

次世代動画像符号化の標準化に向けた フィルタ処理技術 BALF

Block-Based Adaptive Loop Filter toward Next-Generation Video Coding Standardization

渡辺 隆志

中條 健

■ WATANABE Takashi

■ CHUJOH Takeshi

動画像符号化の次世代規格であるHEVC (High Efficiency Video Coding) の標準化が、ISO/IEC (国際標準化機構/国際電気標準会議) のMPEG (Moving Picture Experts Group) とITU-T (国際電気通信連合-電気通信標準化部門) のVCEG (Video Coding Experts Group) の共同組織であるJCT-VC (Joint Collaborative Team on Video Coding) によって開始された。HEVCでは、現行の最新規格であるH.264/AVC (Advanced Video Coding) の2倍の圧縮性能を目標に掲げている。

東芝は、HEVCに向けた新たなフィルタ処理技術として、ブロック適応ループフィルタ (BALF: Block-Based Adaptive Loop Filter) 方式の開発を行ってきた。BALF方式はVCEGの共通参照ソフトウェアに採用されるなど、今後、HEVCの規格策定に貢献することが期待されている有力技術の一つである。

The standardization of High Efficiency Video Coding (HEVC) toward the next-generation video coding standard has been started by the Joint Collaborative Team on Video Coding (JCT-VC), which has been set up by the International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) Moving Picture Experts Group (MPEG) and the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) Video Coding Experts Group (VCEG). The objective of HEVC is to provide double the compression performance of the current H.264/Advanced Video Coding (AVC) standard.

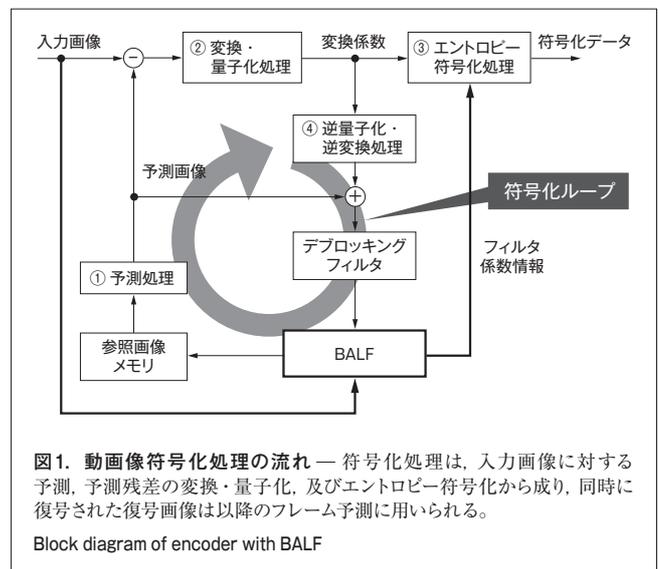
Toshiba has developed a block-based adaptive loop filter (BALF) as a new in-loop filtering technology for HEVC. BALF has already been adopted in the reference software for feasibility studies used by VCEG, and is expected to contribute to HEVC standardization.

1 まえがき

近年、デジタル放送の普及、インターネットの高速化、及びストレージの大容量化に伴い、高解像度 (HD: High-Definition) の動画像を用いたアプリケーションがより身近になってきている。一方、4K2KなどのUltra-HD化や、モバイルアプリケーションでの映像のHD化を見据え、より高効率な動画像符号化技術の要求が高まりつつある。

このような状況を踏まえて、ISO/IECのMPEGとITU-TのVCEGは、これまで別々に活動してきたH.264/AVCに続く次世代の動画像符号化の標準化活動を共同で行うために、JCT-VCと呼ばれる組織を結成した。2010年4月に第1回会合を開催し、HEVC (High Efficiency Video Coding) プロジェクトを本格的に開始した。

ここでは、東芝が次世代の動画像符号化技術としてVCEGで提案を行ってきたブロック適応ループフィルタ (BALF: Block-Based Adaptive Loop Filter) 方式^{(1), (2)}の技術について述べる。この方式は、VCEGの共通参照ソフトウェアに採用されるとともに、HEVCの技術提案募集でも東芝以外の多くの提案で採用されており、次世代の動画像符号化の標準化で有力な技術の一つである。



