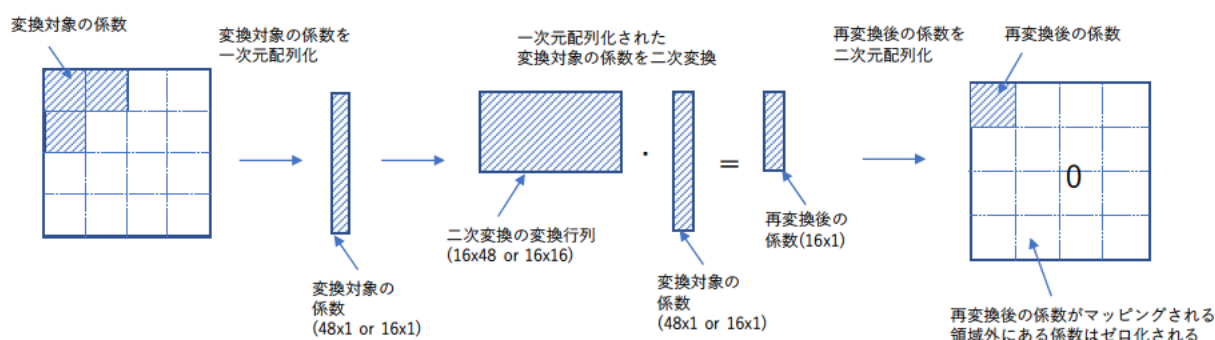


図 6-2 LFNST による変換処理の概要

6.2. 量子化

VVC における量子化に関する技術について述べる。

6.2.1. QP (Quantization parameter)

VVC エンコーダの概略構成(図 2-1)の通り、色差スケール逆変換あるいは整数変換の後に量子化が行われる。基本的な処理方法は HEVC と同じであるが、VVC はより低いビットレートに対応するために量子化パラメータ QP (Quantization parameter)の最大値が 51 から 63 に拡張されている。この拡張に伴い、QP を制御するためのシグナリングが拡張されている。

各スライスの輝度に対する量子化パラメータの初期値は `SliceQpY` で指定される。この `SliceQpY` の最大値が VVC では 63 に拡張されており、PPS で送信される信号 `pps_init_qp_minus26` および PH で送信される `ph_qp_delta` または SH で送信される `sh_qp_delta` によって導出される。`SliceQpY` は次式で求められる。

$$\text{SliceQpY} = 26 + \text{pps_init_qp_minus26} + \text{ph_qp_delta}$$

$$\text{SliceQpY} = 26 + \text{pps_init_qp_minus26} + \text{sh_qp_delta}$$

HEVC と同様に、VVC においても CU レベルで適応的に QP を変更することができる。

色差の QP は輝度の QP から `ChromaQpTable` というマッピングテーブルで変換される。このマッピングテーブルは SPS において、区分線形モデルを用いて指定することができ HEVC のものより細かい制御が可能になっている。図 6-に `Chroma QP` マッピングの例を示す。これは標準化における共通実験で用いられたテーブルをグラフ化したものである。ここで、横軸は輝度の QP であり、縦軸は輝度の QP に対する色差の QP になる。

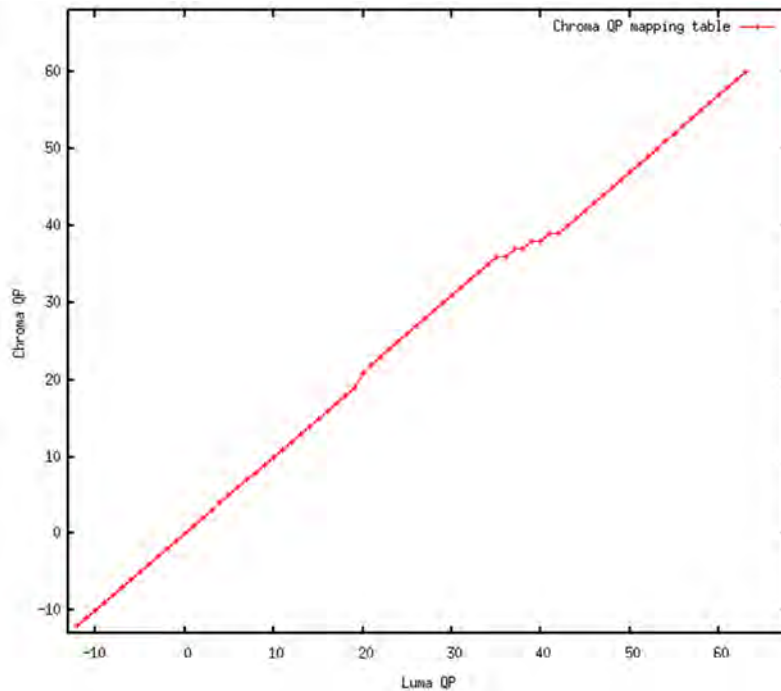


図 6-3 Chroma QP マッピングの例

6.2.2. スケーリング行列

HEVC と同様に、VVC でもスケーリング行列 (QM (Quantization Matrix)、もしくは Quantization Scaling List) を用いて係数毎に量子化値を設定できる。変換の種類増加、および IBC 等の新たな予測モードの追加によりスケーリング行列が拡張されている。また VVC では、デフォルト・モードはすべて量子化値が同じになる平坦なスケーリング行列に変更されている。IBC は、イントラ符号化モードと同じスケーリング行列を使う。このため、定義可能な行列の数は 30 タイプ、DC 値の数は 14 タイプになる。

DC 値は、16x16、32x32、64x64 のスケーリング行列に対して別々に符号化される。

8x8 以下のサイズの TB に対しては、スケーリング行列のすべての要素に対する信号を送ることができる。TB が 8x8 以上のサイズを有する場合、8x8 スケーリング行列の 64 の要素のみ信号が送られて、アップサンプリングすることでより大きいスケーリング行列を再構成して使う。また、シンタックスには正方形のスケーリング行列のみがシグナリングされるため、長方形 TB に対しては長辺が同じ長さとなる正方形のスケーリング行列を短辺方向にダウンサンプリングして生成したスケーリング行列を使用する。

上述の通り、標本点数が 64 の DCT2 変換では 32 以上の係数がゼロになるため、64x64 の TB のスケーリング行列は右下の 4x4 については信号を送信しない。VVC では 2x2 の色差ブロックは SBT を適用するインター予測のみでしか出現しないため、2x2 サイズのイントラ予測色差用スケーリング行列は、デフォルト・モードおよびユーザ定義から取り除かれている。

ユーザ定義のスケーリング行列の信号の送信においては、符号化効率を改善するために、前に符号化されたスケーリング行列の参照、参照スケーリング行列との差分符号化といったシグナリング方法が選択可能となっている。

6.2.3. DQ (Dependent quantization)

VVC は DQ と呼ばれる新しい量子化ツールを採用し、変換係数の圧縮率を向上した。このツールは、処理対象の変換係数と直前に処理した変換係数との間に依存関係を持たせ、それを利用してブロック全体での量子化誤差の最小化を図っている。具体的には、異なる再構成レベルを持つ 2 タイプの

スカラー量子化器(図 6-の Q0 及び Q1)を定義し、係数単位で切り替える。

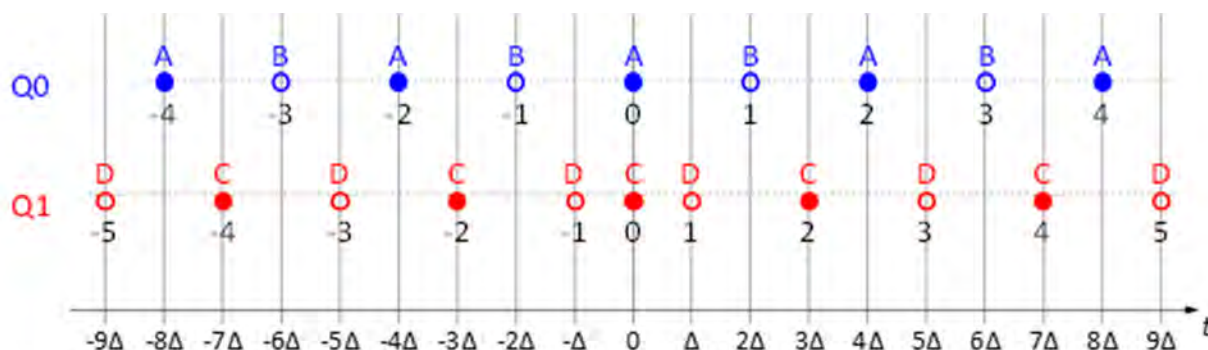


図 6-4 DQ の 2 つのスカラー量子化器

変換係数の復号に用いる量子化器の決定は、状態遷移表(図 6-)に基づく。各係数にて 4 つの状態(0,1,2,3)が定義され、処理対象の変換係数の状態と、変換係数のパリティから、次の処理係数の状態および適用する量子化器(図 6-の Q0 または Q1)を決定する。

