

滑らかな動きを表現する 高フレームレート映像符号化技術

被写体の動き表現の滑らかさに直結するフレームレートは、映像の高臨場感において解像度に並ぶ重要な要素です。本稿では、従来の2倍のフレームレートとなる120 fps (frames per second) 映像サービスを実現するために、膨大な映像データをリアルタイム符号化可能なHEVC (High Efficiency Video Coding) ハードウェアエンコーダ技術について紹介します。

おおもり ゆうや おおにし たかゆき

大森 優也 / 大西 隆之

いわさき ひろえ しみず あつし

岩崎 裕江 / 清水 淳

NTTメディアインテリジェンス研究所

超高臨場感映像サービスへの期待

近年、より臨場感の高い映像の提供に向けて、従来のHDTV (High Definition Television) を超える高解像度の映像サービスに大きな期待が集まっています。家電量販店などでの4K対応TVの販売数増大に呼応するように、4K映像の配信サービスが次々と展開され始め、さらに2016年には8K映像の試験放送が開始されるなど、来たる2020年に向けて超高精細な4K・8K映像をますます身近に楽しめる世界の到来が予想されています。

一方で、高臨場な映像の追求は、4K・8Kといった解像度の向上のみで達成されるものではありません。映像品質の尺度には、解像度以外にもフレームレート、色域、ビット深度などのさまざまな方向性が存在し、これらの要素の組合せによって、臨場感にあふれた映像が生まれます (図1)。

フレームレートのもたらす効果

こうした技術要素の中でも、フレームレート (単位時間当りのフレーム数) は、動画像の本質ともいえる動き表現の滑らかさに直結し、映像の臨場感において解像度に並んで重要な要素と考

えられています。高フレームレート映像は、特にスポーツのような動きの速いコンテンツに対して効果を発揮します。高速な被写体でも、ぼやけ感なく鮮明かつ滑らかな映像表現が可能となり、主観映像品質が大きく向上します⁽¹⁾。従来の映像サービスにおけるフレームレートは最大60 fps (frames per second) ですが、ITU (International Telecommunication Union) により勧告された次世代映像規格「UHDTV (Ultra High Definition Television)」では、120 fpsまでの高フレームレートに対応しました (図

1)。120 fpsの高フレームレートが実現すれば、例えばスポーツ中継やコンサートのライブ映像を迫力ある臨場感で提供することが可能となり、また、VR (Virtual Reality) やE-sportsといった新しい映像サービス分野においてもより自然で現実感にあふれた視聴体験が期待できます (図2)。次世代の超高臨場感映像サービスにおいては、4K・8Kといった高解像度のみならず、従来の60 fpsを超えた高フレームレートが、不可欠な要素といえるでしょう。

