

ブロックパターン情報に基づいたオブジェクトベース符号化

小野 謙治 浜辺 隆二

福岡工業大学 情報工学部 情報工学科

〒811-0295 福岡市東区和白東 3-30-1

TEL 092-606-3131

E-mail hamabe@fit.ac.jp

あらまし 本稿では、ブロックパターン情報を用いたオブジェクト生成と可変ブロック学習動きベクトル量子化 (VLMVQ) による符号化を組み合わせたオブジェクトベース符号化法を提案する。オブジェクト生成においては、可変ブロック内のエッジ情報をブロックサイズごとに検出し、そのときのエッジ情報とブロックサイズ情報を用いて分類とラベリングを行う。VLMVQ は、生成されたオブジェクトごとのコードブックからなるコードブック群が可変ブロックの性質を学習しながら符号化を行うので画質が良好で、汎用性のある符号化方式である。性能評価として、本オブジェクト生成法による領域分割の結果および VLMVQ と BMA の比較を SNR とビットレートの観点から考察する。

キーワード 領域分割, オブジェクトベース符号化, ブロックパターン情報

Object-based Coding based upon Block Pattern Information

Kenji ONO, Ryuji HAMABE

Department of Communication and Computer Engineering, Fukuoka Institute of Technology

3-30-1, Wajirohigashi, Higashiku, Fukuoka 811-0295, Japan

TEL +81-92-606-3131

E-mail hamabe@fit.ac.jp

Abstract This paper presents an object-based video coding, which combines object region merging technique using variable block patterns with VLMVQ (Variable-block Learning Motion Vector Quantization) for coding. In object region merging for segmentation, first, each variable block is detected whether it may be edge block or not. Next, each block is classified and labeled according to edge and block-size information. In VLMVQ, we prepare a codebook for each object. Since the codebooks learn the texture of variable blocks together with coding, quality may be improved. Experimental results show the image segmentation and comparison VLMVQ with BMA from the view point of SNR & bit rate.

key words Image segmentation, Objected-based coding, Block pattern information

1. はじめに

MPEG-1/2 等に代表される従来の動画画像符号化は、画像を一様なブロックに分割し、このブロック単位で動き補償等の処理を行っている。しかし、この方式は画像のテクスチャを反映した分割ではなく、単に一様なブロックに分割しただけなので、同じブロック内に複数の動きを有するブロックのある場合においては、画質の劣化が生じるという問題点がある。現在、このような背景を基に画像からオブジェクトを抽出し、オブジェクト単位で符号化を行うオブジェクトベース符号化方式が研究されている¹⁾²⁾。これは、画像内のテクスチャや動きといった情報によりオブジェクトを生成し、符号化を行う方式で、これは従来方式の問題点を解決することが可能である。

本稿では、まず、オブジェクト生成においては画像内のエッジ情報を可変ブロックサイズごとに検出し、そのときのエッジ情報の有無及びブロックサイズの情報からなるブロックパターン情報を用いて輪郭とオブジェクトに分類しラベリングを行い、更にそのラベル情報に基づく各オブジェクトの生成を行う手法を用いる。次に、オブジェクトベース符号化においては、ブロックパターン情報を基に生成された各オブジェクトに対して、可変ブロック学習動きベクトル量子化 (Variable block Learning Motion Vector Quantization: 以下 VLMVQ) による各オブジェクトに適したコードブック (動きのパラメータ) を用いて符号化を行う手法を用いる。我々は、上述したオブジェクトベースの符号化を実現するために二つの手法を組み合わせた方式を提案する。シミュレーション実験によって、オブジェクト生成された領域分割の結果を示すとともに、オブジェクトベース符号化の性能の評価として、SNR とビットレートの面から VLMVQ とブロックマッチングアルゴリズム (以下 BMA) を比較して考察する。

