

動画像符号化の新規格HEVCに向けた 高効率な重み付き画素値予測技術

High-Efficiency Motion-Compensated Weighted Prediction Technology toward HEVC
Next-Generation Video Coding Standard

谷沢 昭行

山口 潤

中條 健

■ TANIZAWA Akiyuki

■ YAMAGUCHI Jun

■ CHUJOH Takeshi

ITU-T (国際電気通信連合-電気通信標準化部門)とISO/IEC (国際標準化機構/国際電気標準会議)による動画像符号化の現行国際規格であるH.264/AVCの2倍の圧縮効率を目指した標準化プロジェクトが開始され、新規格HEVC (High Efficiency Video Coding)として2013年の標準化完了に向けて活動が続けられている。

東芝は、標準化に先立って圧縮効率の改善を目指した技術開発を進め、HEVCの規格草案策定に向けて多数の技術提案を行ってきた。その一つとして、フェード効果など時間的な画素値変化を含む映像を高効率に予測する重み付き画素値予測 (WP: Weighted Prediction) 技術が規格草案に導入されており、圧縮効率の改善に貢献している。この規格により、タブレット、テレビ (TV)、パソコン (PC) など幅広いデジタル機器への高品質な映像配信の実現が期待されている。

The Joint Collaborative Team on Video Coding (JCT-VC) has been established to develop a next-generation video coding standard, called HEVC (High Efficiency Video Coding), which aims at a substantial improvement in coding efficiency compared with the existing H.264/AVC (Advanced Video Coding) standard. The JCT-VC is planning to finalize the HEVC standardization activity by the beginning of 2013.

Toshiba has been promoting the development of video coding technologies to improve coding efficiency even before this activity and has proposed many types of technologies toward HEVC standardization. One of these technologies, a motion-compensated weighted prediction (WP) technology to predict temporal illumination variations in specific video contents with a fade effect, has been adopted as a draft specification for HEVC. HEVC is expected to realize the distribution of high-quality video contents to various audiovisual devices including tablets, TV sets, and PCs.

1 まえがき

近年、スマートフォンなどの高機能携帯端末や、4K (3,840×2,160画素) テレビなど映像表示デバイスの技術革新により、映像は大画面化及び高解像度化の方向に進みつつある。しかし、増加するユーザー数と高解像度化という二重の負荷増に通信インフラの発展が追いついておらず、映像本来の品質を落とさずにデータ量を削減するための、新しい動画像符号化技術が望まれている。動画像符号化の現行国際規格として知られているITU-TとISO/IECによるH.264/AVCは、幅広いアプリケーションで普及が進んでいるが、圧縮効率と機能面で十分とは言えず、更なる圧縮効率の改善に向けた取組みが進められている。

ISO/IECのMPEG (Moving Picture Experts Group)とITU-TのVCEG (Video Coding Experts Group)は、H.264/AVCの更に2倍の圧縮効率を目指した、動画像符号化の次世代国際規格の標準化活動を行う新組織JCT-VC (Joint Collaborative Team on Video Coding)を2010年に立ち上げた。新規格の名称をHEVC (High Efficiency Video Coding)とし、最初の規格文書が2013年初頭に発行された。

