



映像符号化の標準化動向

日本放送協会 放送技術研究所 テレビ方式研究部 副部長 **岩村 俊輔** いむら しゅんすけ

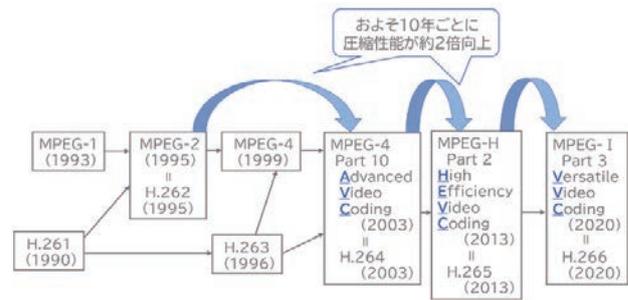


1. 映像符号化標準化の歴史及びプロセス

映像符号化の国際標準化は、通信・放送・ストレージにおける広範な相互接続性と品質の再現性を担保する制度的枠組みとして、ITU-T SG21 WP3 Q.6とISO/IEC JTC 1/SC 29が中核となって推進している。両者は、重複する技術領域において競合ではなく協調を選択し、合同チームを編成して規格策定を行う運用モデルを確立している。これらの国際標準化機関は、プロジェクトごとにJVET (Joint Video Team)、JCT-VC (Joint Collaborative Team on Video Coding)、JVET (Joint Video Experts Team) を組織し、要求条件の抽出、技術の評価・選定、技術統合、勧告・規格文書化の各段階を分担しながら一体的に進める体制をとる。ITU側ではSG21において符号化関連の課題 (Question) としてVCEGを配置し、ISO/IEC側ではJTC 1/SC 29のWG群、とりわけWG5がITU-T SG21と連携するJoint Video Coding Teamsとして機能する。これにより、通信系・メディア系の要件を横断的に扱い、国際的な合意形成を促進している (図1)。

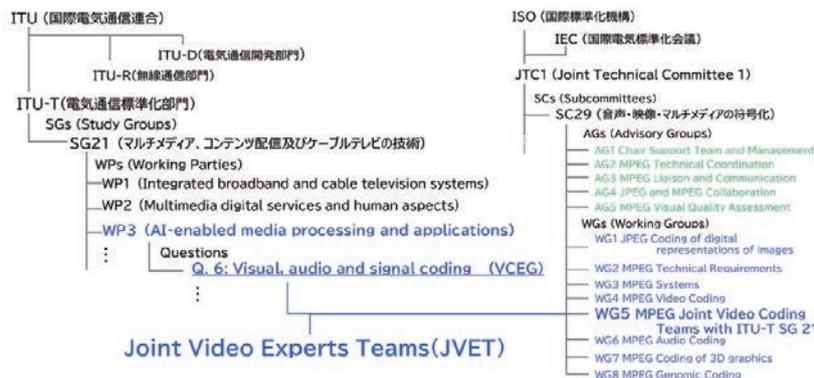
方式の系譜は、H.261 (1990) に始まり、MPEG-1 (1993)、MPEG-2/H.262 (1995)、MPEG-4/H.263 (1996) と進化し、2001年に始動したJVETの下でH.264/MPEG-4 AVC (2003) が策定された。AVCは、符号化効率と実装容易性、プロフィール設計のバランスに優れ、放送・通信・ストレージの現場で広く採用され、産業利用の成熟度を主導した。続いて2010年に発足したJCT-VCがH.265/MPEG-H Part 2 HEVC (2013) を確定し、レンジ拡張、スケーラブル、マルチビュー、

3D、スクリーンコンテンツなどの拡張群 (2014~2016) を整備することで、UHD配信やBS4K/8K等の高精細サービスに適用可能な枠組みを拡張した。さらに2015年、HEVCを凌駕する次世代技術探索を目的にJVET (当初はExploration Team) が始動し、標準化検討の本格化に伴い2017年にExperts Teamへ改称、2020年にH.266/MPEG-I VVCを完了し、同年にH.274/VSEIも整備した。これらの世代的更新は、経験則として「約10年周期で圧縮性能がほぼ2倍向上」するという長期トレンドを形成している (図2)。



■図2. ISO/IEC規格MPEGとITU-T勧告H.26Xの変遷

標準化の進め方は、①要求条件検討、②エビデンス募集 (CfE)、③技術提案募集 (CfP)、④仕様統合・規格文書化という段階的プロセスである。まず要求条件検討では、既存標準が充足し得ないユースケース (超高解像度・高フレームレート・広色域、低遅延・双方向性、スクリーンコンテンツ、可変解像度運用等) を抽出し、画質目標、帯域・遅延制約、複雑度上限、ロバスト性等を含む技術要件を文