2. ブロックパターン情報に基づくオブジェクトベース符号化

本章では、従来方式のように一様に分割したブロック単位での符号化とは異なり、オブジェクトに基づいた符号化を行う。オブジェクトの生成においては、まず符合化対象

フレームに対してエッジの強調操作を行い、次にこのフレームに対して、可変ブロック単位での分割及びエッジ情報をを用いた分類を行うことでオブジェクトの生成を行い、最後に生成したオブジェクトごとに VLMVQ による動きベクトルのコードブックを用いて符号化を行う方式を提案する。ここで、VLMVQ はオブジェクト内の動きの性質を学習するように更新してゆくので、オブジェクトごとに設けた各コードブックは、各オブジェクトの動きの性質を反映することが可能で、汎用性があり、従来方式における一様なブロックの符合化に比較して、画質が改善されると考える。

2.1 オブジェクトの生成

オブジェクトの生成は、ブロックパターン情報に基づいて行う。本方式におけるブロックパターン情報とは、各ブロックのサイズ情報 S_i ($i=0,1,2,3$) 及びエッジの有無情報 E_j ($E_j: j=0,1$) から構成される。ここで、 S_i における $i=0,1,2,3$ の値はそれぞれ、 16×16 、 8×8 、 4×4 、 2×2 のブロックサイズを示し、また E_j における $j=0,1$ の値はそれぞれ、エッジ無し、エッジ有りを示している。ここで、ブロックサイズ 16×16 、 8×8 、 4×4 においては、エッジ情報が存在する場合、それぞれ 8×8 、 4×4 、 2×2 のブロックサイズに再分割を行うので、エッジ有りの情報を持つブロックパターンは存在しない。したがって、ブロックパターンはエッジ無しの 4 種類とエッジ有り (2×2) の計 5 種類となる。

図 1 にブロックパターン情報に基づくオブジェクト生成処理のブロック図を示し、以下にその処理手順を述べる。

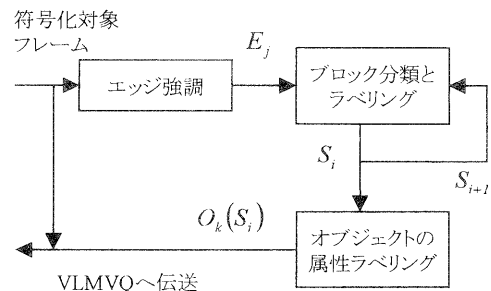


図 1 オブジェクト生成処理の流れ

- (1) エッジ抽出: 符号化対象フレームにグラディエント操作を適用して、エッジを強調する。これは、エッジの部分オブジェクトの輪郭部分と見なすためである。
- (2) ブロック分類とラベリング: (1)の処理を施した符号化対象フレーム画像に対して、 16×16 、 8×8 、 4×4 、 2×2 の4種類の変換ブロックサイズでトップダウン方式による分割を行う。さらに分割したブロックに対してエッジ検出を施し、その情報を基にブロックを分類する。分類の判別条件には、エッジの有無情報を用いる。エッジが検出された場合、ブロックのサイズ情報 S_i は S_{i+1} となる。図2にブロックパターン情報によるブロック分類の具体例を示す。図の全体の正方形を 16×16 のブロックサイズ、網掛け部分をエッジを含んだブロックとする。まず、最大ブロックサイズ 16×16 でエッジの検出を行い、エッジを含むと判断された場合は 8×8 ブロックサイズに再分割を行い、分割されたブロックごとにエッジの再検出を行う。この例では、左側の二つの 8×8 ブロックはエッジを含んでいないので、ブロックパターン情報 $S_1 \cdot E_0$ を与える。右の二つの 8×8 ブロックはエッジを含んでいると判断されるので、 4×4 ブロックサイズで再分割を行い、エッジ検出を行うが、右下半段の二つのブロックはエッジを含んでいないので、ブロックパターン情報 $S_2 \cdot E_0$ を与える。残ったブロックは全てエッジを含んでいると判断されるので、更に 2×2 ブロックサイズで再分割を行う。このときのブロックサイズが最小なので、これによりエッジを含まないブロックにはブロックパターン情報 $S_3 \cdot E_0$ を与え、エッジを含んだブロックにはブロックパターン情報 $S_3 \cdot E_1$ を与えることで、このブロックにおける分類を終了する。
- (3) オブジェクトの属性ラベリング: (2)の手順を繰り返すことでオブジェクトの輪郭を含む部分とオブジェクト部分とに分類することができるが、オブジェクトとして分類された部分に対して、このブロックパターン情報だけではどのオブジェクトに属するのかという識別情報としては不十分である。ここでは、エッジの有無情報を表す数値を基にオブジェクト情報 $O_k(S_i) (i=0, 1, 2, 3, k=0, 1, \dots, K-1)$ のラベルを生成する。これ