2 BALF方式を用いた動画像符号化

2.1 動画像符号化とループフィルタ

図1は、BALF方式を含む動画像符号化処理の流れを示すブロック図である。エンコーダに動画像が入力されると、

1フレームずつ符号化処理が行われる。まず、入力されたフレームに対して周囲の画素や以前のフレームの画素を利用した予測処理を行い、入力画像と予測画像の差分値を算出することでデータ量を削減する(図1の①)。更に、得られた差分信号を変換・量子化処理によって粗い精度に丸め(図1の②)、その後、可変長符号化によって数値を効率良く表現することで、データ量を削減することができる(図1の③)。デコーダでは、符号化データから符号化処理とは逆の手順で復号処理を行って画像を再構成することで、復号画像を得る。ただし、量子化で丸め処理が行われているために元の数値を完全に復元することはできず、入力画像と復号画像の間には誤差が発生する。

エンコーダでは符号化処理と同時にデコーダと同様の復号処理が行われ、得られた復号画像は以降のフレームの予測に用いられる(図1の④)。復号画像を参照して以降の予測を行う処理を符号化ループと呼び、この処理によりエンコーダとデコーダは同一の予測画像を生成することができる。ただし、符号化処理で発生した誤差についても予測処理によって以降のフレームに伝搬してしまう。この問題を解決するための技術としてループフィルタがある。

画像処理におけるフィルタ処理とは、周囲の画素値とフィルタ係数との積を算出し、それらの和を新たな画素値とするものであり、フィルタ係数を変更することで、意図的に画像をぼかしてノイズを除去したり、逆に画像の鮮鋭感を向上させたりすることができる。また、画像に対して適切なフィルタを適用することで画質が向上することが知られている。ループフィルタは符号化ループ内でフィルタ処理を行うため、復号画像の画質が向上するだけでなく、画質の向上した復号画像を以降の予測で用いることで、予測効率も向上する。

当社が開発したBALF方式は、入力画像と復号画像との誤差を低減することを目的としたループフィルタである。BALF方式では、エンコーダで入力画像と復号画像の誤差を低減する線形ウィナーフィルタを設計する。BALF方式を導入することで、復号画像の誤差が低減するとともに以降のフレームの予測効率が向上し、高い圧縮性能を実現することができる。更に、画像を複数のブロックに分割してフィルタ処理を適応的に制御することで、より高い圧縮性能を実現した。以下、BALF方式の詳細について述べる。

2.2 ウィナーフィルタの設計

ウィナーフィルタは、画像中に含まれる誤差を統計的に低減することができるフィルタとして知られており、ノイズ除去など劣化画像の復元で用いられている。当社は、このウィナーフィルタを符号化処理に導入した。一般的な劣化画像の復元では正解となるべき画像が存在しないため、統計的な性質からフィルタを設計するが、符号化処理では正解となる入力画像が存在するため、フィルタ処理後の画像と入力画像との誤差を

必ず最小にすることができる。このとき、最小化すべき誤差は式(1)で表すことができる。

$$E = \sum_x \sum_y (I_{x,y} - D'_{x,y})^2 \quad (1)$$

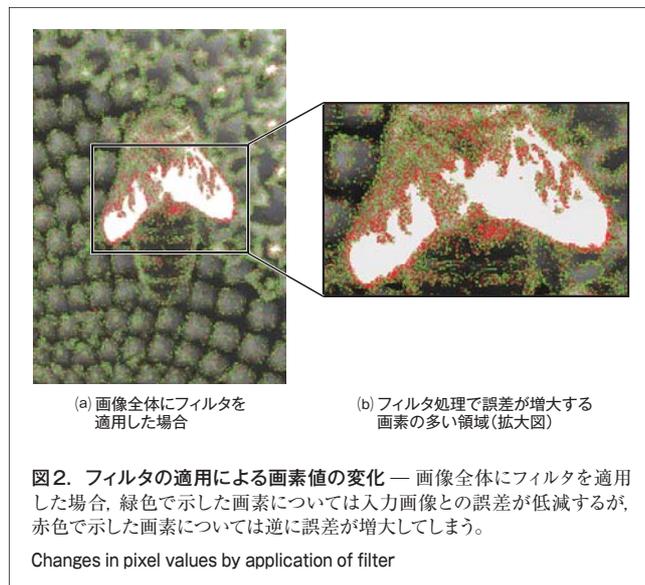
$$= \sum_x \sum_y \left(I_{x,y} - \sum_i \sum_j h_{i,j} D_{x+i, y+j} \right)^2$$

