

動画トラフィックを半減する動画符号化方式 HEVC の標準化完成とモバイルへの展望

スマートフォンの普及により、モバイル動画の利用が急増している。ドコモは、動画トラフィックの半減に繋がる動画符号化方式 HEVC の標準化に参画し、技術提案と標準化推進の両面から高い圧縮率の達成に貢献した。HEVC は、既存方式の半分のデータ量で同等の動画品質を得ることがスマートフォンやタブレット端末を用いた主観評価にて確認されており、移動機端末での動作の実現性も見えてきている。

先進技術研究所
鈴木 芳典 藤林 暁
瀧上 順也

DOCOMO Innovations Inc. フランク ボッセン

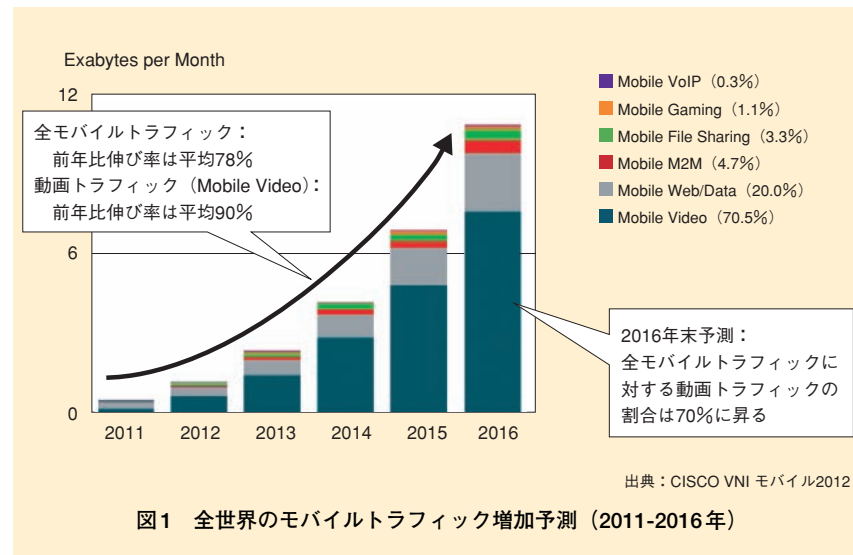
1. まえがき

ISO (International Organization for Standardization)^{*1}/IEC (International Electrotechnical Commission)^{*2} MPEG (Moving Picture Experts Group)^{*3}とITU-T WP (International Telecommunication Union-Telecommunication Standardization sector Working Party) 3/16^{*4}の動画符号化方式 HEVC (High Efficiency Video Coding) が2013年1月に標準化を完了する予定である。HEVC は、JCT-VC (Joint Collaborative Team on Video Coding)^{*5}に参画する世界各国40以上の機関の技術に支えられ、ワンセグ放送、ドコモdマーケットのVIDEOストア/アニメストアサービスやNOTTV^{*6}が利用する動画符号化方式H.264 /AVC (Advanced

Video Coding, 以下、H.264)^{*7}[1][2]と比較して、2倍の圧縮率を達成している[3]。

大画面のスマートフォンの普及に伴う動画の高精細化やLTE端末の普及を背景に、モバイル上の動画トラ

フィックの継続的な増加が見込まれている。図1[4]によれば、モバイルデータのトラフィックは今後も増加し続け、その最大要因が動画 (Mobile Video) であると推測されている。一方、HEVCは動画を従来の



*1 ISO：国際標準化機構。情報技術分野の標準化を行う組織であり、電気および電子通信分野を除く全産業分野に関する国際規格を作成する。
*2 IEC：国際電気標準会議。情報技術分野の標準化を行う組織であり、電気および電気通信の国際規格を作成する。

*3 MPEG：デジタル音声や映像の符号化および伝送方式などの技術標準仕様。ISOとIECの合同専門委員会の作業部会が策定した。MPEG-2では、デジタルテレビ放送やDVDへ適用、MPEG-4では移動端末などの低ビットレート領域まで適用領域を拡大した方式。

2倍以上の高効率で圧縮することが可能であり、これまでモバイル動画の主流であったVGA^{*8}から今後普及が見込まれるフルHD^{*9}を超える高精細動画までを高圧縮する。そのため、HEVCを広く活用することで、動画トラフィックの半減と、高精細動画の利用による動画サービスのさらなる発展が見込める。

ドコモは、HEVCの標準化開始以前から動画素材の提供や標準化の要求条件提案を行うなど中心プレイヤーとして活動しており[5]、標準化開始後も多数の技術採用でHEVCの高い圧縮率の実現に寄与した。また、HEVCの高圧縮達成に欠かせない標準ソフトウェアの主査を務めるなど、標準化の推進にも貢献している。さらに、3GPPに対して、HEVCのモバイル利用に向けた検討を促している[6]。

本稿では、HEVCの概要とともにHEVCへのドコモの貢献について解説する。また、モバイル利用を見据えたHEVCの性能とソフトウェアコーデカの性能について報告する。