東芝は、HEVCの標準化活動がスタートする以前から、来るべき動画像符号化の新規格に向けた圧縮効率改善技術の

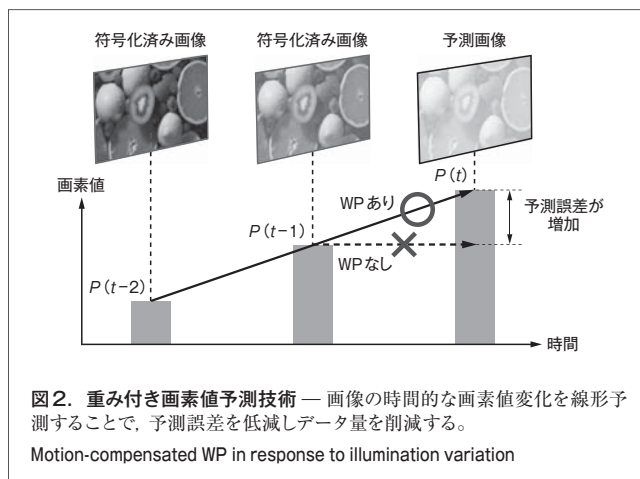
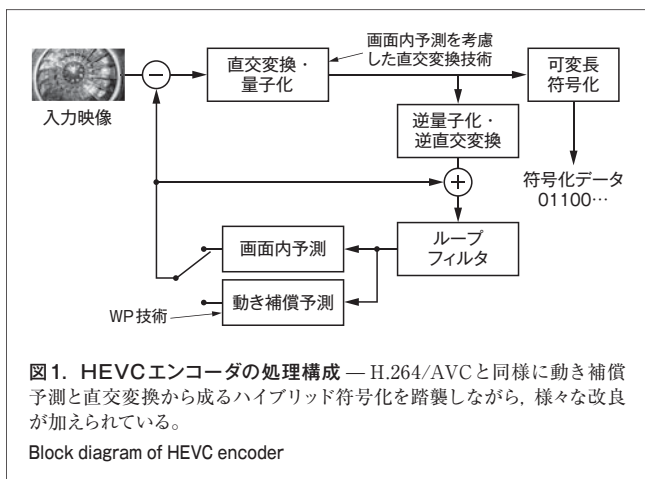
開発を進めており、HEVCの標準化活動のなかで数々の技術提案を行ってきた。重み付き画素値予測 (WP: Weighted Prediction)は、フェード効果など時間的な画素値変化を含む映像を高効率に予測する動き補償予測技術の一つであり、画面内予測を考慮した直交変換は、35種類のモードを持つ画面内予測の誤差を高効率で圧縮する技術として、当社が積極的に提案活動を行ってきた。これらの一部はHEVCの規格草案に採用されており、圧縮効率の改善に大きく貢献している。

ここでは、HEVCの特長と映像圧縮技術の概略の全体像、及び当社が開発したWP技術と直交変換技術について述べる。

2 HEVCの概要

JCT-VCは、2010年にHEVCのひな形となるテストモデル (HM: HEVC Test Model)を策定してから、約3年をかけてHMの完成度向上を進めてきた。2012年7月にISO/IECのMPEGからHEVC国際規格草案が発行され、2013年1月に最終国際規格草案が発行された。今後、スケーラブル符号化や多視点符号化への拡張が行われる予定である。

HEVCは、携帯機器から高精細映像機器までの幅広い応用機器を想定し、スーパーハイビジョン (Super HD)級の解像度まで対応している。また、フレームレートをスケーラブルに



変更する時間スケラビリティへの対応や、ランダムアクセス機能の拡張などが行われ、現行規格から機能的にも改良が図られている。技術の組合せを定義するプロファイルとして、当方は8ビット信号を扱うMainプロファイル、10ビット信号を扱うMain10プロファイル、及び8ビット信号の静止画像を扱うMain Still Pictureプロファイル、の三つが定められている。HEVCの処理構成は、H.264/AVCと同様、動き補償予測と直交変換を用いたハイブリッド符号化を踏襲しており、個々の技術を高解像度向けに拡張する形で、圧縮効率の改善を実現している。大別すると、動き補償予測、画面内予測、直交変換・量子化、逆量子化・逆直交変換、ループフィルタ、及び可変長符号化の六つの処理から成る(図1)。

3 圧縮効率改善技術

ここでは、当社が開発しHEVCに向けて提案した圧縮効率改善技術について述べる。

3.1 WP技術

WPは、当社がH.264/AVCの標準化活動時から継続的に開発を行っている技術であり、既に様々な映像機器で利用されている。一般的なハイブリッド符号化では、過去の画像からの動きを追跡して現在の画像にコピーする動き補償予測を用いることで、時間方向の画像の冗長性を削減している。WPは、動き補償予測された後の予測画素に対して、式(1)で表される乗算型の重み係数 W と加算型のオフセット O を用いて画素補正を行うことで、単純な動きの追跡だけでなく、画面間の時間的な明るさの変化を効率良く予測できる技術である(図2)。

$$P(t) = W \times P(t-1) + O \quad (1)$$

$P(t)$: 時刻 t における画素値

WPが効果的な映像として、映画やドラマなどで頻繁に用いられる白や黒へのフェードアウトシーンなどがある。

H.264/AVCでは、明示的モードと暗黙的モードの2種類の

WPモードが導入されている。明示的モードは重みパラメータをエンコーダで導出し、その情報をデコーダ側に伝送するモードであり、画像に適した任意のパラメータを設定できるメリットがある。しかし、パラメータを伝送するためのデータ量が増加するなどのデメリットもある。一方、暗黙的モードは、画像間の時間距離比を利用してパラメータを一意に導出するモードであり、パラメータをデコーダに伝送する必要がなく、データ量を節約できるメリットがある。しかし、双方向予測でしか利用できないデメリットがある。