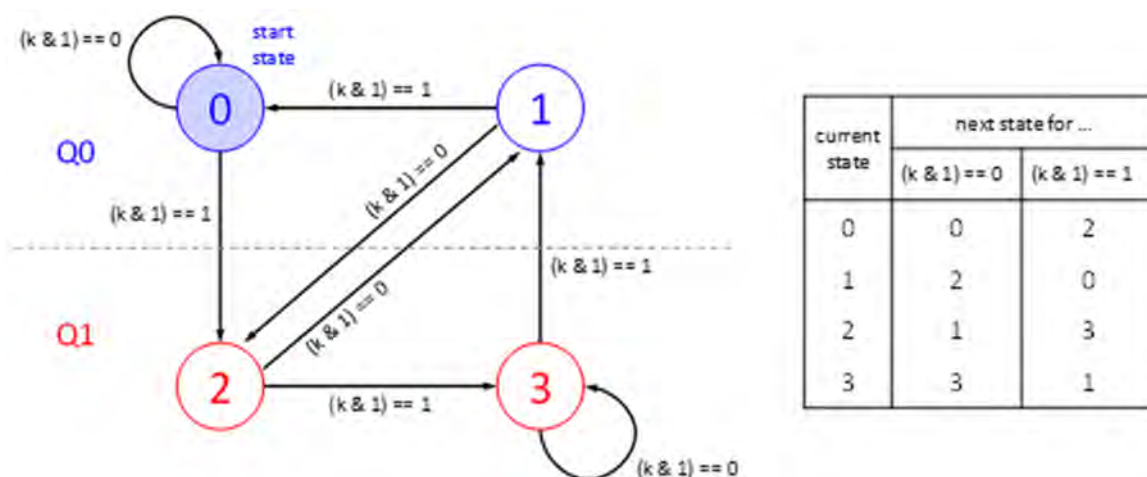


図 6-5 DQ のための状態遷移と量子化器選択

6.3. 変換スキップおよび可逆符号化

HEVC と同様に、VVC においても予測残差信号に対して変換処理を行わずに、画素値に直接量子化を行う変換スキップ(TS; Transform Skip)というモードが導入されている。変換スキップは、エッジを保護するのに有効なモードであり、自然画像ではあまり存在しない急峻なエッジが含まれる CG や PC 画面などのスクリーンコンテンツ(Screen Contents)に効果的である。なお、変換スキップは、輝度、色差共に最大 32x32 画素ブロックまで適用することができる。

HEVC では、可逆符号化をサポートするため、変換処理、量子化およびループ内フィルタ処理を省略する変換量子化バイパスというモードがあった。VVC では、変換量子化バイパスは廃止され、代わりに、非可逆となる符号化ツール(例えば、MTS、LFNST やループ内フィルタ等)を明示的にオフにし、変換スキップと量子化ステップサイズが1以下となる量子化パラメータを用いて変換・量子化する方式に置き換えられた。

図 6-に、HEVC と VVC における逆量子化・逆変換の動作の概略を示す。HEVC では、変換スキップ時、逆変換の代わりにブロックサイズに応じた左ビットシフト(5+log₂nTbS bit 左シフト)、およびビット

深度に応じた右ビットシフト(20-bitDepth bit 右シフト)のスケーリング処理があった(図 6- (a))。VVC では、これら二つのスケーリング処理は逆量子化に吸収され、変換スキップの処理は簡素化されている(図 6- (b))。

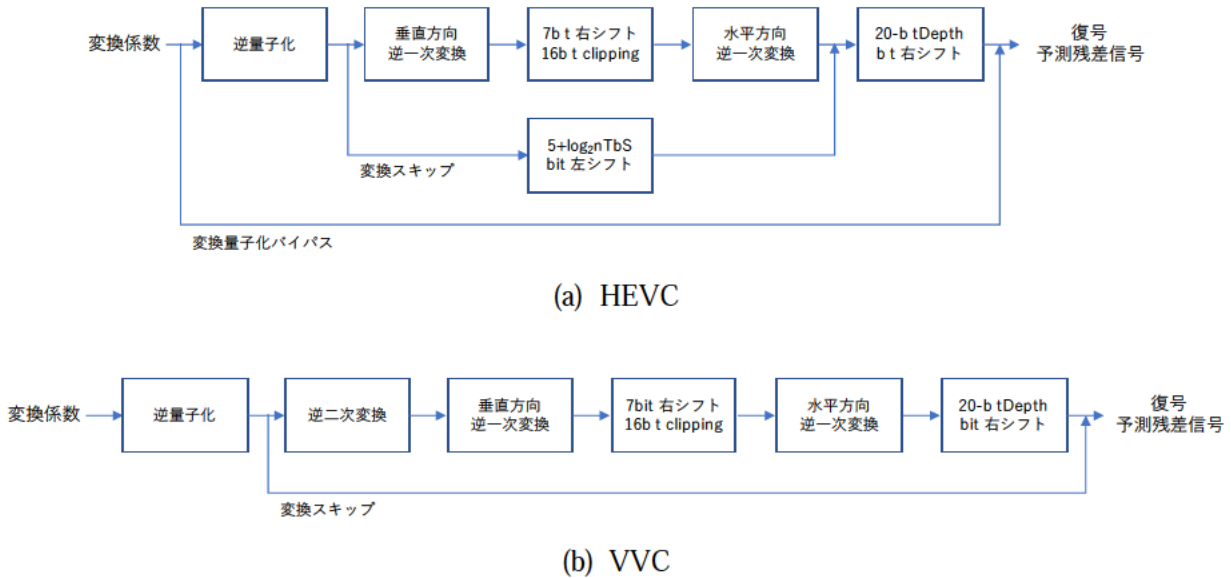


図 6-6 HEVC と VVC における逆量子化・逆変換の動作の概略

7. ループ内フィルタ

ループ内フィルタは、符号化ループ内の量子化処理により発生する符号化ノイズを低減する技術であり、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ギ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。

VVC のループ内フィルタは、HEVC で導入されるデブロッキング・フィルタ(DBF: Deblocking filter) および画素適応オフセット(SAO: Sample Adaptive Offset)に加えて適応ループ・フィルタ(ALF: Adaptive Loop Filter)、輝度マッピング・色差スケーリング(LMCS: Luma Mapping with Chroma Scaling)が導入された。VVC では、LMCS、DBF、SAO、ALF の順でループ内フィルタが適用される。ループ内フィルタはデコーダ全体の処理の中でも比較的複雑な処理であることから、SPS によるシーケンス全体のオン・オフ制御のほか、ピクチャやスライス単位での制御も可能である。

7.1. LMCS (Luma Mapping with Chroma Scaling)

原画像の輝度信号の画素値分布(ヒストグラム)に大きな偏りがある場合には、ほとんど発生しない画素値に割り当てられているコードワードが冗長となる。VVC では、これらの冗長なコードワードを頻出する画素値に割り当てられるコードワードに割り当てなおすことにより、符号化効率を改善する、LMCS (Luma Mapping with Chroma Scaling)が導入されている。LMCS は原画像の画素値分布に基づいた区分線形モデルによる輝度信号のマッピングと、輝度の画素値に応じた色差の残差信号のスケーリングにより構成される。

LMCS では、符号化側で予測処理の前に輝度及び色差の画素値の補正を適用する輝度マッピング・色差スケーリングを行った後に、原画像と異なるダイナミックレンジで予測、変換、量子化を行う。復号側では、デブロッキング・フィルタ処理前に元のダイナミックレンジに復元するように、逆マッピング・色差スケーリングが行われる。図 7-1 に LMCS の処理の概要を示す。

図に示すように、元のダイナミックレンジに復元した状態で復号画像は DPB (Decoded Picture Buffer)

に格納される。DPB に格納された復号画像を用いたインター予測を行う際には、デコーダ側で輝度マッピングを行なわれた上で参照画像として用いられる。

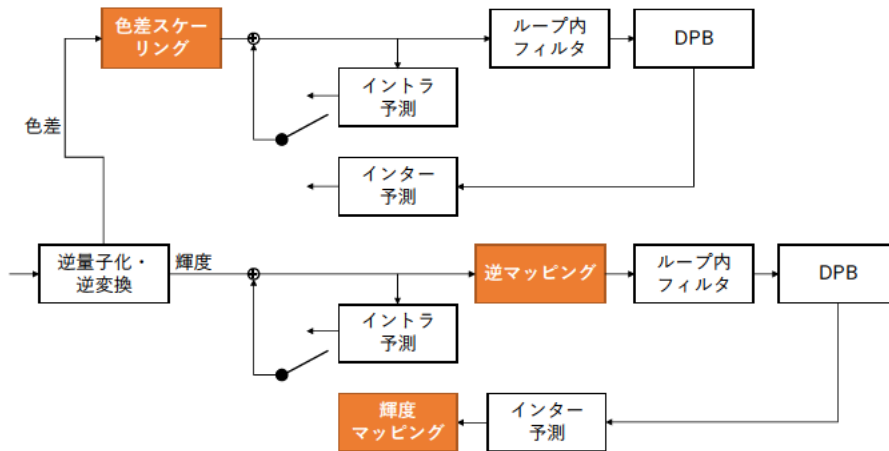


図 7-1 LMCS の概要

7.1.1.1. 輝度マッピング

輝度マッピングは、原画像の輝度信号の取りうる最小値から最大値までを 16 等分した各区分の入力信号値を、区分線形モデルを用いてマッピングし、出力信号を生成する。

デコーダでは、APS にシグナリングされたシンタックス要素を用いて区間線形モデルの各区分における出力信号のコードワード数 $lmcsCW[i]$ を導出する (ただし $i=0 \dots 15$)。導出された $lmcsCW[i]$ をもとに、入力信号および出力信号の各区分の代表点 $InputPivot[i]$ および $LmcsPivot[i]$ 、各区分の輝度マッピング関数及び逆輝度マッピング関数の傾きである $ScaleCoeff[i]$ 、 $InvScaleCoeff[i]$ を導出する。エラー! 参照元が見つかりません。に輝度マッピング関数と逆輝度マッピング関数を示す。ただしエラー! 参照元が見つかりません。の $OrgCW$ は、