2. HEVCの特徴

2.1 動画符号化の基本構造

動画のデータ量は膨大であるが、隣り合う画素や連続する画面間の絵柄は類似している。動画符号化では、この類似性を利用して元のデータ量を圧縮する。

図2に、HEVCとH.264に共通である動画符号化の基本構造を示す。動画を1枚1枚の画像に分け、さら

に画像を「ブロック分割」にて、小さな領域であるブロック(図2の画像内の黒枠)に分ける。ブロックが符号化の処理単位であり、左上の赤枠から1つずつ「予測」と「変換符号化^{*10}」処理が実施される。

予測では、符号化対象のブロックの絵柄に類似する信号(予測信号)を、符号化済みの画像信号から生成する。そして変換符号化では、画像情報の電力が低周波成分に集中しやすいという特徴を利用して、符号化対象のブロックの元信号から予測信号を除いた差分信号をデータ圧縮する。より具体的には、差分信号を周波数変換し、主観的に重要な低周波成分に多くのビット数を割り当てて量子化^{*11}し、残りの多くの高周波成分に少ないビット数を割り当てて量子化することで差分信号の情報量を小さくする。

このように、予測と変換で信号の電力を差分信号の低周波成分に集中させることにより、元信号の情報量を削減できる。しかしながら、予測と変換はブロック単位の処理である

ため、符号化データを「再生」にて復元した再生信号には、モザイクのようにブロック状の境界が見える場合がある。そこで、「品質改善処理」にて、再生信号に平滑化フィルタ^{*12}を掛けてブロック状のノイズを取り除く。品質改善処理後の再生信号は、予測のため「画像メモリ」に保存・管理される。

予測には、画面内予測と画面間予測の2つがある。以下にそれぞれの方式を述べる。

(1)画面内予測

最初に符号化する画像では、符号化対象ブロックの周辺の符号化済み画素を用いる画面内予測にて予測信号を生成する。図3に画面内予測の例(ブロックサイズが8×8の場合)を示す。白塗りと青枠で示したブロック内の画素の予測信号を、水色塗りと黄緑塗りで示した符号化済み画素を用いて生成する。例えば、青枠の画素P(x,y)の予測信号は黄緑塗りの画素Tを引き伸ばして求める。このとき、複数の符号化済み画素とP(x,y)との類似性は画面内の絵柄

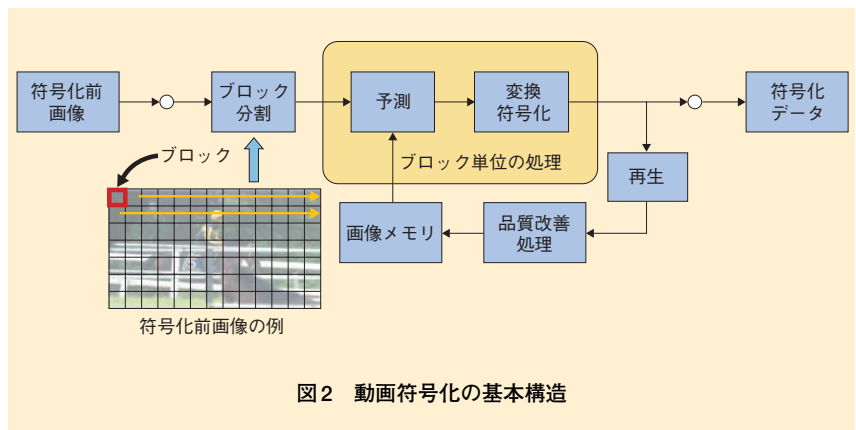


図2 動画符号化の基本構造

*4 ITU-T WP 3/16：電気通信分野における国際連合の専門機関である国際電気通信連合 (ITU) の通信部門において、映像や音声などのメディア符号化方式の標準化を担当している作業グループの1つ。
*5 JCT-VC：ITU-T WP3/16とISO/IEC MPEGが新たな映像符号化方式の検討を行うために制定された合同チーム。両機

関のビデオ符号化専門家グループのメンバーが参加している。
*6 NOTTV：(株)mmbiの商標または登録商標。
*7 H.264/AVC：ITU-TのSG16 (Study Group)のWP3とISO/IECのJCT1 (Joint Technical Committee 1)のMPEGとの合同チームであるJVT (Joint Video Team)で規格化された動画符号化方式の1つ。MPEG-2などの約

2倍の圧縮率を実現し、ワンセグ放送などで標準の動画圧縮形式として採用されている。
*8 VGA：画像の画素数が640×480pixの表示形式。
*9 HD：画像の画素数が1,280×720pixの表示形式。画素数が1,920×1,080pixの表示形式をフルHDと呼ぶ。

に応じて複雑に変化するため、画素Tの候補を黄緑塗り以外に複数用意し、差分値が小さくなる1つを選択する。具体的には、矢印で示した予測方向を複数用意し、ブロック単位で1つの予測方向を選択する。

(2)画面間予測

2番目以降の画像では、符号化対象ブロックに類似する絵柄のブロックを複数の符号化済みの画像（参照画像）を探索して見つける画面間予測が利用できる（図4）。元の位置から類似ブロックまでの移動を「動き