WPはHEVCにも導入されており、H.264/AVCと同様に時間的な画素値変化を効率良く予測することが可能になっている。HEVCへのWP導入に向けて、当社は、重みパラメータを1パスエンコードで算出する手法やパラメータのデータ量を削減するための予測手法を提案し、H.264/AVCからの更なる改良を図っている。

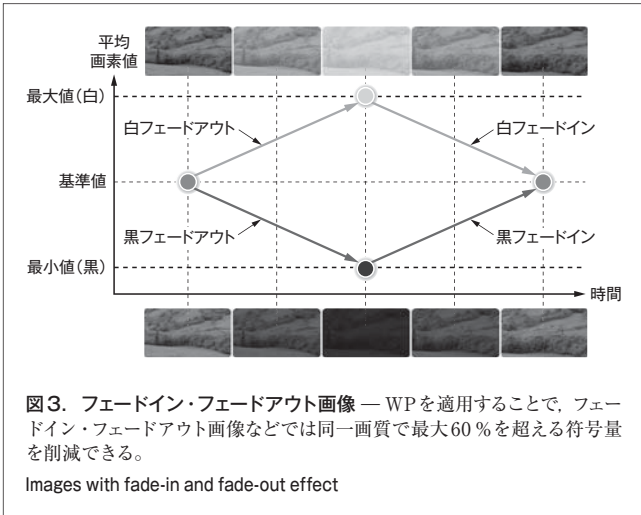
ここでは、HMに用いられている重みパラメータの導出手法について述べる。画素値の変化の傾きに対応する重み係数と、重み係数を乗じた際の誤差を補正するオフセットを、二つの画像から算出するため、式(2)で示す a ブレンドングの枠組みを用いる。

$$P(t) = a(t) \times P(t-1) + (1 - a(t)) \times C \quad (2)$$

$a(t)$: 時刻 t におけるブレンド比率
 C : ブレンドングの目標画素値

画像特徴量として画像を周波数成分に変換した際の直流成分と交流成分を算出し、これらの変化の度合いを比較することにより $a(t)$ を算出する。二つの画像の交流成分の比率により重み係数を算出した後、二つの直流成分と重み係数を用いてオフセットを算出する。ここでは、直流成分として画像の平均値を用い、交流成分として標準偏差を用いる。

一般的な自然映像に、白・黒方向のフェードアウト・フェードイン効果を加えて評価映像を作成し、開発した重みパラメータの導出手法を用いたWPを評価した(図3)。WPを適用し



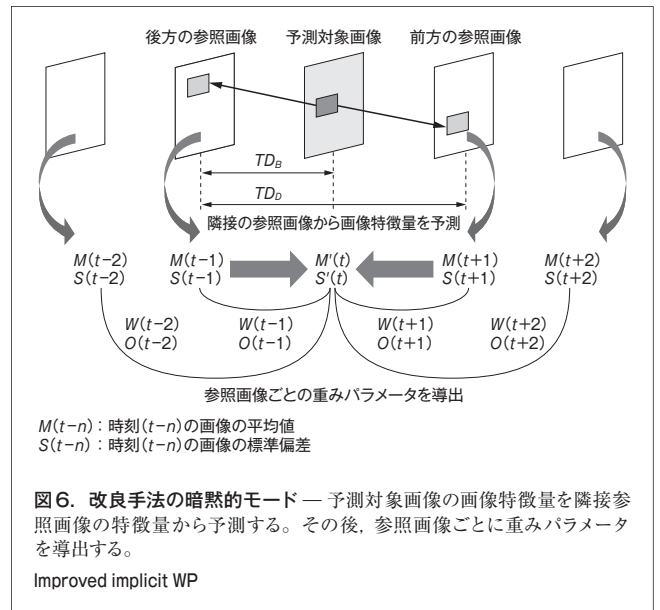
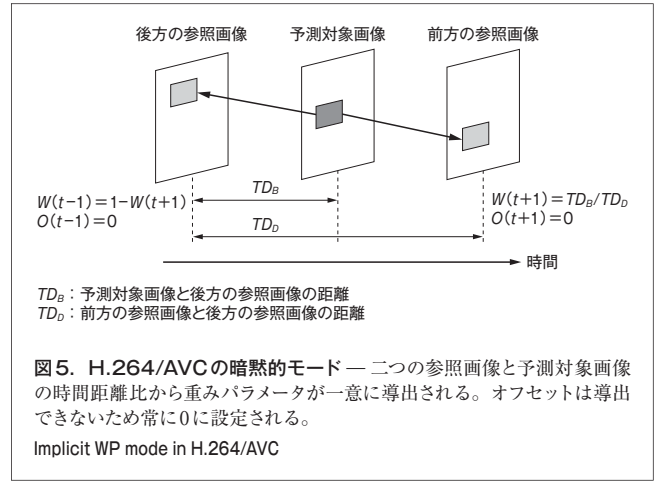
ない場合と比べて、同一画質で最大60%を超える符号量が削減された。このように時間的な画素値変化を伴う映像に対して、開発手法は極めて効果が高いことがわかる。