$$OrgCW = \frac{1 \ll \text{ビット深度}}{16}$$

であり、輝度信号を表すすべてのコードワード数を 16 等分した各区分の入力信号のコードワード数を意味する。

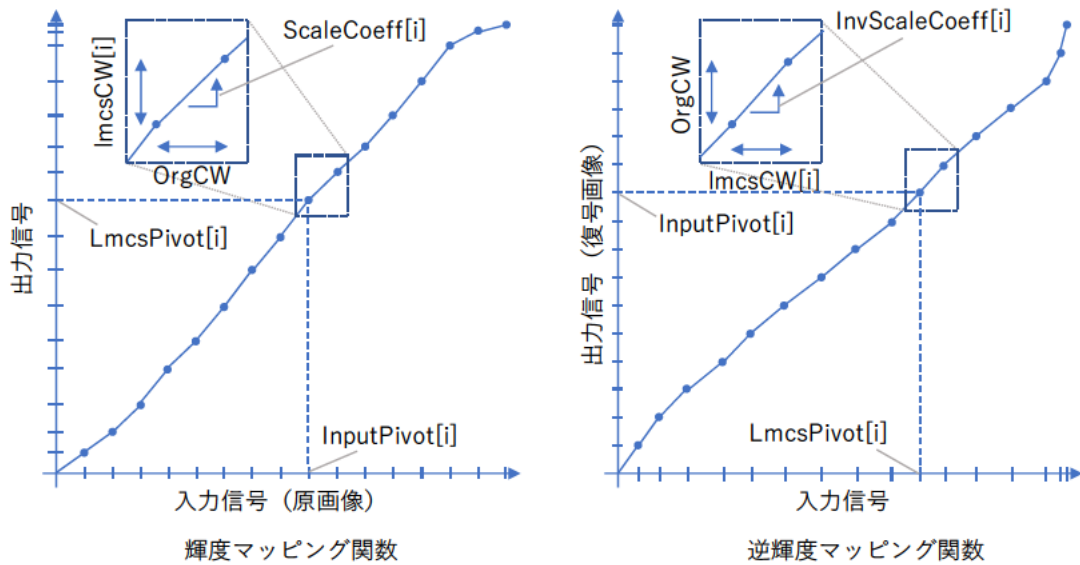


図 7-2 輝度マッピング関数と逆マッピング関数

原信号を Y 、マッピング後の出力信号を Y' としたとき輝度マッピングは下式のように表される。

$$Y' = (\text{LmcsPivot}[\text{idx}Y] + \text{ScaleCoeff}[\text{idx}Y] * (Y - \text{InputPivot}[\text{idx}Y]) + 2^{10}) / 2^{11}$$

ただし、 $\text{idx}Y$ は、 Y が属する輝度マッピング関数の入力信号の区分のインデックスであり $\text{idx}Y = Y / \text{OrgCW}$ によりあらわされる。

同様に、逆輝度マッピングは

$$Y'' = \text{InputPivot}[\text{idx}Y'] + (\text{InvScaleCoeff}[\text{idx}Y'] * (Y' - \text{LmcsPivot}[\text{idx}Y']) + 2^{10}) / 2^{11}$$

である。ただし、 Y'' は逆輝度マッピング後の出力信号であり、 $\text{idx}Y'$ は Y' が属する逆輝度マッピング関数の入力信号の区分のインデックスである。

7.1.2. 色差スケーリング

色差スケーリングは、輝度信号の値に応じたスケーリングを色差の残差信号に対して行う。適応する色差の残差信号のスケーリング係数の導出に同一の輝度信号の復号画像を用いることによる遅延を回避するために、対象の色差ブロックに隣接する復号済みの輝度信号(逆輝度マッピング前)の平均値に応じてスケーリング係数を導出する。

隣接平均輝度に対応する逆輝度マッピング関数の入力信号の区分のインデックスを idx としたとき、色差スケーリングの係数 varScale は、APS にシグナリングされるシンタックス要素により導出される lmcsDeltaCrs を用いて下記のように算出される。

$$\text{varScale} = \text{OrgCW} * \frac{2^{11}}{\text{lmcsCW}[\text{idx}] + \text{lmcsDeltaCrs}}$$

算出された varScale を用いて、下記のように色差の残差信号がスケーリングされたのち、予測画像と合成することで色差の復号信号が得られる。

$$C = C_{\text{pred}} + \text{Sign}(C_{\text{res}}) * \frac{\text{Abs}(C_{\text{res}}) * \text{varScale} + 2^{10}}{2^{11}}$$

ただし C_{pred} および C_{res} は色差信号の予測信号および残差信号である。

7.2. DBF (Deblocking Filter)

予測、変換、量子化処理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・フィルタは、ブロック境界(垂直境界と水平境界)に対し、境界の近傍画素を用いて平滑化処理を行う。HEVCと同様に、任意の位置のブロックの垂直境界に対するフィルタ処理の結果がその直下のブロックの垂直境界のフィルタ処理に影響しないようフィルタ長が設計されているため、ブロックの画面内の全ての垂直境界(もしくは水平境界)の並列処理が可能となっている。また、ブロック単位処理も可能である。

HEVCにおけるCU境界、予測ブロック境界、変換ブロック境界に対するフィルタ処理と同様のコンセプトがVVCのフィルタ処理においても採用されている。VVCでは、HEVCのような階層的な変換ブロックの分割は採用されていないが、SBTやISPを適用したCUでは、CU内部に変換ブロック境界が生じる。また、最大変換ブロックサイズを上回るブロックサイズのCUにおいては暗黙的な変換ブロックの分割が行われるため、同様に変換ブロック境界がCU内に生じる。CUにSbTMVPやアフィンモードが適用された場合には、サブブロックごとの予測が行われるため、CU内にサブブロック境界が存在することになる。VVCではCU境界に加えて、上記の変換ブロック境界、サブブロック境界に対してフィルタ処理が施される。

VVCでは、最大CUサイズが128x128に拡張されており、サイズの大きいブロックの境界での不連続性をより滑らかにするためLTDF(Long Tap Deblocking Filter)が採用されている。ブロックの幅や高さが32以上のブロックに対して、最大で境界から輝度では7画素分、色差では3画素分の画素値をフィルタ処理により更新可能となっており、エッジ分類やブロックサイズにより、境界を跨ぐそれぞれのブロックごとにフィルタ画素数が設定される。また、HEVCでは輝度・色差共に8画素単位の境界にのみフィルタ処理が適用されていたが、VVCでは輝度のみ4画素単位の境界に対してもフィルタ処理が可能となっている。ただし、並列処理を担保するため、ブロックの幅が4以下の場合には、水平フィルタ処理では境界に接する垂直1ラインの値のみが更新される。同様にブロックの高さが4以下の場合には、垂直フィルタ処理では境界に接する水平1ラインの値のみが更新される。

なお、CTUの上側隣接境界のフィルタ処理に必要なCTU外の参照画素については、4画素列(色差については2画素列)までに限定されており、必要なラインメモリ数はHEVC同等となっている。また、他のループ内フィルタ処理と同様に仮想境界(Virtual Boundary)を超えた処理は行われない。

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、境界を跨ぐ2つのブロック(以後、境界の上側もしくは左側に位置するブロックをブロックP、右側もしくは下側に位置するブロックをブロックQと呼ぶ)の予測モードや動きベクトル、変換係数の有無などに応じて規定されるブロック境界強度bS(0~2)が表7-1のように規定される。なお、優先度の高い条件から順に判定し、条件に合致する場合はそれ以下の優先度のチェックは行わない。また、bSの判定は色成分ごとに独立して行う。

表 7-1 境界強度 bS の導出

| 優先度 | 条件 | Y | Cb | Cr |
|-----|-------------------------------|---|----|----|
| 6 | 両方のブロックに対して輝度のBDPCMが適用されている | 0 | 0 | 0 |
| 5 | 少なくともどちらかのブロックの予測モードがイントラ予測もし | 2 | 2 | 2 |

| | | | | |
|---|------------------------------|---|---|---|
| | くは CIIP モードである | | | |
| 4 | 少なくともどちらかのブロックに有意係数が存在する | 1 | 1 | 1 |
| 3 | 2つのブロックの動きベクトルが参照するピクチャが異なる | 1 | - | - |
| 2 | 2つのブロックの動きベクトルの差が0.5画素よりも大きい | 1 | - | - |
| 1 | それ以外 | 0 | 0 | 0 |

算出した bS が 0 となるエッジに対してはフィルタ処理を行わない。 bS が 1 以上のエッジについては、境界付近の勾配およびアクティビティにより算出される dL をもとにエッジ分類 dE (0~3) を算出する。 dE が 0~2 の場合には、HEVC と同様に、フィルタオフ ($dE=0$)、弱フィルタ ($dE=1$)、強フィルタ ($dE=2$) が適用される。 dE が 3 である場合には上記の LTDF が適用される。

bS および量子化パラメータから算出される閾値 tC 、 β の導出方法においては、HEVC と同様であるが、VVC では主な符号化対象となる映像信号のビット深度が 10-bit であることから、 tC を導出するテーブルを構成する値の精度が 10-bit となっている (HEVC では 8-bit)。また、VVC は従来のダイナミックレンジ (SDR) の映像だけでなく、高ダイナミックレンジ (HDR) 映像も対象としており、様々なコンテンツに対して柔軟に DBF をコントロールすることを目的とした、輝度レベルに応じたフィルタ強度の設定を可能とする LADF (Luma Adaptive Deblocking Filter) が導入されている。