ベクトル」として検出し、類似ブロックの信号を予測信号とする。なお、隣接するブロックの動きベクトルは類似するため、差分ベクトルを付加情報として符号化する。

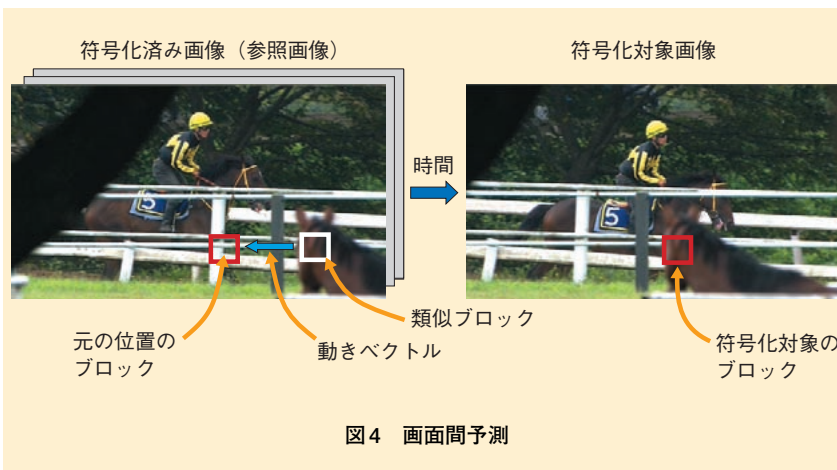
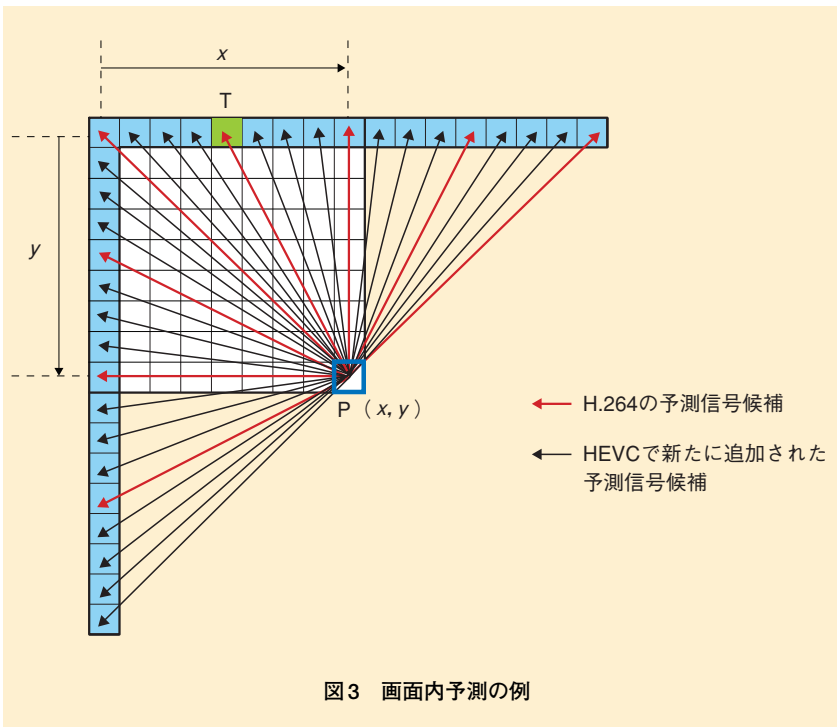
2.2 HEVCと従来技術の違い

(1)ブロックサイズの拡大

H.264以前の従来標準方式では、アナログテレビ放送（720×480pix）より画面解像度の小さい動画を対象として標準化作業を行い、その方式をほとんど変更せずにHD以上の解像度の動画にも適用していた。そのため、従来方式では、仮に図5（a）が最適な圧縮性能を得る画面解像度とブロックサイズの組合せであっても、図5（b）に示す画面解像度4倍の動画に図5（a）と同じブロックサイズを使用せざるを得なかった。つまり図5（b）は、図5（a）のブロックをそれぞれ4つに分けて扱うことに相当するため、予測・変換効率が低下し、最適な圧縮性能は得られない。そこで、従来方式では16×16pixであった基本ブロックサイズを、HEVCでは64×64pixに拡大した。基本ブロックの内部は最小4×4pixまで再分割可能であるため、HEVCでは、例えば図5（c）のように画面解像度に適したブロックサイズが選択できる。

