

2F-2 イントラ符号化にベクター表現を用いた動画像圧縮に関する検討

A Study on Vector Representation on Intra-coding of Video Compression

河村 圭 山本 勇樹 渡辺 裕

Kei KAWAMURA Yuki YAMAMOTO Hiroshi WATANABE

早稲田大学大学院 国際情報通信研究科

Graduate School of Global Information and Telecommunication Studies, Waseda University.

1 まえがき

エッジや均等色領域を多く含む人工的な動画像に、自然画を対象とする符号化を適用すると著しく品質が低下する可能性がある。イントラ符号化に着目すると MPEG ではモスキートノイズや偽色が知られているが、H.264/AVC [1] では均等色領域内のテクスチャ損失やエッジ方向の変化などの品質低下が見られる。

我々は以前より人工的な画像に対してベクター表現を用いる符号化を提案している [2, 3]。ベクター変換技術は低解像度画像における性能が向上し、符号化への適用が現実的になりつつある。計算機の性能向上により、グラフ理論や動的計画法の適用が複雑な画像に対しても実用的になったことが背景に挙げられる。従来のベクター変換の対象画像であったフォントや文書図形と比較して、低解像度で複雑な画像に対してグラフ理論を用いる高品質なベクター表現を生成するアルゴリズムが提案されている [4]。

本稿では、上記のイントラ符号化における品質低下の原因を明らかにすると共に、エッジ領域の符号化にベクター表現を利用して品質低下を抑制する手法を検討する。特に、人工的な画像においてはアンチエイリアシング処理の考慮が必須であり、線画や均等色領域の境界（エッジ）を単純にベクトル化するだけでは品質低下を招く。そこで、ベクター表現により最適な記述が可能なエッジモデルを提案する。

2 フレーム内符号化における品質低下と原因

H.264 のフレーム内符号化では、DCT と量子化の前にイントラ予測が導入された。その結果、テクスチャとエッジの消失による品質低下が発生する。なお、JPEG や MPEG などにおいても DC 成分や AC 成分のフレーム内予測が利用されている。これらはロスレス処理であるため品質には直接影響しない。

H.264 で導入されたイントラ予測は、符号化済みの周辺画素を用いて 4×4 または 16×16 ブロック単位で注目ブロックの予測を行う。均等色領域や 8 つの予測方向に一致するグラデーション、エッジは非常に高い精度で予測が当たり、ほとんど差分信号が発生しない。

均等色領域やグラデーション領域内のテクスチャは差分信号として扱われ、 4×4 DCT と量子化が行われる。 4×4 DCT は 8×8 DCT に比べて周波数分解能が低く、H.264 の粗い量子化により各成分がゼロになりやすい。その結果、テクスチャが消失して、のっぺりした画像となり主観品質が低下する。

予測の当たらなかったエッジを含むブロックでは、差

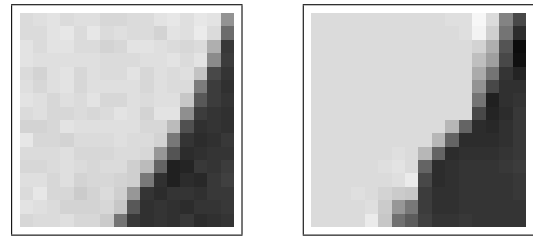


図 1 エッジ領域（入力画像，復号画像（ 16×16 画素， $qp = 32$ ））

Fig. 1 Edge region [Original, Reconstructed of 16×16 pixels at $qp = 32$]

分信号にもエッジが含まれる。これらのブロックでは自己相関は低く、高周波成分を多く含んでいる。さらに、DCT と量子化により高周波成分は消失、または増幅される。特にエッジが消失した際には、イントラ予測によって生成されたエッジのみが再現される。その結果、エッジ方向が変化してブロック間でエッジの連続性が失われ、主観品質も低下する。図 1 にエッジ領域を含む画像の拡大図を示す。なお、違いを明確にするためにコントラストを上げている。テクスチャが失われるだけでなく、エッジの方向が不連続になる様子が確認できる。

3 提案手法

3.1 ベクター表現の導入

テクスチャの消失に対しては、Fidelity Range Extensions (FRExt) [5] において 8×8 DCT 及び重み付き量子化が導入され、低減が図られている。ただし、 8×8 変換にはモスキートノイズの発生が伴う。

エッジ方向の変化に対しては、イントラ予測に起因するので予測方向を多様化する手法が考えられる。しかし、参照可能画素が無い場合には予測画像生成が不可能であり、さらにブロック内に異なる方向のエッジが内包される場合への拡張が困難である。そこで、エッジの位置をベクター表現に変換し、画素値を分離する手法を提案する。本手法は、モスキートノイズの低減も可能であるため、 8×8 DCT 導入に対する問題も解決可能である。

エッジは垂直方向の自己相関が低いいため、離散コサイン変換やアダマール変換など高い自己相関を仮定した変換は効率が悪い。また、エッジに沿って領域を分割する手法が考えられる。しかし、アンチエイリアシング処理により、正確な領域を定義することは困難である。そこで、エッジが持つ高い周波数成分を画素領域で分離し、エッジ位置とアンチエイリアシング処理に利用するフィルタのタップ数により記述する。高周波成分は、エッジと平行方向に高い相関を有するため、ベクター表現によるエッジ位置とフィルタ長のみで記述可能である。入力

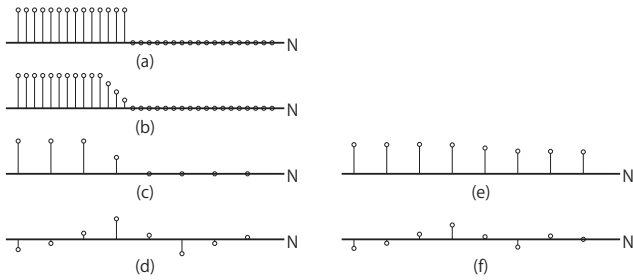


図2 提案するエッジモデル(左)と入力画像に対する高周波成分抽出(右)

Fig.2 Proposed edge model and high-frequency-component of an input image.