■図1. 映像符号化の規格化に係る組織図



書化する。次にCfEでは、標準化の正当性を示す客観的証拠を募る。具体的には、共通テスト条件下で既存標準方式からのビットレート削減率、主観・客観画質改善、ロバスト性、運用適合性を提示させ、技術探索段階の境界条件を整理する。続くCfPでは、要件に即した具体的アルゴリズム群を広く募集し、性能・複雑度・相互作用・実装可能性を評価して採否を決定する。JVETのVVC策定においては、2015年の技術探索開始、2017年～2018年のCfE/CfPを経て、2020年の標準化完了に至るタイムラインが運用された(図3)。



■図3. VVC標準化のタイムライン

会合における採否判断は、テストモデルへの実装と共通テストシーケンス・設定に基づく計測によって支えられる。評価項目は、符号化効率(ビットレート削減率)、画質(主観評価を含む)、処理量(演算時間比、基本演算回数、メモリバンド幅等)であり、性能向上と実装コストのトレードオフを多面的に検討する。類似提案間の比較・相互作用の分析を含め、総合性能・仕様整合性を確認した上で段階的にコア仕様へ採用するのが通例である。

2. ハイブリッドコーディングの要素技術と設計思想

現行の国際標準は、予測・変換・量子化・ループフィルタを統合したハイブリッドコーディングを基盤として設計されている。符号化は、画面をブロック単位に分割し、イントラ(空間)予測とインター(時間)予測によって残差エネルギーを削減し、直交変換・量子化・エントロピー符号化によって符号量を最小化する。復号側では逆量子化・逆変換・ループフィルタにより画質を復号し、次フレーム予測の精度向上に資する。

領域分割は世代ごとに柔軟性を増してきた。AVCでは固定サイズ中心、HEVCではCTU(最大64×64)を四分木分割で可変し、VVCでは最大128×128とし、四分木・二分木・三分木を組み合わせるMTT(Multi Type Tree)構造に加えて輝度・色差で異なる分割木を選択可能にし、絵柄に応じたきめ細かな適応を可能にした。これにより、平坦領域では大きいブロックを、複雑領域では小さく絵柄に適したブロック形状を選択し、予測・変換の効率を高める。

イントラ予測は、AVCの9モード(方向予測8モード+平均値予測)から、HEVCの35モード(方向予測33モード+平均値予測+平面予測)、VVCの67モード(方向予測65モード+平均値予測+平面予測)へと拡充した。さらにVVCでは、色成分間の相関を用いるCCLM(Cross-Component Linear Model)や、参照画素の線形写像により予測画像を生成するMIP(Matrix-based Intra Prediction)を導入し、テクスチャ構造の再現性が向上している。

インター予測は、HEVCで採用されたAMVP(周辺の動きベクトルを基準として差分ベクトルを伝送)とMerge(隣接領域と同一の動きの場合に動きベクトル情報の伝送を大幅に省略)により情報量を抑制する枠組みを基礎としつつ、VVCでは大域的変形を少ないパラメータで表現できるアフィン予測、復号側で動きベクトルを微補正して局所の非線形動きを再現するデコーダ側補正、インターとイントラを合成するCIIP(Combined Inter/Intra Prediction)など、高精度化のための拡張を採用した。これにより、拡大・回転・せん断などの複雑運動を、ビットレートの増加を抑えつつ表現可能である。

VVCの直交変換は、HEVCで採用された正方ブロックのみ(整数DCT-II/DSTV-II、最大32×32)から、非正方拡張(最大64×64)と複数変換の適応切替(DCT-II、DSTV-II、DCTV-III等)による一般化に加えて、低域変換係数に対する二次変換の導入により、残差の周波数分布に合わせたエネルギー集中を更に高めた。

ループフィルタは、デブロッキングフィルタに加え、HEVCで導入されたSAO(Sample Adaptive Offset)、VVCで追加されたALF(Adaptive Loop Filter)により、符号化歪の抑制と次フレーム予測精度の向上を両立する。フィルタの適用/非適用や強度は、境界条件・残差有無・予測モードなどに応じて適応制御される。

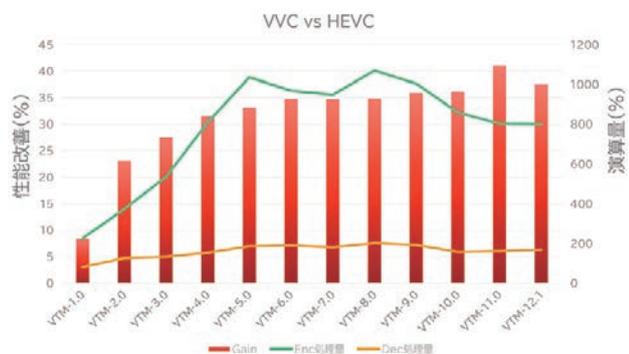
さらにVVCは、実運用における変動する回線品質や視聴デバイスの多様性へ適応するため、Reference Picture Resampling(RPR)を導入した。これは参照画像の解像度を動的に切替えて予測を成立させる機構(適応解像度変更)であり、低ビットレート域では一旦解像度を下げて符号化することで主観画質が改善する。放送のように回線品質が固定される場合でも、符号化難易度の高いシーンで劣化を緩和する効果がある。