ここで、 $I_{x,y}$ 、 $D_{x,y}$ 、及び $D'_{x,y}$ は、入力画像、復号画像、及びフィルタ処理後の画像における画素 (x, y) それぞれの画素値を表しており、 $h_{i,j}$ はフィルタ係数を表している。

$h_{i,j}$ は式(1)から算出することができ、エンコーダでは $h_{i,j}$ と $D_{x,y}$ を用いてフィルタ処理を行って $D'_{x,y}$ を算出する。一方、デコーダでは入力画像を利用することができないため、 $h_{i,j}$ を算出することはできない。そこで、エンコーダは $h_{i,j}$ を符号化し、従来の符号化データに追加してデコーダに伝送する。

2.3 ブロック分割を用いたフィルタ処理の適応制御

画面内の全画素に対してフィルタを設計した場合、画像全体の誤差は最小となる。ただし、フィルタを適用することで誤差が増大してしまう画素も存在する(図2)。これらの画素が以降のフレームの予測で多く利用されると予測画像の誤差も増大するので、予測効率向上のためにループフィルタを導入したにもかかわらず、逆に予測効率が低下することもある。



例として、通常のH.264/AVCで符号化した場合とループフィルタを導入した場合の復号画像の客観画質を図3に示す。ここで、客観画質が高いとは、入力画像と復号画像との誤差が小さいことを意味している。横軸はフレーム番号を示しており、縦軸は客観画質の高さを表す指標であるPSNR(Peak Signal to Noise Ratio)を示している。PSNRが高いほど入力画像との誤差が少なく、客観画質が高いことを意味する。画面全体に対してフィルタを適用し続けた場合、先頭フレームで

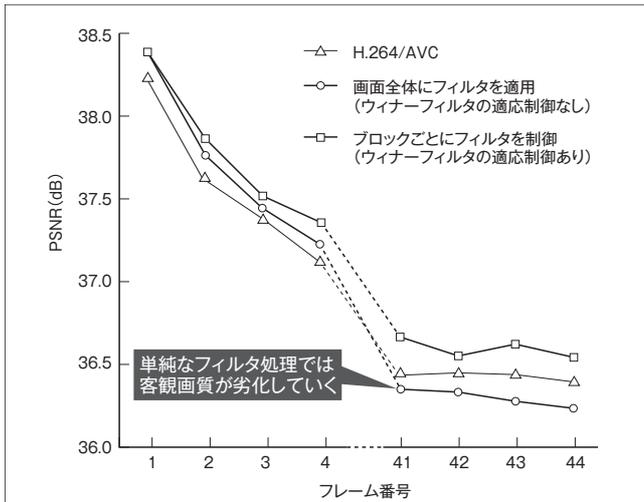


図3. フィルタ処理の制御による客観画質の向上 — 画像全体にフィルタを適用し続けると画質が低下するが、ブロックごとにフィルタの適用を制御すると画質が向上する。

Improvement in objective picture quality according to filter control for blocks

は客観画質が向上するが、後続のフレームになるほど客観画質が低下し、ループフィルタ導入前よりも客観画質が低下してしまう。これは、フィルタ処理により入力画像との誤差が増大した画素を以降の予測で多く参照したことで、誤差が伝搬したためである。

フィルタの適用により誤差が低減する画素にだけフィルタを適用すれば、このような問題を回避することができる。ただし、エンコーダでは入力画像を利用してフィルタを適用した場合の誤差の増減を判定できるが、デコーダではこの判定ができないため、フィルタ係数情報と同様に、フィルタ処理を制御するための情報を符号化データに追加しなければならない。