7.2.1. LADF (Luma Adaptive Deblocking Filter)

HEVC と同様に、フィルタ強度を制御する閾値 β および tC はブロック P およびブロック Q の量子化パラメータの平均 qp_L をもとに算出される。LADF はこれらの閾値を導出する qp_L に対して、輝度レベルに応じた QP オフセットを加えることで、フィルタ強度を柔軟にコントロール可能とする。

対象のエッジに接するブロック P およびブロック Q の輝度信号 (図 7-3) を用いて輝度レベル LL を下記の式を用いて算出する。

$$LL = (p_{0,0} + p_{0,3} + q_{0,0} + q_{0,3}) \gg 2$$

SPS にシグナリングされた複数の輝度レベル帯における QP オフセットをもとに、算出した LL に応じた $qpOffset$ を導出し、 qp_L を補正する。

$$qp_L = ((Qp_Q + Qp_P + 1) \gg 1) + qpOffset$$

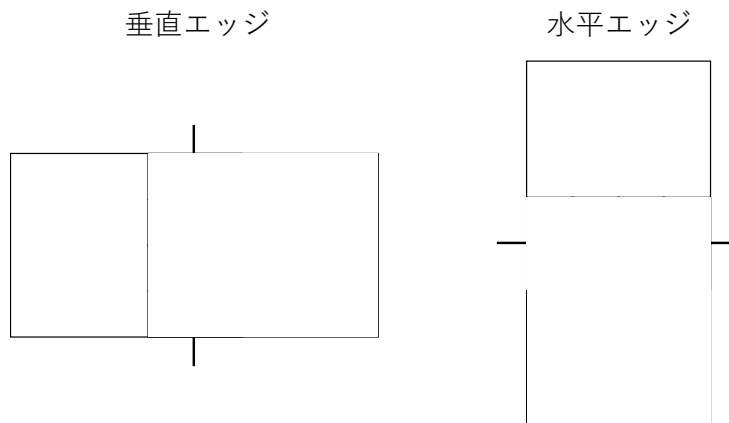


図 7-3 LADF の輝度レベル算出画素

7.2.2. LTDF (Long Tap Deblocking Filter)

LTDF は、従来の HEVC の DBF に比べてフィルタ長の長い Bilinear フィルタである。輝度信号に対しては、ブロック P およびブロック Q のブロックサイズがともに 4 より大きく、また少なくとも一方のブロックサイズが 32 以上であり(垂直エッジの場合は幅、水平エッジの場合は高さ)、かつ境界付近の画素値の変動や勾配が閾値以下である場合に適用される。最大でブロック境界から 7 画素の位置までフィルタ処理可能であり、フィルタ長は境界を挟んだブロックごとに設定される。色差信号に対する LTDF はブロックサイズが 8 以上の信号に対して適用され、最大で 3 画素フィルタ処理可能である。

ブロック P、ブロック Q それぞれのブロックサイズや予測モードに基づいて設定されたブロックごとフィルタ長 (maxFilterLengthP および maxFilterLengthQ) の組み合わせに基づいて、フィルタ処理が行われる。ブロック境界に位置する仮想的な画素 (図 7-4 の middle) を上記の組み合わせに基づいた重み付き平均により算出するとともに、ブロック境界からフィルタ長分離れた画素位置にも同様の仮想的な画素 (図 7-4 の P₇ および Q₃) を近傍 2 画素の平均により算出する。これらの仮想的な画素 2 点の線形補間によって、フィルタ処理後の画素を算出する。

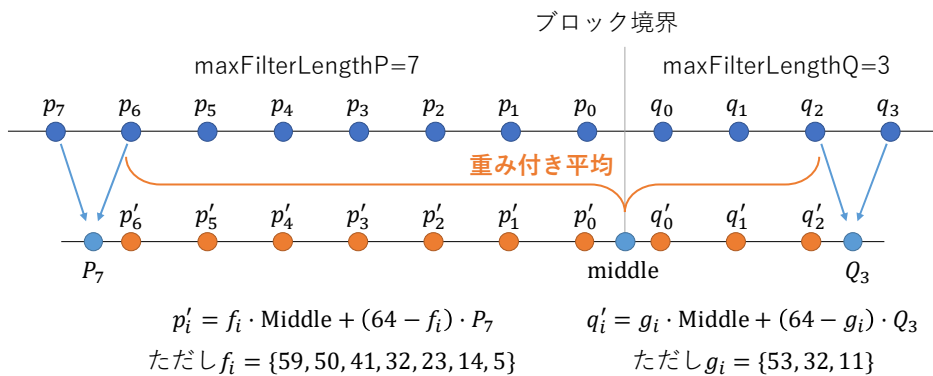


図 7-4 LTDF のフィルタ処理

7.3. SAO (Sample Adaptive Offset)

VVC においても HEVC で導入された SAO (Sample Adaptive Offset) が採用されている。SAO は DBF 適用後の復号画像について画素単位で分類し、分類に応じたオフセットを各画素値に加算することで、復号画像を補正する。HEVC と同様に BO (Band Offset) および EO (Edge Offset) の 2 種類の処理が用意され、CTU 単位で BO、EO、SAO オフのいずれかを輝度および色差に独立して適用可能である。なお、他のループ内フィルタ処理と同様に仮想境界 (Virtual Boundary) を超えた処理は行われない。

7.4. ALF (Adaptive Loop Filter)・CCALF (Cross-Component ALF)

VVC では、新たなループ内フィルタとして輝度および色差に対してそれぞれ独立したフィルタ処理を行う ALF と、色相関に基づく輝度信号を用いた色差信号に対する CCALF (Cross-Component ALF) が導入された。

7.4.1. ALF におけるフィルタ形状

輝度の ALF には 7x7 のダイヤモンドフィルタが、色差の ALF は 5x5 のダイヤモンドフィルタが用いられる。ALF のフィルタ処理では、単純な線形フィルタ処理を行うか、各フィルタ係数の畳み込みの際にクリップ処理を行う非線形フィルタ処理を行うかを選択可能となっている。ただし、クリップ処理に用いるクリップ係数は入力画像のビット深度に応じた 2 のべき乗の値に限られる。各フィルタ係数とクリップ係数を示すフラグは APS にシグナリングされる。

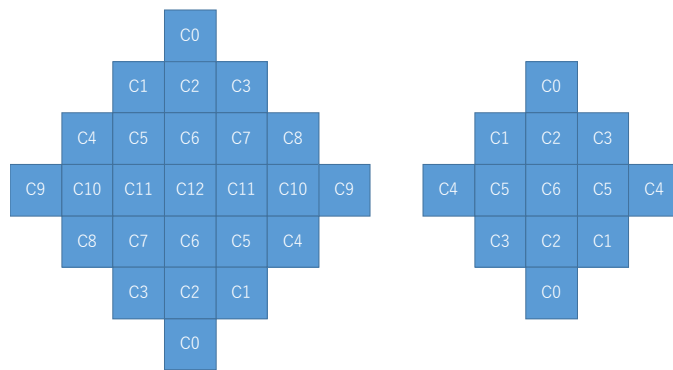


図 7-5 ALF のフィルタ形状

7.4.2. 輝度 ALF におけるクラス分け

輝度の ALF では、画像を 4x4 画素の小領域ごとに方向性およびアクティビティに基づき 25 クラスに分類し、クラスごとにフィルタ処理を切り替えながら適用する。

クラス分けでは、小領域内の画素について水平、垂直、対角、逆対角のそれぞれの方向について、1D ラプラシアンを算出し、それらの比較によりエッジの方向性 D (0~4) およびアクティビティ \hat{A} (0~4) を算出し、対象小領域におけるクラス $filtIdx$ (0~24) を導出する。

$$filtIdx = 5D + \hat{A}$$

クラス分けにおける演算量を低減させるため、1D ラプラシアンは図 7-6 における灰色の画素位置のみとしている。ただし、図中赤枠が対象の小領域である。

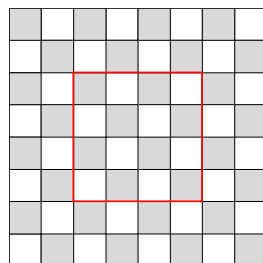


図 7-6 クラス分けに用いる画素位置

7.4.3. フィルタの変形

上記クラス分けにおける対象小領域の垂直・水平・対角・逆対角方向の 1D ラプラシアンにより得られた対象領域におけるエッジ強度の強い方向の組み合わせに応じてダイヤモンドフィルタを転置、垂直反転、90 度回転することで、少ないフィルタセットで多様な方向のテクスチャに対するフィルタ処理 (25 クラスのフィルタセット×4 パターン) が可能となる。