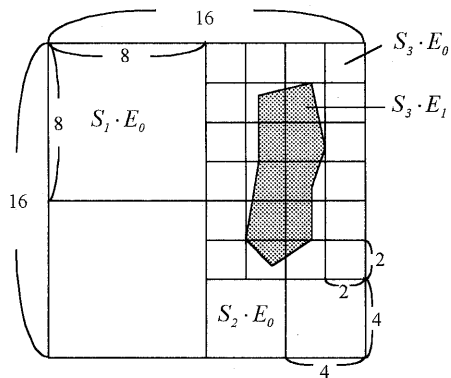


図2 ブロックパターン情報に基づく分類

により、各画素はどのオブジェクトまたは輪郭部分に属し、ブロックサイズはどのサイズでどのオブジェクトに属しているのかを認識することができる。

- (4) VLMVQへ伝送: 生成したオブジェクト情報 $O_k(S_i)$ と符号化対象フレームを VLMVQ の処理過程へ伝送する。

以上の手順によりオブジェクトを生成する。

2.2 可変ブロック学習動きベクトル量子化(VLMVQ)

オブジェクトに基づいた符号化は、オブジェクト生成時に用いたオブジェクト情報 $O_k(S_i)$ のラベルに基づいて、VLMVQにより符号化を行う。

2.2.1 可変ブロック学習動きベクトル量子化

一般に動画像では、隣接するブロックの動きは時間的と空間的の両方において相関が高いことが知られている。従って、本手法では符号化対象ブロックの動きベクトル $\vec{\delta}$ を算出後、次に動き補償を行う対象ブロックにおいても $\vec{\delta}$ と類似した動きベクトルを持つと仮定し、 $\vec{\delta}$ に近いコードワードが出現するようにコードブック内の動きベクトルを更新する。つまり、オブジェクトごとにこのコードブックを用意した場合、各オブジェクト内の動きの性質をそれぞれのコードブックが学習するように更新してゆくの、一様なブロックによる符号化に比較して画質の改善が可能となると考えられる。また、本方式を可変ブロックサイズに対応さ

せたのは、前節 2.1 にて生成されるオブジェクトの画素がそれぞれ可変なブロックで構成されたオブジェクトのオブジェクト情報 $O_k(S_i)$ を有しているので、その情報を基に符号化を実行するには、符合化レベルでも可変にする必要があるからである。

図 3 に VLMVQ による処理のブロック図を示す。あらかじめ探索領域内の動きベクトルから適当に抽出した動きベクトルで構成されたコードブックをオブジェクトの数 K 個分を用意する。まず、オブジェクト情報 $O_k(S_i)$ により可変ブロックを生成する。次に、可変入力ブロックとそれに対応した一つのコードブック $C_l^k (l=0,1,2,\dots,L-1$: コードワードのインデックス) 内の動きベクトル δ_l^k に対応する量子化対象ブロックとの誤差が最小となる動きベクトル δ_l^k を選択し、そのときの予測誤差 d_{min} を求める。ここで、 d_{min} は平均二乗誤差である。この d_{min} とあらかじめ設定しておいた閾値 T_{hd} との比較を行い、 d_{min} が T_{hd} 以下のときはコードブックから選択された動きベクトル δ_l^k を採用し、 T_{hd} を超えた場合はコードブック内に最適な動きベクトルが存在しなかったとして BMA によるリフレッシュ操作で動きベクトル δ_{BMA}^k を算出し、これを採用する。次に、 δ_l^k または δ_{BMA}^k の性質を次期コードブックが学習するように更新する。ここで、図 3 の入力ブロックは可変ブロックだが、探索領域は入力ブロックのサイズに関係なく、 16×16 ブロックサイズのみ探索領域とする。理由は、探索領域を極度に小さくした場合は、探索領域外へと移動するオブジェクトが存在する場合において、BMA による動きの補正を用いてもコードブックをそのオブジェクトに対応させることが困難となるからである。