しかし、画素ごとにこれらの情報を追加する場合、一般的な8ビット入力画像では、フィルタ処理の制御情報だけで画像の1/8のデータ量が必要となる。一方、動画像符号化では画像のデータ量を数十分の1から数百分の1まで圧縮するため、このような制御をすることは現実的ではない。そこでBALF方式では、画像を所定のサイズを持つブロックに分割し、ブロックごとにフィルタ処理の制御を行う(図4)。

ここで、式(2)によって、 n 番目のブロック B_n に対してフィルタ制御情報である f_n を得ることができる。すなわち、 f_n が1となる場合にはフィルタを適用し、 f_n が0となる場合にはフィルタを適用しない。エンコーダは f_n の値を符号化データに追加し、デコーダに伝送する。

$$f_n = \begin{cases} 1 & \left(\sum_{x,y \in B_n} \{D_{x,y}^i - I_{x,y}\}^2 < \sum_{x,y \in B_n} \{D_{x,y} - I_{x,y}\}^2 \right) \\ 0 & \left(\sum_{x,y \in B_n} \{D_{x,y}^i - I_{x,y}\}^2 \geq \sum_{x,y \in B_n} \{D_{x,y} - I_{x,y}\}^2 \right) \end{cases} \quad (2)$$

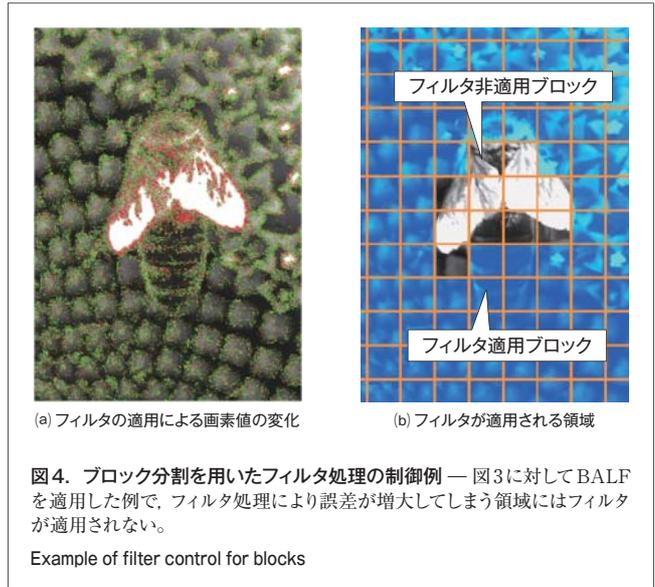


図4. ブロック分割を用いたフィルタ処理の制御例 — 図3に対してBALFを適用した例で、フィルタ処理により誤差が増大してしまう領域にはフィルタが適用されない。

Example of filter control for blocks

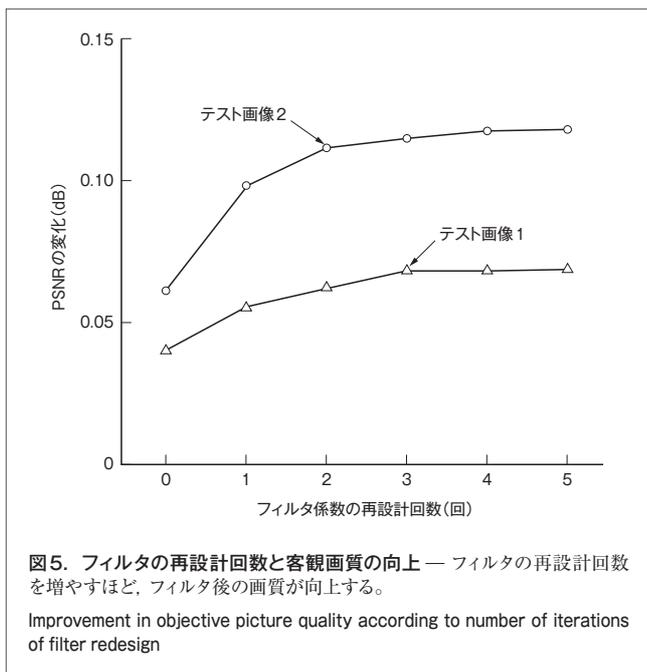
このとき、ブロックのサイズによって誤差の低減効果と送信すべきフィルタ制御情報のデータ量は変化し、各フレームで最適なブロックサイズは異なる。そのため、BALF方式ではフレーム単位でブロックサイズを変更している。各フレームでのブロックサイズは、所定のブロックサイズ候補から、動画像符号化で一般的に用いられているRD (Rate Distortion) コストに基づいて決定される。RDコストは、入力画像との誤差の大きさと発生するデータ量の重み付き加算で表されており、RDコストのもっとも小さくなるブロックサイズを選択することで客観画質とデータ量の最適なバランスをとった処理が実現できる。