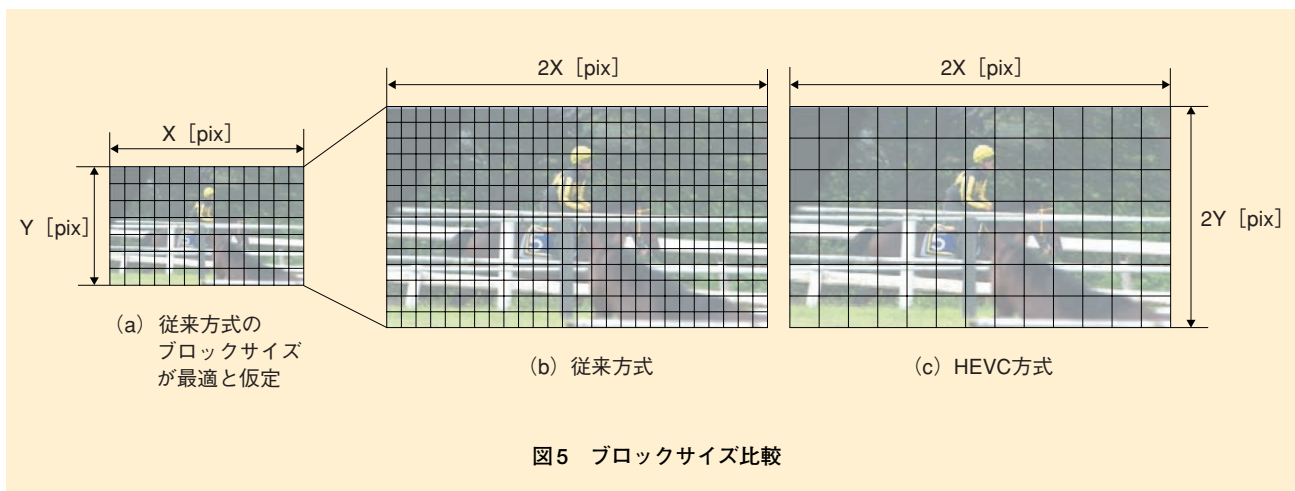
(2)画面内予測の改善

H.264の画面内予測では、選択可能な予測方向は図3に赤矢印で示す8方向のみであった。HEVCでは、



*10 変換符号化：動画のような連続的データを、数学的処理によって周波数軸上の特定成分のみを持った離散的なデータに変換してから符号化すること。情報量を圧縮することが可能。
*11 量子化：変換処理にて生成される離散データの値を、飛び飛びの値である粗い区間の代表値に対応づけること。歪を許し

ながら大幅に情報量を削減することが可能。
*12 平滑化フィルタ：信号の高帯域をカットすることにより、ノイズを除去するフィルタ。



カメラ技術の進歩と共に高精細化される決め細やかな画像を効率良く予測するため、黒矢印を加えた33個に選択可能な予測方向を増やした。この改良により、高精細画像の予測性能が向上した。

(3) 動画品質のさらなる改善

再生信号に発生するリンギング^{*13}やごま塩状のノイズを取り除くため、HEVCの品質改善処理(図2)にはブロックノイズ除去後の再生信号を元信号に近づける仕掛けが追加されている。改善された信号は予測に利用されるため、予測性能を高める効果もある。

2.3 ドコモの貢献

HEVCの高圧縮を実現するため、筆者らは画面内・画面間予測や変換符号化を改善する10を超える技術を考案し、HEVCへの採用を実現した。

(1) 技術採用での貢献

代表的な2つのHEVC採用技術に

ついて述べる。

① 参照画像管理方法

参照画像管理方法は動画の途中再生に伴う画面間予測の性能を高める技術である。ユーザは動画を鑑賞するとき、受信動画の先読み再生や放送チャンネルの切り替えなど、動画を途中から再生する処理を頻繁に行っている。動画の途中再生を可能にするため、動画符号化では、任意の画像から符号化データを再生できるように連続する画像を圧縮する。図6に、連続する画像の符号化/表示順と画面間予測に用いる参照画像の管理手順の例を示す。上段の四角は各画像を示しており、中段と下段は符号化済みの画像を参照画像として保存する画像メモリ(図2)の状態を示している。

動画の途中再生を可能にするため、符号化データには通常1~2秒ごとに画面内予測のみで

符号化したランダム・アクセス点(I)を設ける。また、早送り再生を可能にするため、再生が必須で参照画像として画像メモリに蓄積される画像(P)と、再生せずに読み飛ばし可能な画像(B)を設ける。画面間予測には表示順で未来の画像を参照画像に含めることで予測性能が高まるという性質があるため、通常、Bの符号化順を1画面分遅らせる(図6上段)。例えば、 B_2 の予測に P_1 に加えて表示順で未来の P_3 が使用可能となる(A)。

I_5 から途中再生を開始するとき、 P_7 の予測処理に I_5 以前に符号化した画像は使用できない。そこで、H.264では、 I_5 の再生直後に、画像メモリに保存していた参照画像を使用できなくしていた(図6中段の青四角)。そのため、 $P_1 \sim P_7$ を全て再生する場合においても B_4 の予測に I_5 しか

*13 リンギング：画像のエッジ付近に擬似的に発生するノイズで、本来存在しないエッジが重なって見える。不要な高周波成分が原因で発生する。

使用できなかった (B)。

筆者らは、画面間予測の性能は利用する参照画像が多いほど高いという性質に着目し、 I_5 以前に符号化した画像を予測に使用できなくするタイミングを P_7 の再生後に遅らせる (図6下段の青四角) ことで、 B_4 の予測に複数の参照画像を利用可能 (C) とする手法を考案した。この手法により、圧縮率が平均4%、最高で8%程度まで向上した (2倍の圧縮率は50%の圧縮率向上に相当)。