以上のことを反復することで、オブジェクトごとの動きの性質を学習するようにコードブックが更新される。

2.2.2 コードブックの学習³⁾

図 4 に VLMVQ におけるコードブックの学習過程のフローチャートを示す。本方式では、選択した最近傍参照ベクトル δ_l^k は学習せずに残したほうがより良い学習結果が得られると考え、コードワードに動きベクトルの使用頻度カウンタ $F_l^k(t)$ を設置し、このカウンタが最小のコードワード、即ち使用頻度の最も低いコードワードを学習する。

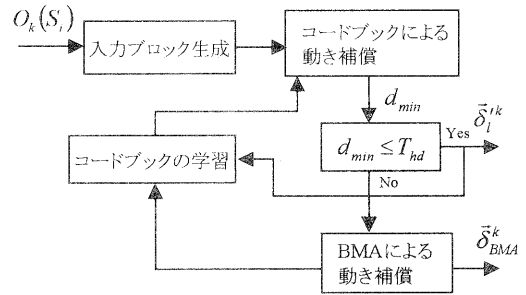


図3 VLMVQ の処理手順

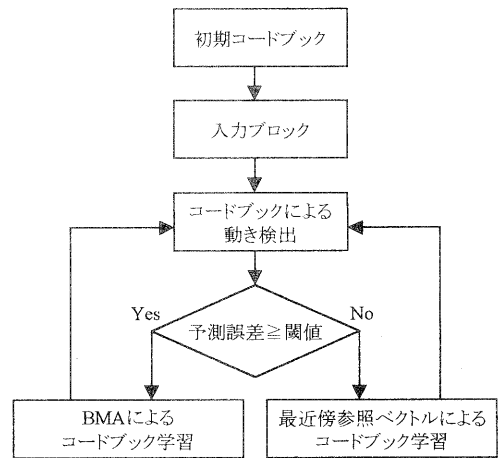


図4 VLMVQ におけるコードブックの学習過程

また学習としては、閾値の値によって選択された動きベクトルまたは BMA による動きベクトルでコードブックの更新を行う 2 通りを設ける。

この処理を反復することによって、最適な動きベクトルに似て、かつ次ブロックで選択される可能性の高いコードワードが複数個準備され、少ないコードワードでもオブジェクトごとに局所的に最適化された汎用性のある良好なコードブックを構成することができる。学習過程を以下に示す。

- (1) 符号化対象探索領域から適当に初期コードブックを用意する。
- (2) 入力ブロックに対して、コードブックによる動きの検出

を行い、誤差最小のブロックを示す動きベクトル $\vec{\delta}_i^{k'}$ とそのときの予測誤差を求める。このとき選択された動きベクトル $\vec{\delta}_i^{k'}$ のカウンタ $F_i^k(t)$ を+1 する。

- (3) 予測誤差とあらかじめ設定しておいた閾値とを比較する。このとき閾値を下回ったならば、最近傍参照ベクトル $\vec{\delta}_i^{k'}$ によるコードブックの学習を行い、閾値以上であれば、BMA によるコードブックの学習を行う。
- (4) 最近傍参照ベクトルによるコードブックの学習は、次式により最も使用頻度の低いコードワード $\vec{\delta}_m^k(t)$ を更新する。このときカウンタ $F_m^k(t)$ を+1 する。

$$\vec{\delta}_m^k(t+1) = \vec{\delta}_m^k(t) + \alpha(t) \cdot (\vec{\delta}_i^{k'}(t) - \vec{\delta}_m^k(t))$$

ここで $\alpha(t) = 1 - F_m^k(t) / F_i^{k'}(t)$ とする。 (1)

この更新係数 $\alpha(t)$ により、更新用ベクトル $\vec{\delta}_i^{k'}(t)$ の使用頻度が高い場合、 $\vec{\delta}_m^k(t)$ の近づく確立が高くなる。ただし、更新後のコードワード $\vec{\delta}_m^k(t+1)$ がコードブック内に存在するときは更新を行わない。これは、同じコードワードが連続して選択された場合に起こり易いコードブックの収束を防止するためである。