標準化における要素技術の採否は、テストモデル(HEVCではHM、VVCではVTM)への実装と、あらかじめ定められた共通テスト条件での比較評価により検証される。評

価は単に符号化性能の向上にとどまらず、複数の観点から総合的に行われる。具体的には、提案技術を参照モデルに組み込んだ場合の符号化効率（ビットレート削減率）に加え、エンコード時間及びデコード時間といった処理量指標が測定される。さらに、これらの時間指標だけでは十分に把握できない要素として、並列化のしやすさやメモリアクセス構造、制御分岐の複雑さなど、ハードウェアやソフトウェアでの実現容易性についても専門的観点から審議される。

これらの評価結果は、個々の指標を独立に判断するのではなく、符号化性能向上と実装コストとのトレードオフを考慮した総合評価として整理される。その過程では、類似提案との比較や他ツールとの相互作用、将来的な最適化余地なども含めた検討が行われ、最終的に規格へ採用するか否かが決定される。

VVCの標準化においては、このような参照モデルを用いた評価と審議が、年4回程度開催されるJVET会合を通じて継続的に実施され、技術探索段階の開始から標準化完了まで、おおむね5年程度を要した。その間、VTMは版を重ねるごとに機能拡張と整理が進められ、方式全体としての性能とバランスが段階的に洗練された。その結果、最終的なVVC規格では、共通テスト条件においてHEVC参照モデル(HM)比でおおむね30~45%程度のビットレート削減が確認され、世代進化として期待された圧縮性能向上が達成された。標準化段階でハードウェア実装の容易性についても丁寧に議論されたことにより、標準化完了後1年足らずで8K解像度に対応したVVCデコーダチップが市場に登場しており、実装面での立ち上がりは早かったと言える。



■ 図4. VVC標準化における符号化性能推移

3. 主観評価と放送品質の所要ビットレート推定

映像符号化方式の放送サービスへの実用化を進めるにあたっては、符号化効率の改善率という指標のみでは運用設計は困難である。放送品質の所要ビットレートを、主観

評価によって定量化する必要がある。2022年2月に報告された電波産業会 (ARIB) の映像符号化方式作業班の取り組みでは、MOS (Mean Opinion Score) のしきい値として、「MOS 3.0未満の映像が存在しない」「ほぼすべて (7種以上) の映像でMOS 3.5以上」という条件を満たすビットレートを所要値として推定する方法が採用された。評価シーケンスは、SDR/HDRを含む4K及びHDの多様なコンテンツ群から構成され、信頼区間 (95%) を明示した上で品質の安定性を検証した。実験から、2K映像の所要ビットレートは約7Mbps、4K映像の所要ビットレートは約30Mbpsと見積もられた。また、実験結果の詳細な分析により2K映像については約5Mbps以上、4K映像については約22Mbps以上で放送品質が担保されることが推定され、運用上の帯域設計に示唆を与えた。

主観評価の意義は2つある。第1に、客観指標 (PSNR, SSIM, VMAF等) だけでは捉えきれないアーチファクトや視覚的妥当性を検知できる点である。第2に、放送等の固定ビットレートのメディアにおいて運用設計で重要となる「最悪ケース」を把握できる点である。平均的な改善率が高くとも、一部コンテンツで品質劣化が顕著ならば、サービス全体のQoEが棄損する。したがって、評価群の裾野を広げ、動き量・テクスチャ・コントラスト・ダイナミックレンジの異なる多様なシーケンスを含めることが重要である。ARIBにおける主観評価実験は、放送運用を想定して設計されており、VVC導入による効果をMOSのしきい値により明確化し、実際の運用に反映可能な形で評価可能にした点で有用である。

RPRやスクリーンコンテンツ向け拡張のようにユースケース特化のツールは、主観評価において効果の現れ方が顕著となる。例えば、低ビットレート域における適応解像度変更は、輪郭破綻の抑制や文字の可読性維持に寄与し、ニュース・字幕・テロップなど情報が多く含まれる放送コンテンツにおいてQoE向上の効果は大きい。これらの効果は、単純なPSNRの改善値だけでは十分に記述できないため、運用実験と主観評価の併用が不可欠であると考えられる。

4. マルチレイヤ符号化とパーソナライズ化サービス

VVCでは、映像フォーマット、ビット深度など、想定されるアプリケーションや運用条件に応じて複数のプロファイルが規定される形で標準が構成されており、汎用的な映像サービスを対象としたMain 10プロファイルをはじめとして、プロフェッショナル用途や高ビット深度映像に対応する複数