同一ビットレートでのWP適用ありなしの復号画像の比較を図4に示す。開発手法による画質(c)が、WPを適用しない映像(b)と比べて大幅に改善されていることがわかる。

3.2 WPを改良した暗黙的重みパラメータ導出技術

暗黙的模式では、双方向予測で用いる二つの参照画像と予測画像の時間距離比を算出し、重み係数を算出していた(図5)。しかし、この方法ではオフセットを利用できないだけでなく、同じ参照画像を用いた双方向予測や単方向予測で暗黙的模式が利用できない問題がある。

そこで、3.1節の重みパラメータ導出手法を暗黙的模式に拡張した¹⁾。暗黙的模式では、エンコード側とデコード側で同一の重みパラメータ導出処理を行うため、入力画像が利用できないうへ、復号処理が完了していない画像から画像特徴量を計算できない。そのため、既に復号処理が完了している二



つの参照画像からそれぞれの画像特徴量を計算した後、内挿又は外挿予測することで予測画像の画像特徴量を導出する。予測画像の画像特徴量が求めれば、3.1節と同じ枠組みで二つの画像の重み係数とオフセットを導出できる(図6)。この手法により、単方向予測にもWPを用いることが可能になる。

3.1節と同様に、H.264/AVCベースの暗黙的模式と開発手法を評価した。現行手法が適用できない単方向予測だけを用いる低遅延符号化条件において、同一画質で最大50%の符号量が削減できた。

3.3 画面内予測を考慮した直交変換技術

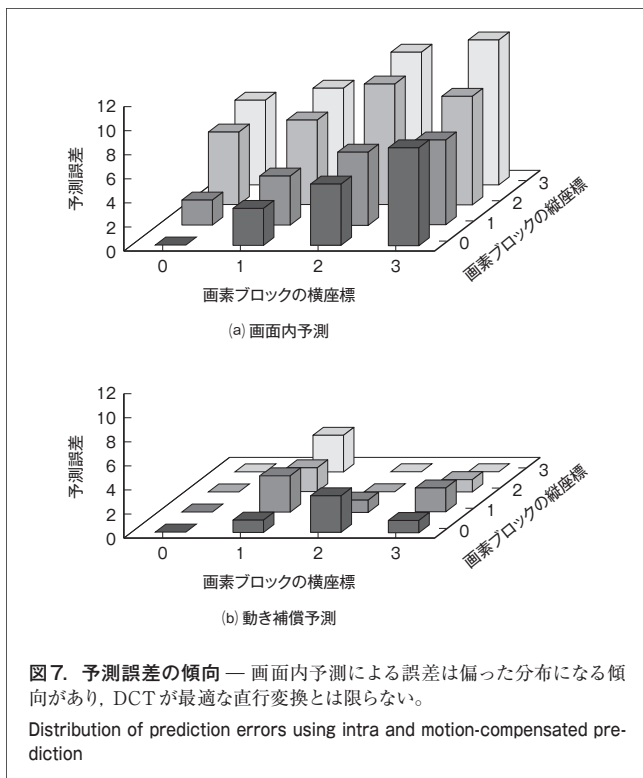
HEVCの画面内符号化では、画像の空間相関を利用して画面内予測と直交変換を行うことで圧縮率を向上させている。画面内予測は、参照画素の平均値を予測値として用いる平均値予測、予測画素位置に応じて参照画素の内挿処理を行う平面予測、及び参照画素を最大33方向に外挿する方向予測の合計35モードがある。エッジや平坦部などの局所的な特徴