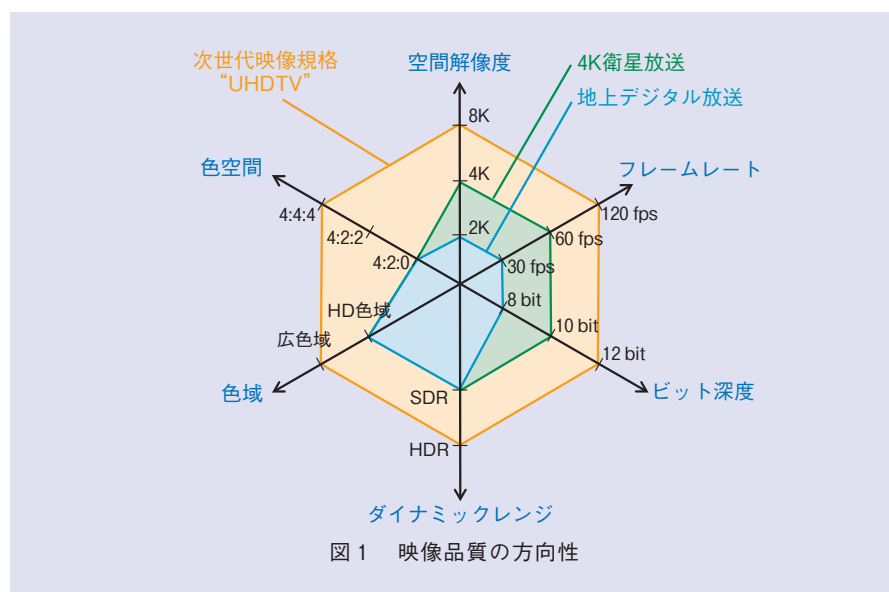


図1 映像品質の方向性



図2 高フレームレート映像サービス

リアルタイムエンコードの必要性

解像度やフレームレートの向上は映像データサイズの増大に直結するため、膨大なデータを高品質のまま伝送可能な映像符号化技術が必要となります。さらに、スポーツ中継のようなライブ映像配信や映像素材の伝送用途では、リアルタイム性に加え、低遅延、小型、低消費電力などの性能も要求されます。これらの要求を同時に満たすためには、ハードウェア型のエンコーダが不可欠となります。

NTTメディアインテリジェンス研究所では、エンコーダLSI (Large Scale Integration) の開発を通じて、リアルタイム映像符号化技術の研究開発を進めてきました。2015年には、最新の映像符号化国際標準H.265/HEVC (High Efficiency Video Coding) に対応した、リアルタイムエンコーダLSI「NARA」を開発しました⁽²⁾。HEVCは従来の国際標準H.264/AVC (Advanced Video Coding) の2倍の符号化性能を持ち、情報量の膨大な高臨場映像の放送・配

信に適した技術です。NARAは、1チップで4K映像のHEVC符号化処理を行い、さらに4チップの連携動作により8K映像のエンコードにも対応するため、4K・8Kの超高精細映像をリアルタイムかつ高効率に圧縮・伝送することが可能です。一方で、フレームレートという観点では、最大60 fpsまでの対応となっていました。

高フレームレート120 fps対応リアルタイムエンコーダ

今回、NTTメディアインテリジェンス研究所では、高フレームレート120 fps映像を実時間で高効率に圧縮が可能な、120 fps対応リアルタイムHEVCエンコーダを開発しました⁽³⁾。カメラ・ディスプレイといった撮像系・表示系ではすでに高フレームレート対応機器が展開し始めている一方で、これまで高フレームレート（特に120 fps）に対応したリアルタイムエンコーダは存在しませんでした。そこで今回、既開発のNARAを倍速で動作させることで、フルFDながら従来の2倍のフレームレートとなる120 fps

対応を実現しています。さらにそのエンコーダ装置を複数協調動作させることで、4K/120fpsの高解像度かつ高フレームレートの映像のリアルタイム符号化が可能となります。

時間方向スケーラビリティへの対応

今回開発したエンコーダ装置は、120 fps映像サービス展開の観点から「時間方向スケーラビリティ」機能も実装しています。「時間方向スケーラビリティ」とは、異なるフレームレートへの拡張性・適応能力であり、今回の場合では既存の60 fps映像システムへの下位互換性を意味します。

120/60 fpsの時間スケーラビリティを持った映像エンコーダによる、120 fps映像配信の例を図3に示します。符号化された120 fps映像データは2種類のレイヤからなるビットストリームとして伝送されます。一方は、120 fpsのフレームが1枚ずつ間引かれた60 fps相当のベースレイヤ、もう一方は、120 fpsと60 fpsの差分となるエンハンスメントレイヤと呼ばれます。一般的な60 fps対応受信機の場合、ベースレイヤのストリームのみを受け取りデコードすることで、60 fps映像を再生します。一方、120 fps対応受信機では、ベースレイヤとエンハンスメントレイヤの両方のストリームを受け取り、組み合わせてデコードすることで、120 fps映像を再生します。このようなスケーラビリティは120 fpsエンコーダが備えるべき重要な機能であり、既存の60 fps映像サービスを保持しながら新たに120 fps対応を展開していくことが可能となります。特に国内デジタル放送においては、電波産業会のARIB STD-B32という標準規