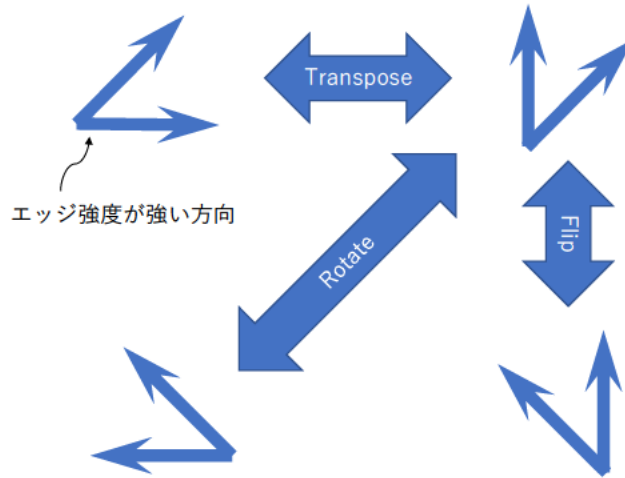


図 7-7 エッジ強度の強い方向の組み合わせに応じたフィルタの変形

7.4.4. フィルタ処理

輝度信号に対するフィルタとしては、既定の 16 種類のフィルタセット(各 25 クラス)を用いるか、APS にシグナリングされたフィルタセットを用いるかをシンタックスにより指定可能である。色差信号に対するフィルタは既定のフィルタセットは用意されていないが、APS に複数の色差用フィルタを格納可能であり、これらのフィルタのうち、どのフィルタを用いるかをシンタックスにより指定可能である。なお、色差信号に対するフィルタ処理においては、輝度で行うようなエッジ強度に応じたクラス分けは行われない。

輝度・色差信号に対するフィルタ処理は、 $R(i, j)$ をフィルタ前の画素値、 $R'(i, j)$ をフィルタ後の画素値としたとき、

$$R'(i, j) = R(i, j) + \left(\left(\sum_{k \neq 0} \sum_{l \neq 0} f(k, l) \cdot K[R(i + k, j + l) - R(i, j), c(k, l)] + 64 \right) \gg 7 \right)$$

となる。ただし、 $f(k, l)$ はフィルタ係数であり、 $K[]$ はクリップ関数、 $c(k, l)$ はクリップ係数である。また、 k および l の範囲は、フィルタ長に応じる。

7.4.5. CC-ALF

CC-ALFは色成分間の相関を用いた復元フィルタの一種であり、対象の色差信号と同位置の輝度の復号画素 8 画素に対するフィルタ処理結果を色差信号に加算する。ALF との並列処理を可能とするため、フィルタ処理には、ALF を適用前の輝度信号を用いる。

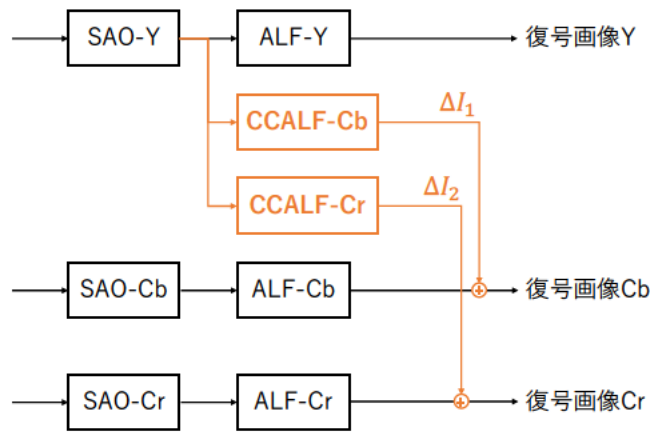


図 7-8 CCALF の処理概要

具体的なフィルタ処理に用いる輝度信号の画素位置は図 7-9 の赤い画素である。これらの輝度信号を用いて、青い画素位置の色差信号を補正する。なお、CC-ALF では Cb と Cr 成分それぞれに対して独立してフィルタのオン・オフ制御が可能である。

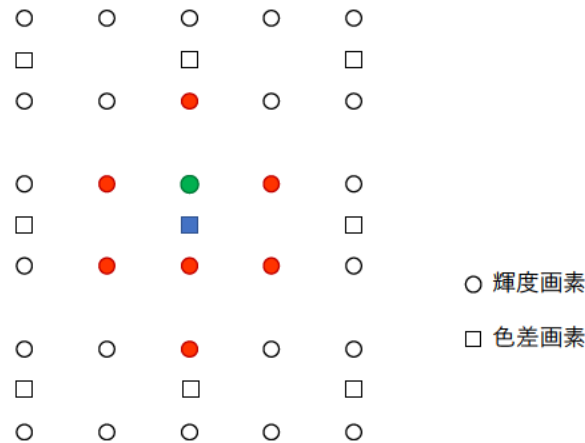


図 7-9 CC-ALF のフィルタ処理に用いる画素位置

図 7-8 に示すように、輝度に対するフィルタ結果を Cb および Cr 信号に加算する。図中の ΔI_i は下記のように算出される。

$$\Delta I_i(x, y) = \sum_{(k, l) \in S_i} c_i(k, l) \cdot (I_L(k, l) - I_L(x', y'))$$

ただし、 S_i は図 7-9 の赤で示す輝度のフィルタサポート範囲であり、 (x, y) は図 7-9 における青で示す対象フィルタ画素位置である。 $I_L(k, l)$ は赤で示す輝度の復号画素の値、 $I_L(x', y')$ は緑で示す輝度の復号画素の値、 c_i はフィルタ係数である。

7.4.6. CTU 境界における処理

ラインメモリ削減のため、CTU 境界より 4 画素(色差については 2 画素)上に ALF 境界を定義し、ALF 境界を跨いだ領域の画素値を用いることなくクラス分けおよびフィルタ処理が行われる。図 7-10 に示すように、ALF 境界に近い小領域においては、オレンジの画素を用いることなく方向性・アクティビティの算出を行う。なお、同様に、対象の小領域が ALF 境界の下側に位置する場合も、ALF 境界の上側の画素を用いることなく処理が行われる。

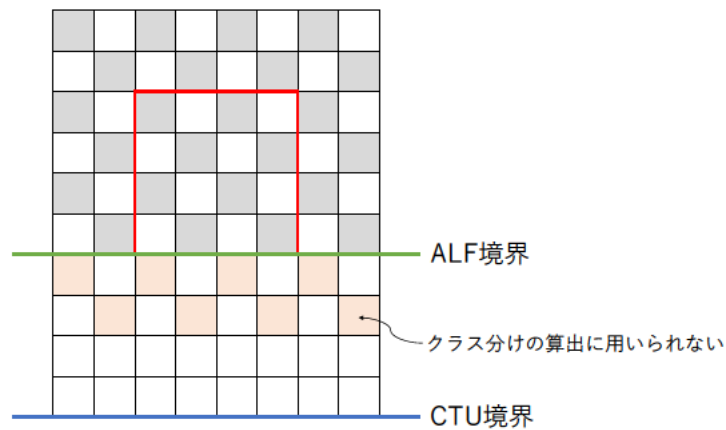


図 7-10 ALF 境界付近におけるクラス分け

また、フィルタ処理においては、ALF 境界を付近の画素における歪みを防ぐために ALF 境界を跨いだ画素だけでなく、ALF 境界を跨いだ画素の原点に対して点対称の位置(同一のフィルタ係数を適用する位置)の画素についてもフィルタ処理に用いないよう規定されており、これらの画素位置については近傍画素のパディングにより補完された上でフィルタ処理が行われる。なお、ALF においても他のループ内フィルタ処理と同様に仮想境界を跨いだ処理は行われない。

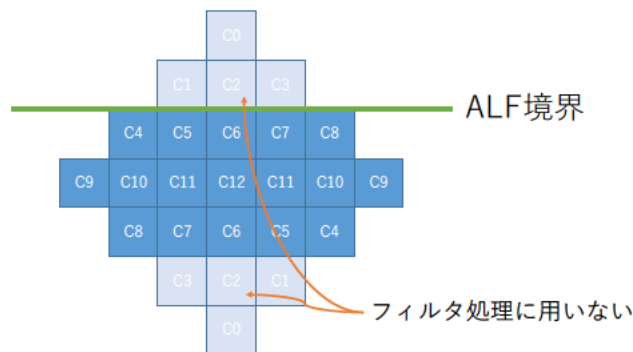


図 7-11 ALF 境界におけるフィルタ処理

8. エントロピー符号化

エントロピー符号化は、HEVC 同様に、スライスヘッダ層以上では、固定長もしくはゴロム符号が用いられ、スライスデータ層以下のブロック単位のシンタックスでは算術符号化方式の CABAC (Context-based Adaptive Binary Arithmetic Coding) が用いられる。なお、HEVC と同様に CABAC 以外の方式には対応していない。また、ブロック単位の残差係数をエントロピー符号化するためのシンタックス変換処理(残差係数符号化)について、符号化効率の向上を図って HEVC から大きく変更されたため、ここで詳しく説明する。

8.1. CABAC 符号化

CABAC は、シンタックス要素を所定の方法により二値化(Binarization)し、二値化ビット列のビット(bin)ごとに、0 もしくは 1 の出現確率を制御するモデル(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、符号化対象シンタックス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各 bin に符号を割り当て、さらに符号結果に応じて出現確率の更新を行う。なお、対象の bin によっては、コンテキストを用いずに出現確率を常に 1/2 に固定して符号化を行う bypass も用いられる。

VVC の CABAC 処理の基本構成は上記の通りほぼ HEVC と同様であるが、コンテキストの切り替え方式の改良に伴い、周辺の参照情報を格納するメモリ等がやや増加している。また、HEVC では各コンテキストの出現確率を 1 つの予測器のみで制御していたが、VVC では、長期的な確率変動に基づいて出現確率を更新する予測器と、短期的な確率変動に基づいて出現確率を更新する予測器との 2 つを用いて、重み付け平均を取ることで処理対象 bin に割り当てる出現確率を導出する。一方で、出現確率の更新量を定義したルックアップテーブルが削除され、単純な演算式で更新量を特定することが可能となっている。