②動き情報共有

動き情報共有は画面間の動きをより効率良く符号化するための技術である。動物体の端では、図7の赤枠のように1つのブロックが動きの異なる2つの領域を含む場合がある。この例では、水色枠で囲ったように、赤枠のブロックの左半分領域と、白枠のブロックが同じ動きをもつ。従来は、赤枠のブロックを2分割し、個別にそれぞれの動きを符号化していた。

そこで、筆者らは、水色枠を1つの動きで予測するため、白枠のブロックの動きの情報を赤枠の左半分領域に共有する手法を考案した。この手法によれば、赤枠の左半分領域の信号を白枠のブロックの動きを参照して予測できる。隣接する2ブロックの動きを共有して符号化する技術の導入により、圧縮率

が平均3%、最高で5%程度まで向上した。

(2)標準化の推進にかかわる貢献

HEVC標準化の推進にかかわるドコモの貢献について述べる。まず、筆者らは、標準化作業に用

いる動画素材の提供やHEVCの性能目標に関する要求条件の提案を行うなど、HEVCの標準化開始以前から標準化の推進に貢献してきた。HEVCの標準化開始に先立って実施した提案募集では、ドコモ提案方式

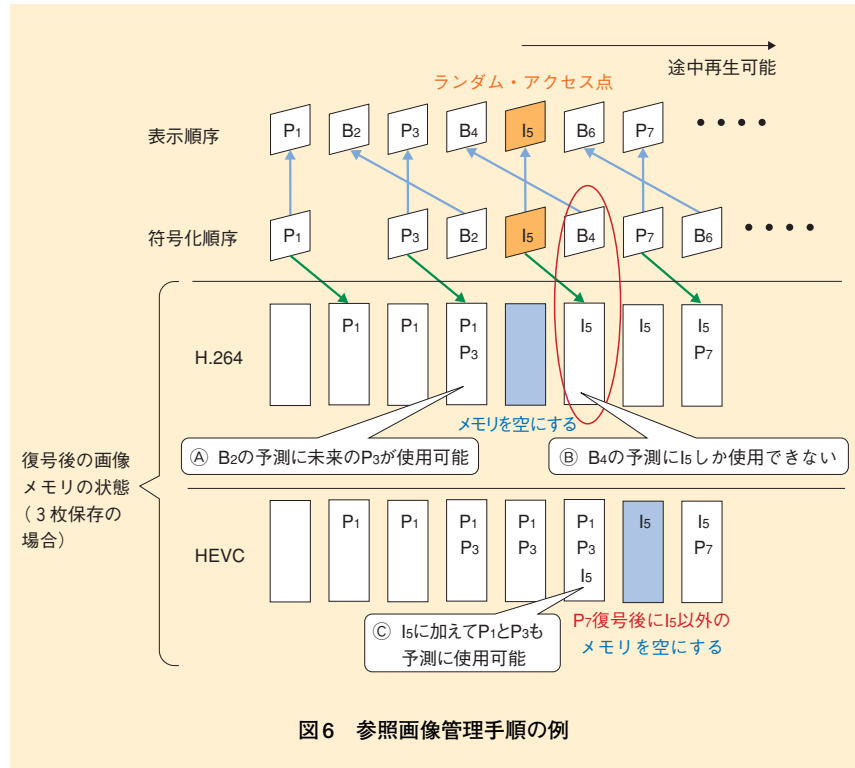


図6 参照画像管理手順の例

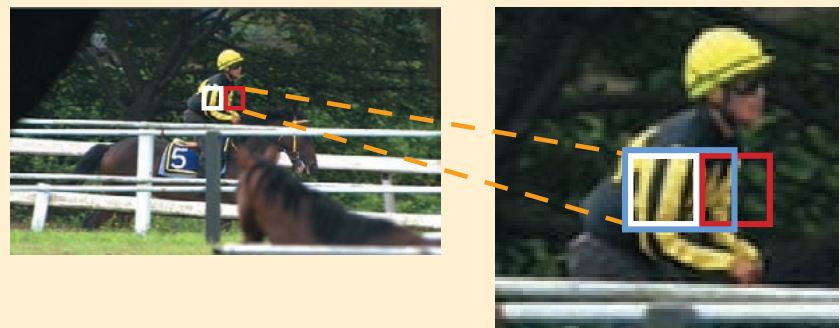


図7 動き共有予測

は応募のあった27方式中、差のない上位5方式に評価された[5]。

標準化開始後、ドコモはHEVC方式の高圧縮達成に欠かせない標準ソフトウェアの主査を務めた。ソフトウェア実装が難しい技術の見直しや規格草案文書とソフトウェアの不一致の解消を促すなど、HEVCの完成度向上に重要な役割を果たした。

また、ドコモは主観評価試験を通してモバイル動画サービスに必要となるHEVCの機能を検証し、HEVC利用時の機能構成を規定するプロファイル^{*14}の提案を行った[7]。現在のHEVCのプロファイルはドコモ提案を参考にして策定されている[8]。