格により、120 fps映像放送における符号化ストリームは上記のような30 fps互換性を持つことが必須となっています。

符号化ストリームに時間スケラビリティを持たせるためには、エンコーダ側で多くの変更・追加要素が必要となります。例えば、通常の配信・放送では映像のビットレートを一定量に制御して伝送することが多いですが、60 fps受信機へのスケラビリティを考えた場合、120 fpsデータとその一部である60 fpsデータの両方でビットレートが一定となる必要があります。これは、ベースレイヤとエンハンスメントレイヤに対する二重のビットレート制御機能を意味します。また、60 fps相当のデータが全体のストリームの中で偏った位置に存在した場合、60 fps受信機としてはデコードが追いつかずバッファ溢れが生じるおそれがあります。そのため、ベースレイヤの符号化データが全体的に偏りなく位置するような、符号化の順序変更が要求されます。こういった機能の追加・変更は、通常はLSI内部の符号化回路を大きく修正する必要がありますが、今回新規に開発したエンコーダでは、NARAのファームウェア構造を活用することでこの問題を解決しています。NARAのファームウェアは階層的な構造をしており、主要な3つの層から構成されますが、特に最上位層においては、HEVCの基礎的な機能をモジュールベースで入れ替えることのできる機構を設けています。時間方向スケラビリティで必要となる追加機能は、すべてこのカスタム部の修正のみで対応できる部分です。このため、本エンコーダでは、NARAのプログラム可能なインタフェースを通した簡単な

修正のみで、時間スケラブル機能の実装を実現しています。

エンコーダ協調動作による4K/120fpsへの対応

本エンコーダは、1台につきNARAを1つ搭載し、2K/120fpsに対応しています。さらに本装置は、複数台を協調動作させることで、4K/120fpsリアルタイムエンコードを行う機能を有しています。4K/120fpsエンコーダのシステム構成を図4に示します。4K/120fpsエンコーダは4つの2K/120fpsエンコーダで構成され、それぞれの2K/120fpsエンコーダが同期

信号によりタイムスタンプを共有することで、協調した動作を行います。入力として、2K/120fpsエンコーダは、4K映像が田の字に4分割された2K映像を受け取りますが、現状の非圧縮映像入出力インタフェースは60 fpsが主流であるため、2K映像を120 fpsのままで受け取るのは困難です。そこで本装置では、NARAに備わっていたマルチチャンネル入力機能を利用することで問題の解決を図っています。4分割された2K/120fps映像をベースレイヤとエンハンスメントレイヤに分離して2つの2K/60fps映像として扱うことで、既存の60 fpsインタフェースによる入