CABAC のコンテキストおよびエンジンの初期化方法は基本的に HEVC と同様であり、スライスおよびタイルの先頭 CTU の処理開始時、もしくは WPP 使用時の各 CTU 列の先頭 CTU の処理開始時に初期化が実施される。なお、3.2.2 節でも説明した通り、WPP における先頭 CTU のコンテキストの初期化処理は、初期値に設定するのではなく上隣接 CTU の処理終了時のコンテキストの値がコピーされる。

8.2. 残差係数符号化

変換・量子化後の残差係数の符号化方法については、VVC では 2 種類の方式が規定されている。変換を適用するブロックについては RRC (Regular Residual Coding)、変換をスキップする TS もしくは BDPCM を用いたブロックについては TSRC (Transform Skip Residual Coding) が使用される。いずれの方式も、HEVC 同様に、輝度もしくは色差の 1 つの変換ブロックを複数の 4x4 係数のサブブロックに分割し、サブブロック単位で係数をスキャンして符号化を行うが、RRC は低周波数領域(ブロックの左上)に向かって値の大きな係数が集中して発生する特性に基づく方式であり、一方、TSRC は係数の発生位置の偏りがなく比較的値の大きな係数が連続した領域に発生する特性に基づく方式である。さらに、いずれの方式も、8.1 節で説明した CABAC のワースト処理量を抑制するために、各変換ブロックにおいてコンテキストを用いた CABAC 処理回数の上限を制限し、制限超過後の係数についてはコンテキストを用いない bypass 処理のみで符号化する仕様となっている。また、色差の 2 つの成分のうち、片方の成分の残差係数のみを符号化し、もう一方の成分の残差係数は他方を変換して流用する JCCR (Joint Coding of Chroma Residual) と呼ばれる方式も新たに導入されている。

8.2.1. RRC (Regular Residual Coding)

RRC は変換を適用したブロックに対して使用される残差係数符号化方式であり、低周波数領域(ブロックの左上)に向かって非ゼロ係数の値が次第に大きくなる傾向を利用して、図 8-1 のように逆スキャン順で 4x4 サブブロックおよびサブブロック内の係数がスキャンされる。また各シンタックスの CABAC 処理におけるコンテキスト切り替え等に参照される周辺情報も、それに合わせて処理対象係数の右下 5 係数位置となっている。

図 8-1 RRC のスキャンおよび周辺参照位置

RRC では、まず、処理対象変換ブロック内で最も右下に位置する非ゼロ係数の位置 (LastSignificantCoeff) を指定し、そこを起点に逆スキャン順で非ゼロ係数を持つ 4x4 サブブロックを指

定して、4x4 サブブロックごとに係数の符号化を行う。

図 8-2 に RRC における 4x4 サブブロックごとの係数符号化の処理フローを記載する。

係数ループ 1 は、0 から 3 までの係数値かどうか、および係数値の奇偶を表現するシンタックスで構成されている。sig_coeff_flag は係数値がゼロか非ゼロかを示し、abs_level_gt1_flag は係数の絶対値が 1 より大きいかどうかを示し、par_level_flag は係数値が奇数か偶数かを示し、abs_level_gt3_flag は係数の絶対値が 3 より大きいかどうかを示す。par_level_flag は VVC で新たに導入された概念であり、0 節で説明した DQ の状態遷移を、係数値を復号することなくシンタックス値から直接制御可能とするために導入された。これらのシンタックスは CABAC のコンテキストを用いて符号化され、係数ごとに使用した回数がカウントされる。使用した回数が閾値以下の場合には係数ループ 1 の次の係数の処理を継続するが、閾値を超過した場合はループを中断する。係数ループ 1 以外は全て bypass 処理で CABAC が行われるため、これにより、処理対象変換ブロックで必要となるコンテキストを用いた CABAC 処理のワースト回数が制限可能となる。なお、閾値は $28 \times$ (処理対象変換ブロック内の 4x4 サブブロック数) となっている。

係数ループ 1 の処理を実施し、かつ係数の絶対値が 3 より大きかった係数については、係数ループ 2-1 によって abs_remainder (係数ループ 1 で表現できなかった残りの係数の絶対値) を符号化する。abs_remainder の符号化方法は HEVC と同様にゴロム・ライス符号が用いられるが、ライス部分のビット長を決定するライスパラメータは、図 8-1 で説明した周辺 5 係数の合計値に従って 4 通りの中から選択される。

係数ループ 1 の処理を実施しなかった係数については、係数ループ 2-2 によって dec_abs_level (係数の絶対値の全体) を符号化する。符号化方法は abs_remainder とほぼ同様でありゴロム・ライス符号を用いて符号化される。ただし、dec_abs_level には係数値 0 も符号化対象として含まれるため、係数値 0 に必要以上に小さなビット長が割り当てられてしまうことを回避するため ZeroPos 処理が適用される。ZeroPos 処理は、図 8-1 で説明した周辺 5 係数の合計値と 0 節で説明した DQ の状態遷移のステートに従って係数値 0 に割り当てられるビット長の切り替えを行う処理である。

最後に係数ループ 3 で、coeff_sign_flag を符号化することで係数の正負を指定する。なお、HEVC と同様に 4x4 サブブロック内の係数絶対値の総和の奇偶を用いて、4x4 サブブロック内の逆スキャン順で最後の非ゼロ係数の正負符号を隠蔽する SDH (Sing Data Hiding) もサポートされている。ただし、0 節で説明した DQ を使用する場合は、SDH は使用禁止となる。


```

for (係数ループ 1) {
    sig_coeff_flag
    abs_level_gt1_flag
    par_level_flag
    abs_level_gt3_flag
    コンテキスト処理回数判定
}
for (係数ループ 2-1) {
    abs_remainder
}
for (係数ループ 2-2) {
    dec_abs_level
    ZeroPos 処理
}
for (係数ループ 3) {
    coeff_sign_flag
}

```

図 8-2 RRC の 4x4 サブブロック係数符号化

8.2.2. TSRC (Transform Skip Residual Coding)

TSRC は変換をスキップする TS もしくは BDPCM を用いたブロックに対して使用される残差係数符号化方式であり、RRC とは異なり非ゼロ係数の発生の偏りがないため、図 8-3 のように通常のスキャン順で 4x4 サブブロックおよびサブブロック内の係数がスキャンされる。また各シンタックスの CABAC 処理におけるコンテキスト切り替え等に参照される周辺情報も、それに合わせて処理対象係数の左上 2 係数位置となっている。

図 8-3 TSRC のスキャンおよび周辺参照位置

TSRC では、LastSignificantCoeff を使用することなく、直接 4x4 サブブロックごとに非ゼロ係数を持つかどうかを指定し、係数を持つ 4x4 サブブロックごとに符号化を行う。

図 8-4 に TSRC における 4x4 サブブロックごとの係数符号化の処理フローを記載する。

係数ループ 1 は、-1 から+1 までの係数値かどうか、および係数値の正負と奇偶を表現するシンタックスで構成されている。sig_coeff_flag は係数値がゼロか非ゼロかを示し、coeff_sign_flag は係数の正負を示し、abs_level_gt1_flag は係数の絶対値が 1 より大きいかどうかを示し、par_level_flag は係数値が奇数か偶数かを示す。TSRC では 0 節で説明した DQ は使用禁止であるが、RRC とシンタックスを揃えるために par_level_flag が使用されている。これらのシンタックスは CABAC のコンテキストを用いて符号化され、RRC と同様に、係数ごとに使用した回数をカウントし、使用した回数に応じて係数ループ

プ 1 を継続するか中断するかを判定する。

さらに係数ループ 2 では、-9 から+9 までの係数値かどうか、`abs_level_gt3_flag` から `abs_level_gt9_flag` のシンタックスによって指定される。係数ループ 2 のシンタックスも CABAC のコンテキストを用いて符号化されるため、係数ループ 1 と同様に、係数ごとに使用した回数をカウントし、使用した回数に応じて係数ループ 2 を継続するか中断するかを判定する。なお、判定に使用する閾値は係数ループ 1 と係数ループ 2 とで共通であり、RRC と同様に $28 \times$ (処理対象変換ブロック内の 4×4 サブブロック数)となっている。

係数ループ 1 の処理のみを実施し、かつ係数の絶対値が 1 より大きかった係数、もしくは、係数ループ 2 の処理まで実施し、かつ係数の絶対値が 9 より大きかった係数については、係数ループ 3-1 によって `abs_remainder` (係数ループ 1 および係数ループ 2 で表現できなかった残りの係数の絶対値) を符号化する。`abs_remainder` の符号化方法は RRC と同様にゴロム・ライス符号が用いられるが、ライス部分のビット長を決定するライスパラメータは、図 8-3 で説明した周辺 2 係数の合計値に従って 4 通りの中から選択される。さらに、ここでは LevelMapping と呼ばれる係数の絶対値の変換処理が行われる。具体的には、図 8-3 で説明した周辺 2 係数の大きい方の値を予測値とし、処理対象係数値と予測値が同じ場合は `abs_remainder` として値 1 に相当する短いビット長に振り替えて符号化する。これにより、変換スキップの特徴である比較的大きな係数値が連続して発生する場合に、符号化効率を向上させることが可能となる。