に基づいて最適なモードを選択することで、空間方向の画像の冗長性を削減できる。直交変換は、図1に示した量子化や可変長符号化と組み合わせて、画面内予測や動き補償予測で発生した予測誤差を圧縮しやすい信号に変換する技術である。現行規格の直交変換では、予測誤差に対して水平・垂直方向に離散コサイン変換(DCT: Discrete Cosine Transform)を行い、画像信号を2次元の周波数成分に変換している。DCTは、自然画像などの主に低周波数成分から構成される画像に対して有効な変換手法であり、空間相関の高い画像に対してカルーネン・レーベ変換(KLT: Karhunen-Loeve Transform)の近似として扱えることから、従来の動画像符号化規格で広く用いられてきた直交変換である。

しかし、画面内予測によって発生する画素ブロック内の予測誤差の分布は、空間相関の影響で偏る傾向を持っており、必ずしもDCTが最適な直交変換とは限らない。例えば、動き補償予測によって発生する予測誤差の分布は画素位置によらずほぼ一様である(図7(b))。一方で、画面内予測によって発生する予測誤差は参照画素からの距離が大きくなるにつれて増大する傾向を持っている(図7(a))。

この偏りを利用して、画面内予測のモードごとに最適な変換行列をあらかじめ設計しておき、モードごとに切り替えることによってDCTよりも圧縮しやすい信号に変換できる。

当社は、直交変換行列を1種類追加するだけで符号量を削減する新直交変換手法を開発した⁽²⁾。開発手法では、画素ブロック内の予測誤差の傾向を次の二つに分類する。



- (1) 画素ブロック内で予測誤差が一様
- (2) 画素ブロック内で予測誤差が変化

(1)の傾向を示す場合にはDCTを、(2)の傾向を示す場合にはあらかじめKLTに基づき設計した変換行列を適用する。このように、画面内予測の予測誤差の傾向に従って直交変換を切り替えることで、符号量を削減しながらメモリの増大を抑えることが可能になる。

開発手法をHMに組み込んで符号量削減効果を評価したところ、全モードに対してDCTだけを用いる場合と比べて、同一画質で2%の符号量が削減できた。

4 あとがき

H.264/AVCがタブレットやTVなどのデジタル機器に幅広く浸透し、映像を身近に、簡単に視聴できる環境が整いつつある。動画像符号化の新規格HEVCの実用化により、この流れが更に加速し、今までになかった新しいサービス形態やアプリケーションへの展開が期待される。

当社は、今後もHEVC標準化への参画企業として積極的に活動していくとともに、HEVCを用いた新しい製品やサービスに向けた研究開発を継続していく。

文献

- (1) Tanizawa, A. et al. "Multi-directional Implicit Weighted Prediction Based on Image Characteristics of Reference Pictures for Inter Coding". Proc. 2012 IEEE International Conference on Image Processing (ICIP 2012). Orlando, FL, USA, 2012-10, IEEE, 2012, p.1545 - 1548.
- (2) Yamaguchi, J. et al. "One-Dimensional Directional Unified Transform for Intra coding". Proc. ICIP 2011. Brussels, Belgium, 2011-09, IEEE, 2011, p.3681 - 3684.



谷沢 昭行 TANIZAWA Akiyuki

研究開発センター マルチメディアラボラトリー研究主務。
画像信号処理及び動画像符号化処理の研究・開発に従事。
映像情報メディア学会会員。
Multimedia Lab.



山口 潤 YAMAGUCHI Jun

研究開発センター マルチメディアラボラトリー。
画像信号処理及び動画像符号化処理の研究・開発に従事。
Multimedia Lab.



中條 健 CHUJOH Takeshi, D.Eng.

研究開発センター マルチメディアラボラトリー主任研究員、
博士(工学)。動画像符号化処理の研究・開発に従事。電子
情報通信学会、映像情報メディア学会、IEEE会員。
Multimedia Lab.