3. HEVC の性能

3.1 HEVC の基本性能

現在主流のH.264とHEVCの圧縮

性能の比較検証結果を報告する。具体的にはH.264とHEVCを対象に、主観的な品質が同程度となるデータ量を比較した。

本検証では、動画の非専門家24名にH.264とHEVCで圧縮されたさまざまなデータ量の動画を見て、それぞれの動画を5段階の絶対評価(1が非常に悪い、5が非常によい)で評価してもらった。この検証方法は、ITUで定める標準の評価方法[9]に基づいている。検証結果を図8に示す。図8の横軸は動画の解像度と種類、縦軸は動画の主観品質に対する24名の評価の平均値(MOS(Mean Opinion Score)^{*15}値)を示しており、紺の棒グラフがH.264、赤の棒グラフがHEVCで圧縮した動画の評価結果を表している。各棒グラフ内の数値は各動画のデータ量を示

し、棒グラフ上の区間記号は評価値の信頼の度合いを示す。2つの棒グラフの区間が重複するとき、統計学的には2つの動画の主観品質に差がないとみなされる。この結果から、HEVCが、動画の種類・解像度によらずH.264の半分以下のデータ量で同程度の主観品質を達成していることが確認できる。なおこの結果は、タブレット・ハンドセット型などの携帯端末に加え、PCディスプレイでも同様の結果が得られることが確認されている。

3.2 ドコモサービスへのHEVC活用

まえがきに示したとおりドコモの動画サービスにおいてもH.264が利用されている。そのため、HEVCの高い圧縮性能により、現在サービス

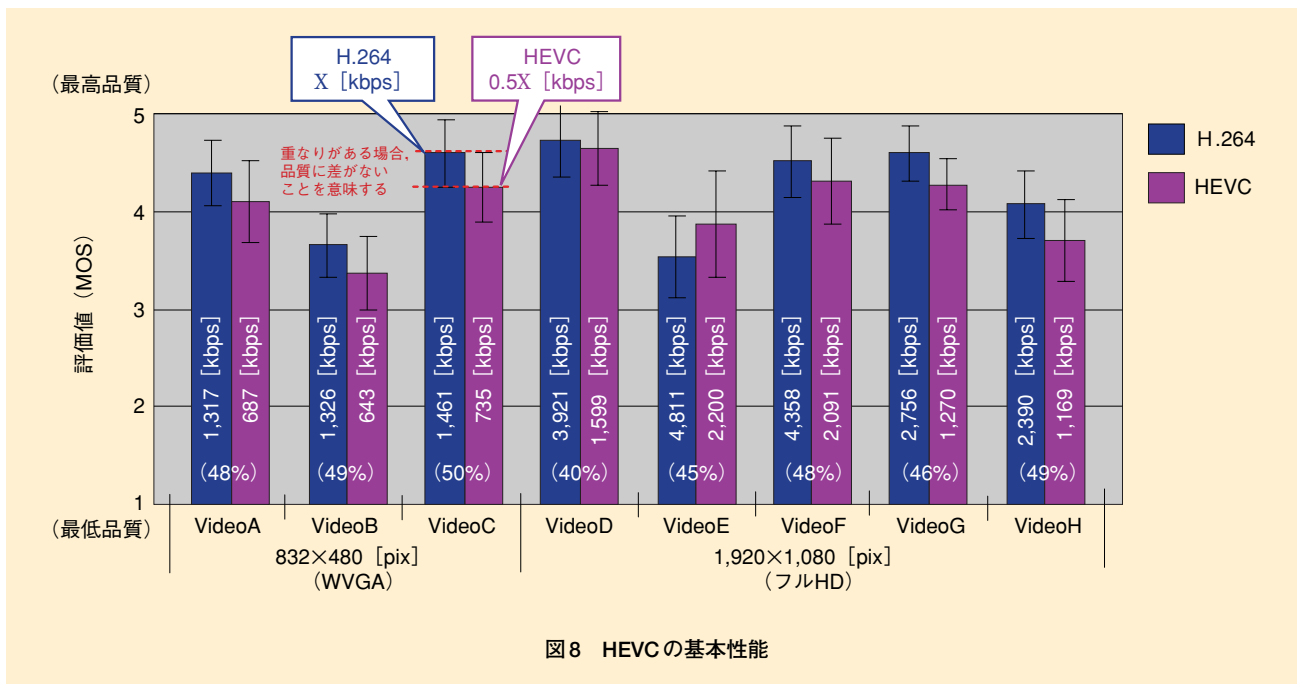


図8 HEVCの基本性能

*14 プロファイル：用途に応じて端末間の相互接続を確保するため、符号化機能全体のサブセットを標準化の合意により定義したもの。

*15 MOS：広く用いられる一般的な主観品質尺度の1つ。複数の被験者による主観的な評価を平均した値。

提供している半分以下のデータ量でも現在と同程度の主観的な品質を達成することが期待できる。つまり、現在のサービス品質を保ちながら伝送データ量の軽減に繋がると考えられる(表1)。また、伝送データ量を現状と変えずに、より高品質な動画コンテンツ配信も可能となる。例えば現在VIDEOストアで「すごく綺麗」と設定されるコンテンツ(WVGA (Wide-VGA)^{*16}、毎秒30コマ)で設定されるデータ量(1.5Mbps)でより高品質なHD動画(720p、毎秒30コマ)の提供が可能となる。またHEVCのアニメに対する効果はさらに高く、200kbpsでも非常に高品質なアニメ動画(WVGA、毎秒24コマ)の提供が可能となる。