係数ループ 1 も係数ループ 2 も実施しなかった係数については、係数ループ 3-2 によって `abs_remainder` (ここでは係数の絶対値の全体) と `coeff_sign_flag` を符号化する。`abs_remainder` の符号化方法は係数ループ 3-1 のものと共通である。また、`coeff_sign_flag` は、係数ループ 1 のものとは異なり bypass 処理によって符号化される。なお、TSRC では SDH (Sing Data Hiding)、および 0 節で説明した DQ は共に使用禁止となっている。

```
for(係数ループ 1){
    sig_coeff_flag
    coeff_sign_flag
    abs_level_gt1_flag
    par_level_flag
    コンテキスト処理回数判定
}
for(係数ループ 2){
    abs_level_gt3_flag
    abs_level_gt5_flag
    abs_level_gt7_flag
    abs_level_gt9_flag
    コンテキスト処理回数判定
}
for(係数ループ 3-1){
    abs_remainder
    LevelMapping 処理
}
for(係数ループ 3-2){
    abs_remainder
    coeff_sign_flag
}
```

図 8-4 TSRC の 4×4 サブブロック係数符号化

なお、TS もしくは BDPCM を用いたブロックでもスライスヘッダに記述するフラグによって TSRC ではなく RRC を使用するように切り替えることも可能となっている。例えば、6.3 節で説明した可逆符号化(ロスレス符号化)や超高ビットレートでの符号化時には、非ゼロ係数の個数および係数値が非常に大きくなり、変換スキップの特性に基づいた TSRC が十分な性能を出せず、RRC の方が高性能となる場合がある。そのような条件において TSRC を RRC に切り替えて使用することで符号化効率の改善を図ることが可能となる。

8.2.3. JCCR (Joint Coding of Chroma Residual)

JCCR は色差の 2 つの成分のうち、片方の成分の残差係数のみを符号化し、もう一方の成分の残差係数は他方を変換して流用する残差係数符号化方法である。JCCR には、通常の Cb、Cr の残差係数を個別に符号化する方法に加えて 3 種類のモードがある。3 種類のモードのいずれを使用するかは、Cb および Cr の残差係数を符号化するかどうかを示す coded_flag の組合せによって指定される。表 8-1 にそれぞれのモードにおける残差係数の生成方法をまとめる。なお、表中の SignFlag(正負値)はピクチャヘッダに記述される信号である。また、6.2.1 節で説明した、輝度 QP から色差 QP への変換に用いる ChromaQpTable は、通常の Cb 用、Cr 用とは別に、JCCR 用のテーブルを独立して指定することが可能となっている。

表 8-1 JCCR モード

| cb_coded_flag | cr_coded_flag | 残差係数生成方法 |
|---------------|---------------|--|
| 1 | 0 | Cb の残差係数 = シンタックスで符号化 Cr の残差係数 = (SignFlag × Cb の残差係数) >> 1 |
| 1 | 1 | Cb の残差係数 = シンタックスで符号化 Cr の残差係数 = SignFlag × Cb の残差係数 |
| 0 | 1 | Cr の残差係数 = シンタックスで符号化 Cb の残差係数 = (SignFlag × Cr の残差係数) >> 1 |

9. スクリーンコンテンツ符号化ツール

スクリーンコンテンツ符号化ツールは、パソコンのデスクトップ画面やゲームやアニメの CG 映像などのスクリーンコンテンツの符号化効率向上に特化したツールであり、IBC (Intra Block Copy)、BDPCM (Block Differential Pulse Coded Modulation)、Palette mode、ACT (Adaptive Color Transform)、TS (Transform Skip)が挙げられる。いずれも、HEVC 第 2 版の SCC 拡張プロファイルとして導入されていたツールであるが、VVC の処理方式に合わせて処理の簡略化および最適化が行われている。なお、Palette mode および ACT は、Main10 4:4:4 プロファイル群のみで使用可能なツールであるためここでは概要のみを説明する。また、TS についてはスクリーンコンテンツに限らず基本的な符号化ツールとして使用されるためここでは説明を省略する(6.3 節参照)。

9.1. IBC (Intra Block Copy)

IBC は、イントラ予測の特別なモードであり、輝度および色差のブロックに対し、処理対象ブロックと同じピクチャの処理済み周辺領域からブロック単位で参照画像をコピーして予測画像を生成するモードである。例えば、同じようなテクスチャパターンが繰り返されるようなグラフィック画像において高い符号化効率を得ることができる。IBC は、4x4 から 64x64 までの CU サイズで使用可能であり、CU 単位で BV (Block Vector)を指定することで参照画像を決定する。

BV はインター予測における AMVP モードやマージモードと同様に、処理済みブロックの BV から予

測ベクトル候補リストを作成し、AMVP モードでは選択した予測ベクトルに差分ベクトルを加算することで BV を決定し、マージモードでは選択した予測ベクトルをそのまま BV として用いる。予測ベクトル候補リストの生成方法はインター予測に比べて大幅に簡略化されており、Spatial BV、History-based BV、Zero BV のみで構成される。また、BV は輝度ブロックも色差ブロックも常に整数画素精度に丸められるため、フィルタ処理なしで参照画像の画素がそのままコピーされる。

IBC で参照可能な周辺領域は、図 9-1 に示すように、処理対象 CU が属する 64x64 領域(VPDU)から符号化順で 3 つ前までの 64x64 領域の範囲に限定される。つまり、4 つ分の 64x64 領域を記憶するメモリを用いて 64x64 領域単位の FIFO 処理で管理することが可能であり、インター予測のように外部メモリにアクセスする必要がなく、通常のイントラ予測の周辺参照画素と同様にハードウェアロジックの内部メモリとして実装することができる。そのため、メモリアクセスのレイテンシ課題やフィルタ処理のための処理遅延が発生しないため、通常のイントラ予測処理と同じパイプラインステージに配置することが可能である。

図 9-1 IBC で参照可能な周辺領域

9.2. BDPCM (Block Differential Pulse Coded Modulation)

BDPCM は、イントラ予測の特別なモードであり、輝度および色差のブロックに対し、水平もしくは垂直方向のイントラ予測を行い、生成された予測差分信号を、変換処理を行わずに量子化し、水平もしくは垂直方向に隣接する量子化後の残差係数の差分値を符号化するモードである。例えば、水平もしくは垂直のどちらか一方に極端に相関の高い画素構成を持つようなブロックなどにおいて高い符号化効率が期待できる。

水平もしくは垂直のどちらの方向を用いるかを示すフラグは、CU 毎に輝度ブロックおよび色差ブロックでそれぞれ独立して符号化される。また BDPCM を選択した CU では、TS かどうかを示すフラグを符号化することなく強制的に TS として扱い、残差係数(正確には残差係数の差分値)の符号化方式は TS と同じ TSRC 方式(8.2.2 節参照)が適用される。また DBF 処理において、処理対象境界の両側が BDPCM であった場合は DBF の適用がキャンセルされる。

9.3. Palette mode

Palette mode は、Main10 4:4:4 プロファイル群のみで使用可能なツールであり、図 9-2 のように、表現可能な画素値の中から幾つかの代表値のみをパレットとして持たせ、その中から処理対象画素に最も近い代表値を選択し、選択した代表値のパレットインデックスを用いて符号化するツールである。パレットに適切な代表値が存在しない場合は、Escape モードを選択して量子化した処理対象画素の値を直接符号化する。また、ブロックの処理完了時にパレットの更新が行われ、常に選択される可能性の高い

代表値がパレットに登録されている仕組みとなっている。これにより、特定の画素値のみが使用されているグラフィック画像などにおいて高い符号化効率を実現することが可能となる。

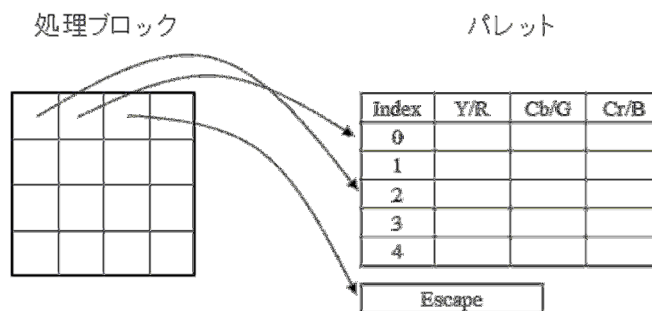


図 9-2 Palette mode の概要

9.4. ACT (Adaptive Color Transform)

ACT は、Main10 4:4:4 プロファイル群のみで使用可能なツールであり、4:4:4 フォーマット(例えば RGB)の符号化時に、予測処理後の予測差分信号に対して下記の変換式を用いて RGB カラー空間から YCgCo カラー空間に変換し、変換された予測差分信号に対して変換処理および量子化処理を行うツールである。例えば、予測差分信号が輝度成分のみに集中する可能性の高いブロックなどで符号化効率の向上を図ることができる。なお、復号時には逆量子化処理および逆変換処理を行った残差信号に対して逆の変換処理を行って RGB カラー空間の予測差分信号を生成する。なお、ACT はスクリーンコンテンツ向けに特化された符号化ツールではないため、YCgCo カラー空間に適した画像であれば通常コンテンツの符号化においても使用される。

$$\begin{aligned} Co &= R - B; \\ t &= B + (Co \gg 1); \\ Cg &= G - t; \\ Y &= t + (Cg \gg 1); \end{aligned}$$

10. プロファイルとレベル

VVC は、HEVC や AVC と同様に、モバイル配信から放送、素材蓄積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。