ここで再び図3に着目すると、ブロックごとにフィルタ処理を制御した場合には、すべてのフレームで高い客観画質を示しており、ブロックによるフィルタ処理の制御が有効であることが確認できる。

2.4 フィルタの再設計

BALF方式では、まず画面全体に対してウィナーフィルタを設計して適用し、入力画像との誤差が低減するブロックにだけフィルタを適用するように制御を行う。ただし、このときのフィルタは画面全体に対して設計されたものであり、実際にフィルタを適用する領域に最適化されたものではない。そこで、フィルタを適用する領域だけでフィルタの再設計を行う。更に、ここで設計されたフィルタを用いてフィルタを適用するブロックを再度判定し、またフィルタを設計するというプロセスを繰り返すことで、フィルタ後の客観画質が向上することが確認されている。

図5は、2種類の画像に対してフィルタの再設計回数とPSNRの関係を示しているが、どちらの画像でも、再設計回数が増加するほどフィルタ後の画像のPSNRが向上していることが確認できる。



3 BALF方式による圧縮率の向上

BALF方式をH.264/AVCに導入して、実際に符号化を行った場合の効果について述べる。

表1は、様々なHD画像に対して符号化を行い、H.264/AVCで符号化した場合と比較して、同一客観画質を実現するために必要なデータ量をどの程度削減できているかを示したものである。ここではフィルタの再設計回数は3回としている。

これらの画像に対しては、BALF方式を導入することで最大33.9%、平均で22.3%のデータ量を削減できる。現在JCT-VCで標準化が行われているHEVCでは、H.264/AVCの半分の

表1. H.264/AVCと比較したBALFによるデータ量削減率
Bitrate reduction using BALF compared with H.264/AVC

画像名	画像サイズ (画素)	データ量削減率 (%)
Raven	1,280×720	19.4
ShuttleStart		16.2
SpinCalendar		13.4
Plane	1,920×1,080	33.9
Waves		25.5
BlueSky		22.1
Sunflower		31.2
ToysandCalendar		16.4
平均		22.3

データ量で同等の客観画質を目指している。すなわち、様々な方式を導入することで50%のデータ量を削減することを目標としており、BALF方式は次世代の標準規格に向けて有効な技術であると言える。

4 あとがき

ここでは、次世代動画画像符号化の標準化に向けて当社が開発を行っているBALF方式について述べた。

BALF方式は、ウィナーフィルタをループフィルタとして用い、ブロックごとにフィルタ処理を制御することで、復号画像の客観画質を向上させるとともに、入力画像との誤差がそれ以降のフレームの予測に伝搬することを抑制する。

この方式は、VCEGで用いられている共通参照ソフトウェアに採用されており、また、HEVCの技術提案募集でも多数の提案に含まれている注目技術の一つである。当社がHEVCの標準化に貢献できるよう、更に検討を進めていく。

文献

- (1) Watanabe, T., et al. "In-loop filter using block-based filter control for video coding". IEEE International Conference on Image Processing 2009 (ICIP2009), Cairo, Egypt, 2009-11, IEEE Signal Processing Society, 2009, p.1013-1016.
- (2) Chujo, T., et al. "Improvement of block-based adaptive loop filter". ITU-T SG16/Q.6. San Diego, 2008-10, VCEG-AJ13.



渡辺 隆志 WATANABE Takashi

研究開発センター マルチメディアラボラトリー。
動画画像符号化技術に関する研究・開発に従事。電子情報通信学会会員。
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、
工博。動画画像符号化技術に関する研究・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.