これによりドコモにとっては、伝送帯域の効率化利用、急増する動画トラフィックの緩和に繋がる。ユーザにとっては、再生までの待ち時間短縮や閲覧中の動画途切れの頻度軽減など、より快適な動画視聴が期待できる。またコンテンツ配信者にとっても、高品質な動画のモバイル配信が可能となるなどのメリットが期

待できる。さらにHEVCを活用すれば、表1に示すとおり、フルHD動画やフルHDの4倍もの画素をもつ高精細な4K^{*17}動画のようなより高品質、高コマ数のなめらかな動画もLTE網の実伝送レートで配信可能となる。

3.3 HEVCのさらなる活用

ドコモの動画サービス以外にも、Hulu^{*18}などの動画オンデマンドサービスやYouTube^{*19}に代表される動画共有サイトの閲覧もモバイルで数多く利用されるようになった。

そのため、HEVCの普及後も、インターネット上には既存の方式(H.264、MPEG2など)で圧縮されたコンテンツが数多く存在し、すぐには動画トラフィックの軽減に繋がらないおそれがある。

そこで、このようなインターネットを流れる既存方式で圧縮された動画コンテンツを一度再生して再度HEVCで圧縮することで、さらなる動画トラフィック軽減の可能性を検討した。

具体的には、現在のモバイル網における動画トラフィックの大部分を

占めるYouTube動画に対し、HEVCを用いて再圧縮を行どれくらいデータ量削減が見込めるか検証した。検証では実際の複数の人気ジャンルのYouTube動画を対象とし、各YouTube動画をHEVCで再圧縮した際に、どの程度までデータ量を減らしても主観的な品質が保てるかを3.1節と同じ評価方法[9]で調査した。

動画の非専門家20名に元のYouTube動画とそれらをさまざまなデータ量でHEVCを用いて再圧縮した動画をPCディスプレイ上で見て、各動画の品質を5段階の絶対評価で評価してもらった。

図9に評価結果を示す。図9の横軸はコンテンツの解像度と種類、縦軸が動画の品質に対する評価の平均値(MOS)を示し、区間記号は3.1節と同様の意味をもつ。図9の結果からわかるように、YouTube動画で提供されるデータ量をHEVC利用により同程度の主観品質を保ちながら元の約45%のデータ量に再圧縮可能である。また動画によってはデータ量を元の30%まで削減しても主観品質に差異がないコンテンツも存在した。

この結果から、例えばインターネット上を流れる動画の伝送データ量をモバイルネットワークに流れる前にHEVCの再圧縮にて削減できれば、主観品質を保ちながらさらなる動画トラフィックの軽減が期待される。

表1 HEVCが実現する動画伝送データ量

解像度 [pix]	コマ数 [fps]	データ品質の実現に必要なデータ量 [kbps]
360p	30	300
WVGA		750
HD		1,000
フルHD		2,000
4K		10,000