画像と復号された高周波成分の差分には、低周波成分やテクスチャが含まれる。

3.2 エッジモデル

ベクトル化の逆処理であるラスタライズにおいて、アンチエイリアシング処理の簡易手法として、オーバーサンプリングとその平均値を採用するアルゴリズムがある。これは、画素内の複数の位置で輝度値を計算し、それらの平均を注目画素の輝度値とする手法である。

これをモデル化した手順を以下に述べる。まず、 n 倍のオーバーサンプリングを行い、アンチエイリアス処理として m タップの平均値フィルタを導入する。ダウンサンプリングにより目標のサンプリング結果を得る。次に、低周波成分は DCT を適用することを考慮し、DCT によるハイパスフィルタを導入する。図 2(a)~(d) に一連の流れを示す。図中では $n = 4, m = 4$ とし、左から 3 画素 (14 サンプル目) にエッジがあると仮定する。本モデルは、エッジ位置と平均値フィルタのタップ数 m によってエッジを構成する高周波成分を記述可能である。

一方、入力画像からエッジ位置とタップ数の算出には以下の手順に従う。入力画像に対してモデルと同じハイパスフィルタを適用して高周波成分を得る(図 2(e), (f))。エッジの最尤位置やタップ数を変えたモデルと入力画像の高周波成分について相関を取り、エッジの位置とタップ数を決定する。なお、モデルでは複数のタップ数を考慮しているが、予備実験により n 倍のオーバーサンプリング時には、 $m = n$ タップの平均値フィルタのみを用いてエッジの位置を得られることが確かめられている。同時に、タップ数を大きくするとエッジ位置の不確かさが高くなるため、後段のベクトル化における誤差が大きく許容されることになる。

本モデルは、エッジ位置をベクター表現に変換するとともに、アンチエイリアシング処理をタップ数により調整できる。従って、拡大表示時にはタップ数を小さくして高周波成分を残し、主観品質の向上が可能である。

4 実験と考察

提案したエッジモデルが、入力画像に対して適切に選択されるか確認する実験を行う。本実験では、 8×8 DCT

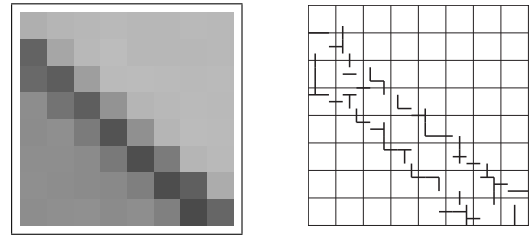


図3 エッジの検出例(入力画像と検出されたエッジ位置)

Fig.3 Experimental results (input images and detected edges).

を利用し、4 倍でオーバーサンプリングするモデルを利用した。また、計算量を抑えるために 8 画素内に含まれるエッジの本数は 2 本までと仮定した。

モデルは 1 次元のまま用い、画像に対して行、列それぞれに対してエッジ位置の推定を行った。また、平面としての連続性を考慮するために、隣接する行(列)を用いて補完した行(列)も利用した。

入力した画像と、エッジモデルを適用して得られたエッジ位置を図 3 に示す。相関係数の絶対値が 0.85 以上、振幅が大きい時に有意であるとしてエッジの位置を記述した。4 倍でオーバーサンプリングして計算したため、 $1/4$ 画素精度で位置が推定されていることが確認できる。線幅 1 画素程度の線画がアンチエイリアシング処理によりぼやけていても、入力画像から主観的に得られるエッジの位置を正確に推定している。

5 むすび

本稿では、H.264 におけるエッジ消失について原因を明らかにするとともに、エッジ領域の符号化にベクター表現を適用する手法を検討した。アンチエイリアシング処理を考慮し、オーバーサンプリングと平均フィルタによるエッジモデルを提案した。実験により、提案したモデルがアンチエイリアシング処理を考慮してエッジを推定できることを確認した。

参考文献

- [1] ITU-T Recommendation H.264 IS, "Advanced Video Coding for Generic Audiovisual Services," 2003.
- [2] O. Nakagami, T. Miyazawa, H. Watanabe, H. Tominaga, "A Study on two-layer coding for animation images," IEEE International Conference on Multimedia Expo (ICME) 2002, WedAmPO3: Compression II, Aug. 2002.
- [3] 河村, 山本, 渡辺, "動画像圧縮におけるベクター表現の符号化に関する検討", 情処研報 2005-AVM-51, no.15, Dec. 2005.
- [4] "potrace," <http://potrace.sourceforge.net/>
- [5] G. Sullivant et al., "Draft Text of H.264/AVC Fidelity Range Extensions Amendment," JVT-L047d12, Jul. 2004.