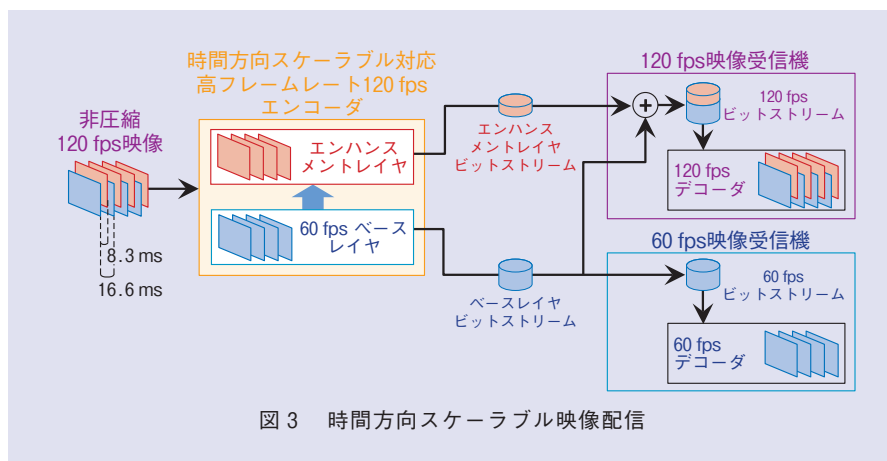


図3 時間方向スケラブル映像配信

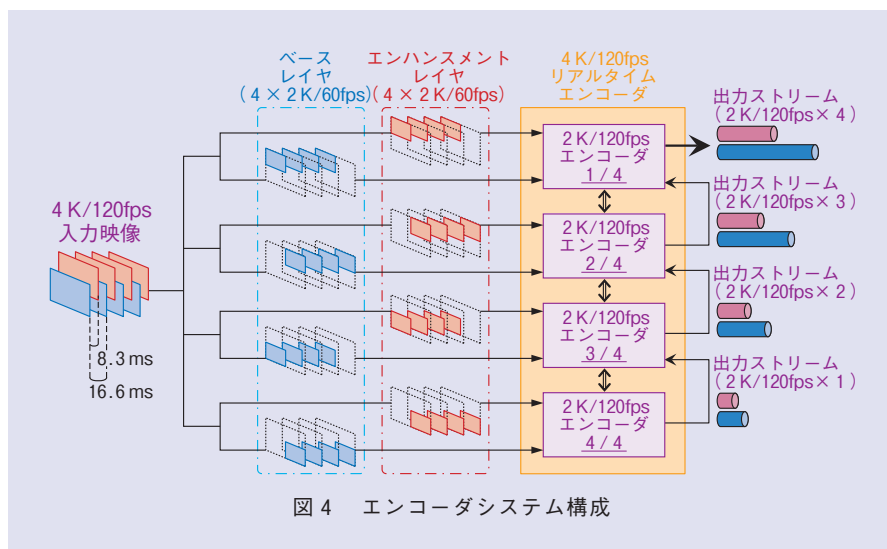


図4 エンコーダシステム構成

力を可能としています。エンコーダは、受け取った2つの60 fps映像から120 fps映像を再構成し符号化を行います。また出力においても、エンコーダ装置の持つマルチチャンネル多重化・出力機能が活用されます。NARA内部の多重化部により、ベースレイヤとエンハンスメントレイヤのデータは異なるビデオパッケージIDが付加されるとともに、NARA外部に設置した多重化部では、異なる装置から出力されたストリームを多重化することができます。本装置ではこの機能を利用し、各2K/120fpsエンコーダの出力部を多段接続することで、4本の2K/120fpsビットストリームを1本のストリームに多重化し出力することが可能となります。

今回開発した4台並列4K/120fpsリアルタイムエンコーダを図5に示します。NTT R&Dフォーラム2017にて、本装置を用いた4K/120fpsリアルタイムエンコードのデモンストレーションを行いました(図6)。膨大なデータとなる4K/120fps原画像を、リアルタイムに600分の1のデータサイズまでHEVCで圧縮しながら伝送し、受信側ではデコード映像をプロジェクタを用いて150インチの大画面で映し出しました。特にサッカーのようなスピード感のある映像では、従来の60 fpsと比較してより動きが滑らかで高い臨場感を感じるという声を多数いただいています。

今後の展開

ディスプレイ技術やカメラ性能の向上に伴い、映像の高品質化はさまざまな側面において進んでいくと予想されます。そのような中、高品質映像をより効率良く圧縮・伝送するリアルタイム符号化技術は、ますます必要とされ



図5 4台並列4K/120fpsリアルタイムエンコーダ



図6 4K/120fpsリアルタイムエンコードデモンストレーション

と考えられます。NTTメディアインテリジェンス研究所では、引き続き映像圧縮のキーデバイスとなるエンコーダ装置の研究開発に邁進し、高品質映像放送・配信サービスの拡大に貢献していきます。

参考文献

- (1) M. Emoto, Y. Kusakabe, and M. Sugawara : "High-Frame-Rate Motion Picture Quality and Its Independence of Viewing Distance," *Journal of Display Technology*, Vol.10, No.8, pp.635-641, 2014.
- (2) T. Onishi, T. Sano, Y. Nishida, K. Yokohari, J. Su, K. Nakamura, K. Nitta, K. Kawashima, J. Okamoto, N. Ono, R. Kusaba, A. Sagata, H. Iwasaki, M. Ikeda, and A. Shimizu : "Single-chip 4K 60 fps 4:2:2 HEVC video encoder LSI with 8K scalability," *VLSI Circuits 2015*, Kyoto, Japan, June 2015.
- (3) Y. Omori, T. Onishi, H. Iwasaki, and A. Shimizu : "A 120 fps high frame rate real-time HEVC video encoder with parallel configuration scalable to 4K," *COOL Chips 20*, Yokohama, Japan, April 2017.



(左から) 大西 隆之/ 大森 優也/
岩崎 裕江/ 清水 淳

NTTメディアインテリジェンス研究所では、より臨場感にあふれた次世代映像サービスのタイムリーな提供を実現すべく、今後も継続的な研究開発を進めています。

◆問い合わせ先

NTTメディアインテリジェンス研究所
画像メディアプロジェクト
TEL 046-859-8583
FAX 046-859-2829
E-mail omori.yuya@lab.ntt.co.jp