* 16 WVGA：画像の画素数が832×480pixの表示形式。

* 17 4K：画像の画素数が3,840×2,160pixまたは4,096×2,340pixの表示形式。

* 18 Hulu：Hulu, LLCの商標または登録商標。

* 19 YouTube：Google, Inc.の商標または登録商標。

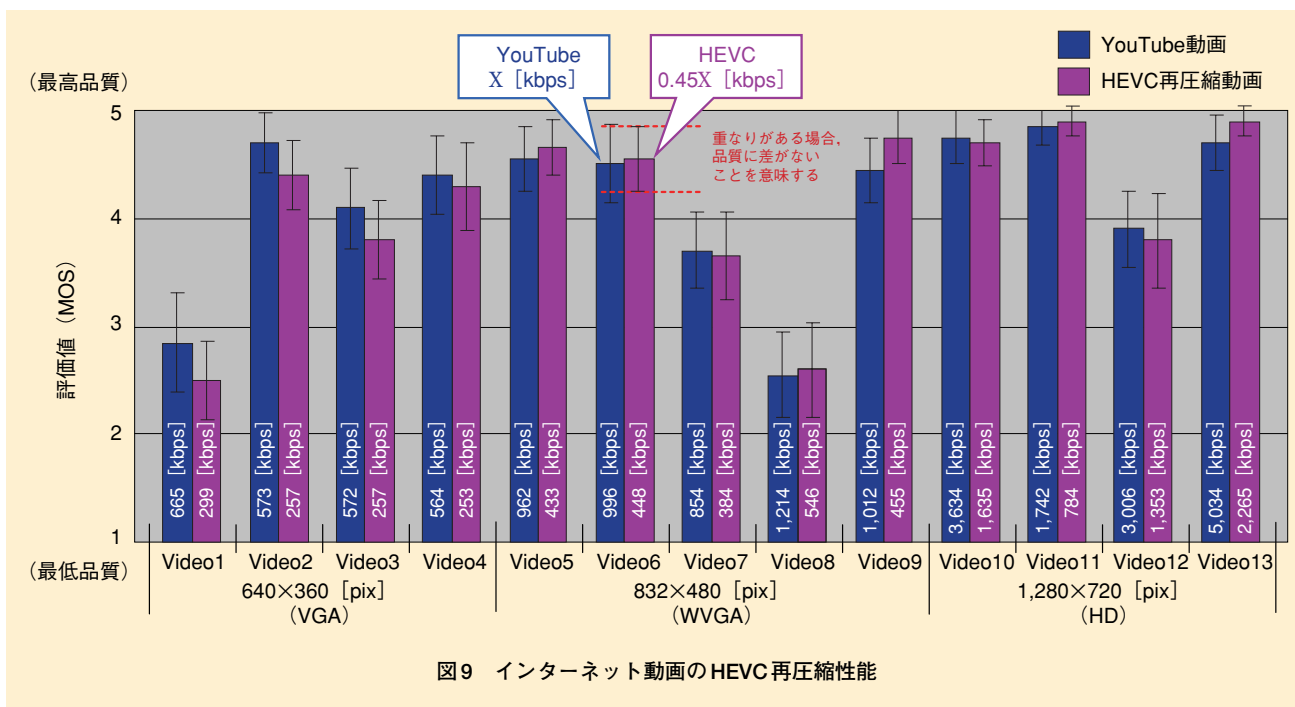


図9 インターネット動画のHEVC再圧縮性能

4. HEVCソフトウェアデコーダ開発

以上のようにHEVCがモバイル環境においてH.264に比較して2倍の圧縮率を達成しさまざまなメリットをもたらす可能性を秘めることが確認されたが、モバイル環境で活用するためには、HEVCを移動機端末環境で動作させる必要がある。

その1つの取り組みとして、演算量や必要メモリ量の検証を主目的として、Android™*20 端末で主流のARMプロセッサ*21 上で動作するHEVCソフトデコーダおよびビデオプレーヤーアプリの開発に取り組んできた。

開発を進めた結果、規格草案DIS (Draft International Standard) 仕様の

HEVCデコーダは、Dual Core*22 1.5GHzのARMプロセッサ上でソフトウェアでもHD品質のHEVC動画(720p, 毎秒30コマ, 1Mbps)を実時間再生可能であることが実証された。

またHEVC動画再生時のCPU使用率は約60%であり、HD動画の8時間以上の連続再生を確認した。このことから、消費電力の観点からも実用レベルであることが実証されている。今後は、より実用環境での開発・検討を進める予定である。

5. あとがき

2013年1月に標準化が完了する動画符号化方式HEVCの特徴とHEVCへのドコモの貢献について解説した。また、HEVCにはモバイル動画

の伝送データ量を同品質のまま従来の半分に減らす効果があることを実際の主観評価にて確認するとともに、ドコモ動画サービスでのHEVCの効果やモバイルでのさらなる展望を示した。

今後は、HEVCの世界的な普及のため、3GPPへの働きかけやモバイルサービスへのHEVC導入に向けた活動を進めていく。

文献

- [1] ISO/IEC 14496-10:2009 : “Information technology – Coding of audio-visual objects – Part 10 : Advanced Video Coding,” May 2009.
- [2] ITU-T Recommendation H.264 : “Advanced video coding for generic audiovisual services,” Mar. 2005.
- [3] V. Baroncini, J.-R. Ohm and G. Sulli-

* 20 Android™ : スマートフォンやタブレット向けのオペレーティングシステム、ミドルウェア、主要なアプリケーションからなるソフトウェアプラットフォーム。米国Google, Inc.の商標または登録商標。

* 21 ARMプロセッサ : 英ARM社が開発したARMアーキテクチャを採用したプロセッサの総称。スマートフォンへの搭載率が

高い。

* 22 Dual Core : 1つのパッケージに2つのCPUコアを集積したプロセッサのこと。

- van : "Report on preliminary subjective testing of HEVC compression capability," JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-H1004, 8th Meeting : San José, CA, USA, 1-10 Feb. 2012.
- [4] Cisco Systems, Inc. : "Cisco Visual Networking Index : 全世界のモバイル データ トラフィックの予測、2011 ~ 2016 年アップデート," Feb. 2012.
- [5] 鈴木, ほか : "高品質・高効率な映像配信を実現する次世代映像符号化方式の標準化動向とドコモの提案技術概要," 本誌, Vol.18, No.4, pp.38-45, Jan. 2011.
- [6] Qualcomm, NTT DOCOMO, et al. : "New Work Item Description - High Efficiency Video Coding," 3GPP, S4-121208, TSG SA WG4 70th Meeting, Aug. 2012.
- [7] T.K. Tan, A. Fujibayashi, Y. Suzuki and J. Takiue : "Objective and subjective evaluation of HM5.0.," JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-H0116, 8th Meeting: San José, CA, USA, 1-10 Feb. 2012.
- [8] M. Horowitz, A. Ichigaya, K. McCann, T. Nishi, S. Sekiguchi, T. Suzuki, T. K. Tan, W. Wan and M. Zhou : "Straw man for Profiles and levels," JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JCTVC-H0734, 8th Meeting: San José, CA, USA, 1-10 Feb. 2012.
- [9] Recommendation ITU-R BT.500-11 : "Methodology for the subjective assessment of the quality of television pictures."