のプロファイルが定義されている。これらのプロファイルのうち、Multilayer Main 10プロファイルは、複数解像度の映像を効率的に圧縮・伝送することを目的とした、いわゆるスケーラブル符号化をサポートするプロファイルである。同プロファイルでは、低解像度映像を符号化したベースレイヤ（BL）と、高解像度成分を表現するエンハンスメントレイヤ（EL）からなる階層構造を用い、両レイヤ間の相関を積極的に活用する。このレイヤ間相関を利用することで、各レイヤを独立に符号化する場合と比べて、およそ35%程度の符号化効率改善が得られることが確認されている。また、受信端末は回線品質や表示能力、利用状況に応じて復号するレイヤを選択できるため、同一のビットストリームから、端末ごとに適切な品質や解像度の映像を提供することが可能となる。これにより、放送や配信サービスにおいて、視聴環境に適応した柔軟なサービス設計が実現される。例えば、複数解像度サービスでは、移動体向け伝送路でBLを確実に受信し、固定受信向け伝送路でELを追加受信することで、受信耐性とプレミアム品質を両立できる。また、放送通信融合では、放送波でBLを、通信でELを配信し、帯域の最適利用と品質差別化を同時に達成する。さらに、メインの番組（メインコンテンツ）に手話、字幕、多言語、チーム別解説、速報テロップ等の付加情報を重畳した映像（サブコンテンツ）を、メインコンテンツとともにマルチレイヤ符号化する「コンテンツレイヤリング」を行うことで、視聴者が受信側で嗜好に応じたコンテンツの選択が実現可能となる。このように、マルチレイヤ符号化を用いると、従来の放送システムでは実現できない様々なパーソナライズされたサービスが実現できる技術として実用化に期待が寄せられている。

国内標準化の進展として、ARIB STD-B32（映像・音声・多重）及びSTD-B60（MMTトランスポート）が2025年3月改訂でVVCを採用し、Multilayer Main 10の運用も包含した。これにより、次世代地上デジタル放送でパーソナ

ライズ化サービスの導入が制度面から可能となり、実装・運用設計の具体化が可能となった。NHK技研ではコンテンツレイヤリングに対応したVVCライブエンコーダを開発し、CMAF/MMTによる伝送が可能なシステムの検証を進めている。

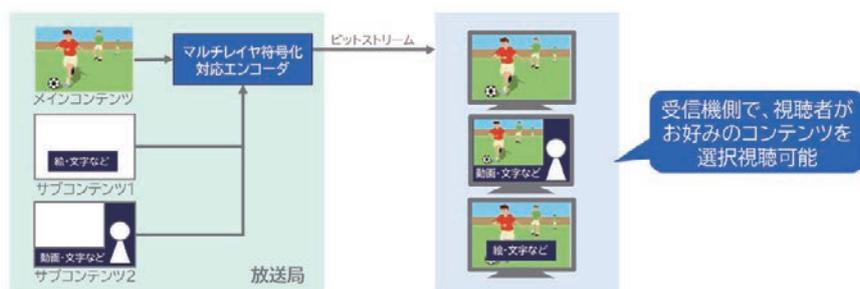
5. ポストVVC標準化動向

JVETでは2021年ごろ以降、VVCを超える圧縮性能を目標に評価用ソフトウェアEnhanced Compression Model (ECM) を用いた「ポストVVC」方式の技術探索を進めている。探索段階では性能志向で符号化ツールを選定し、従来エンコーダで採用が難しかったデコーダ側演算の大きいツールや、ニューラルネットワーク（AI）ベースの手法も検討対象に含めており、対VVC比28%の性能改善が報告されている。2025年7月会合に発行されたCfEに対して5件の応答が寄せられ2025年10月会合にてVVCからの大きな性能改善が確認されたため、今後CfPを経て標準化プロセスへ移行する見通しである。

上記のようにポストVVCでは、これまでのハイブリッド構造にAIを組み込む試みが進められており、これは従来の予測や変換を中心とした符号化手法に、学習ベースの推論を加えることで、より高い圧縮性能や適応性を目指す動きである。ただし、AIツールの導入には課題が多い。演算資源や電力消費の増加、端末の多様化への対応に加え、標準規格としてネットワークモデルの更新といった運用面の問題も想定される。

こうした背景の中で、ポストVVCの検討では、実際のサービスに適用するための現実的な条件として、遅延やデコーダの複雑度をどこまで許容するかが重要となる。特に、クラウドゲーミングやXR（拡張現実）といった新しいサービスでは、超低遅延と高品質が必須であり、これらの要求が次世代符号化方式の設計に大きな影響を与えるだろう。

（2025年11月5日 情報通信研究会より）



■ 図5. コンテンツレイヤリング