VVC の第1版が対応するプロファイルは 6 種類定義されている。映像フォーマットは、コンシューマ向け(4:2:0 色差フォーマット、10-bit 画素値精度)とスタジオ業務向け(4:4:4 色差フォーマット、10-bit 画素値精度)である。動画像に対しては、これらに対してシングルレイヤとマルチレイヤ符号化用の符号化ツールが組み合わせられ、合計 4 種類のプロファイルが定義されており、加えて、静止画用のプロファイルが 2 種類定義されている。また、13 種類のレベルを定義している。

10.1. プロファイル

表 10-1 VVC 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 10 | 標準的なツール群で構成される | (ハイレベル) |

| プロファイル名 | 概要・想定用途 | 主な仕様 |
|---------|---|--|
| | <p>代表的なプロファイル</p> <p>想定用途: 放送、蓄積メディア、配信など</p> | <ul style="list-style-type: none"> ● 4:2:0 およびモノクロクロマフォーマット ● 8-bit から 10-bit 画素サポート ● I/P/B スライス ● 時間方向階層符号化 ● Subpicture ● Tile ● Slice ● WPP ● GDR ● Scaling window ● RPR ● VB <p>(ブロック分割)</p> <ul style="list-style-type: none"> ● 4x4~128x128 CU ● 4分木、3分木、2分木 ● Dual tree ● Local dual tree <p>(イントラ予測)</p> <ul style="list-style-type: none"> ● 輝度イントラ予測(DC 予測, Planar 予測, Angle 予測/WideAngle 予測(93 方向)) ● 色差イントラ予測(DC 予測, Planar 予測, 水平予測, 垂直予測, DM 予測) ● MRL ● CCLM ● PDPC ● MIP ● ISP <p>(インター予測)</p> <ul style="list-style-type: none"> ● MV: 輝度 1/16、色差 1/32 ● MC: 輝度 8-tap、色差 4-tap ● AMVP モード(AMVP, AMVR, SMVD) ● スキップモード ● マージモード(Regular Merge, MMVD, CIIP, GPM, DMVR) ● サブブロックモード(アフィンモード, PROF, SbTMVP) ● 動き補償(MC, SIF, Wrap around MC, RPR) ● 予測画像生成(BDOF, BCW, WP) <p>(変換量子化)</p> <ul style="list-style-type: none"> ● TB : 2x2~64x64, 正方形、長方形 ● 変換(MTS(DCT2, DST7, DCT8), SBT, |

| プロファイル名 | 概要・想定用途 | 主な仕様 |
|-----------------------------|---|--|
| | | LFNST) ● 量子化(QP, QM, DQ) ● TS (ループ内フィルタ) ● LMCS(輝度マッピング, 色差スケーリング) ● DBF(LADF, LTDF) ● SAO ● ALF, CC-ALF (エントロピー符号化) ● CABAC ● 残差係数符号化(RRC, TSRC, JCCR) (スクリーンコンテンツ用) ● IBC ● BDPCM |
| Main 10 Still Picture | Main 10 プロファイルに対し、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 10 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 |
| Multilayer Main 10 | Main 10 プロファイルに対し、スケーラビリティ機能を追加したプロファイル 想定用途: 広範な民生用途 | Main 10 プロファイルに以下に示すスケーラビリティに関するツール群を追加したもの ● 空間 ● SN ● マルチビュー |
| Main 10 4:4:4 | Main 10 プロファイルに対し、クロマフォーマットを 4:4:4 まで引き上げたプロファイル 想定用途: スタジオ業務(素材伝送、素材蓄積)用途 | Main 10 プロファイルに、4:4:4 および 4:2:2 フォーマットのサポートを追加したもの (スクリーンコンテンツ用) ● Palette mode ● ACT |
| Main 10 4:4:4 Still Picture | Main 10 4:4:4 プロファイルに対し、静止画像に特化したプロファイル 想定用途: デジタルカメラなど | Main 10 4:4:4 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 |
| Multilayer Main 10 4:4:4 | Main 10 4:4:4 プロファイルに対し、スケーラビリティ機能を追加したプロファイル 想定用途: 広範な業務用途 | Main 10 4:4:4 プロファイルに以下に示すスケーラビリティに関するツール群を追加したもの ● 空間 ● SN ● マルチビュー |

10.2. レベル

VVC 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 10-2 VVC 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 (*1) | | 最大 CPB サイズ (*2) | | ピクチャ内最大スライス数 | 最大タイル数 | 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 | | | | Main tier | High tier |
| 1 | 36 864 | 552 960 | 128 | - | 350 | - | 16 | 1 | 1 | 2 | 2 |
| 2 | 122 880 | 3 686 400 | 1 500 | - | 1 500 | - | 16 | 1 | 1 | 2 | 2 |
| 2.1 | 245 760 | 7 372 800 | 3 000 | - | 3 000 | - | 20 | 1 | 1 | 2 | 2 |
| 3 | 552 960 | 16 588 800 | 6 000 | - | 6 000 | - | 30 | 4 | 2 | 2 | 2 |
| 3.1 | 983 040 | 33 177 600 | 10 000 | - | 10 000 | - | 40 | 9 | 3 | 2 | 2 |
| 4 | 2 228 224 | 66 846 720 | 12 000 | 30 000 | 12 000 | 30 000 | 75 | 25 | 5 | 4 | 4 |
| 4.1 | 2 228 224 | 133 693 440 | 20 000 | 50 000 | 20 000 | 50 000 | 75 | 25 | 5 | 4 | 4 |
| 5 | 8 912 896 | 267 386 880 | 25 000 | 100 000 | 25 000 | 100 000 | 200 | 110 | 10 | 6 | 4 |
| 5.1 | 8 912 896 | 534 773 760 | 40 000 | 160 000 | 40 000 | 160 000 | 200 | 110 | 10 | 8 | 4 |
| 5.2 | 8 912 896 | 1 069 547 520 | 60 000 | 240 000 | 60 000 | 240 000 | 200 | 110 | 10 | 8 | 4 |
| 6 | 35 651 584 | 1 069 547 520 | 60 000 | 240 000 | 60 000 | 240 000 | 600 | 440 | 20 | 8 | 4 |
| 6.1 | 35 651 584 | 2 139 095 040 | 120 000 | 480 000 | 120 000 | 480 000 | 600 | 440 | 20 | 8 | 4 |
| 6.2 | 35 651 584 | 4 278 190 080 | 240 000 | 800 000 | 240 000 | 800 000 | 600 | 440 | 20 | 8 | 4 |

表 10-3 CpbNalFactor と MinCrScaleFactor の値

| プロファイル | CpbNalFactor | MinCrScaleFactor |
|--|--------------|------------------|
| Main 10, Main 10 Still Picture, Multilayer Main 10 | 1 100 | 1.00 |
| Main 10 4:4:4, Main 10 4:4:4 Still Picture, Multilayer Main 10 4:4:4 | 2 750 | 0.75 |

(*1) NAL レイヤでの最大ビットレート(bps)は

$$(\text{最大ビットレート}) \times \text{CpbNalFactor}$$

である。

(*2) NAL レイヤでの最大 CPB サイズ(bit)は

$$(\text{最大 CBP サイズ}) \times \text{CpbNalFactor}$$

である。

(*3) 最大圧縮率は

$$(\text{最大圧縮率}) \times \text{MinCrScaleFactor}$$

である。

表 10-4 最大ピクチャレートと代表的な画像フォーマット (Main tier)

| 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 | 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 | 3616 | 1536 | 5 603 328 | - | - | - | - | - | - | - | 47.7 | 95.4 | 190.8 | 190.8 | 300.0 | 300.0 |
| 3672x1536 | 3680 | 1536 | 5 701 632 | - | - | - | - | - | - | - | 46.8 | 93.7 | 187.5 | 187.5 | 300.0 | 300.0 |
| 3840x2160 | 3840 | 2160 | 8 355 840 | - | - | - | - | - | - | - | 32.0 | 64.0 | 128.0 | 128.0 | 256.0 | 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 | 4096 | 2304 | 9 437 184 | - | - | - | - | - | - | - | - | - | - | 113.3 | 226.6 | 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 |

High tier では最大ピクチャレートが 960 となるため、表 10-4 で最大ピクチャレートが 300 となっている箇所については、表 10-2 の輝度画素数最大値(秒当たり)と輝度サイズに基づいて最大ピクチャレートを算出する。

11. 参考文献

- [1] ITU-T Recommendation H.266: Versatile video coding (07/2020).
- [2] ITU-T Recommendation H.274: Supplemental enhancement information messages for coded video bitstreams (07/2020).
- [3] J. Chen, Y. Ye, S. Kim, “Algorithm description for Versatile Video Coding and Test Model 10 (VTM 10)”, JVET-S2002, JVET 19th Meeting: by teleconference, 22 June – 1 July 2020.
- [4] G.J.Sullivan, et.al., “General Video Coding Technology in Responses to the Joint Call for Proposals on Video Compression With Capability Beyond HEVC”, IEEE Transaction on Circuits and Systems for Video Technology, vol. 30, issue. 5, May 2020.
- [5] 高村誠之, “ISO/IEC JTC 1/SC 29/WG 11 (MPEG)における次世代映像符号化標準化動向”, NTT ジャーナル, 2019.4.
- [6] 市ヶ谷敦郎, “次世代映像符号化方式の標準化動向”, NHK 技研 R&D, No.177, 2019.9.
- [7] ITU-T Recommendation H.265: High Efficiency Video Coding (11/2019).
- [8] 大久保榮, 鈴木輝彦, 高村誠之, 中條健ほか, “H.265/HEVC 教科書”, インプレスジャパン, 2013.10.18.