- (5) $\vec{\delta}_{BMA}^k$ で更新するときは、(1)式により $\vec{\delta}_m^k$ を更新し、カウンタ $F_m^k(t)$ を+1 する。このとき、あまり変化しないように更新係数 $\alpha(t)$ を $\alpha(t) = 1 / F_i^{k'}(t)$ とする。

3. シミュレーション実験

実験に用いた動画データは、Claire、Flower Garden (352×240 画素、グレイスケール、60フレーム) である。ブロックサイズは 16×16、8×8、4×4、2×2 画素の 4 種類からなる可変ブロックサイズで、探索領域は、それぞれ垂直・水平方向に±15 画素の半画素精度を用いる。予測は前方向予測のみとし、参照フレームは原画像である。各オブジェクトに与える VLMVQ のコードブックサイズは 4 ビット(コードワード数 16)とする。閾値は平均二乗誤差で、予備実験によって決定した $T_{hd} = 40$ を用いる。これは VLMVQ におけるブロックマッチング回数を 5%未満に抑えることができる値である。以上の条件で、まずブロックパターン情報による領域分割の結果を提示する。次に、ブロックパターン情報により生成したオブジェクトに対して、



図 5(a) 原画像



図 5(b) 領域分割(Claire)

そのオブジェクトに基づく符号化を VLMVQ と BMA とで行い、SNR 及びビットレートでの比較を行う。尚、実験を行った動作環境は、CPU: Pentium III (700MHz) /OS: Windows98 である。

3.1 ブロックパターン情報による分割結果

図4(a) (b)に Claire のブロックパターン情報による分割結果を示す。(a)は Claire の原画像、(b)は Claire の領域分割結果である。(b)の結果から、ブロックをベースに領域分割を行っているため輪郭部分がギザギザに成っているが、画像全体としては良好な分割を確認できる。

3.2 オブジェクトに基づく VLMVQ と BMA との比較

表1は Claire に対して、ブロックパターン情報によるオブジェクトを基に VLMVQ と BMA の符号化を行った際の SNR とビットレートの比較を示したものである。この表より Claire において VLMVQ は BMA に比較して、SNR では平均約

3.66dB の劣化が見られるが、ビットレートに関しては、平均約 0.155bpp と上回った値を示している。また、Flower Garden においても VLMVQ は BMA に比較して、SNR では平均約 3.21dB の劣化が見られるが、ビットレートに関しては、平均約 0.293bpp と上回っている。

表 1 各動画像の平均 SNR 及びビットレート

	SNR (dB)		ビットレート (bpp)	
	BMA	VLMVQ	BMA	VLMVQ
Claire	45.49	41.83	0.257	0.102
Flower Garden	30.72	27.51	0.488	0.195

図 6(a) (b)に最も SNR の差が著しいフレーム番号 13 の視覚評価を示す。(a)は BMA による再生画像で約 47.67 (dB)、(b)は本方式による再生画像で約 39.38 (dB)である。この画像においては、本方式が約 8.29 (dB)と下回っているが、視覚的には十分に良好な画像が得られている。

4. むすび

本稿では、ブロックパターン情報に基づくオブジェクトベース符号化として、オブジェクトの生成においては画像内のエッジ情報を可変ブロックサイズごとに検出し、そのときのエッジ情報の有無及びブロックサイズの情報からなるブロックパターン情報による分類とラベリングを行い、更にそのラベル情報に基づいて各オブジェクトを生成するという手法、及び生成した各オブジェクトに対して VLMVQ のコードブックを用いて符号化を行う手法を示した。符号化の性能評価としては、Claire 及び Flower Garden に対して、SNR 及びビットレートの観点から BMA との比較で行った。結果として、SNR に関してはそれぞれ約 3.66 (dB)と約 3.21 (dB)の劣化が見られたが、ビットレートに関しては、それぞれ約 0.155 (bpp)と約 0.293 (bpp)だけ上回った。また、視覚評価としては、SNR 上の劣化はあるが Claire 及び Flower Garden の両画像ともに良好な画質であることを確認した。

今回のシミュレーション実験では、小領域の除去に



図 6(a) BMA による再生画像
(SNR = 47.67dB)

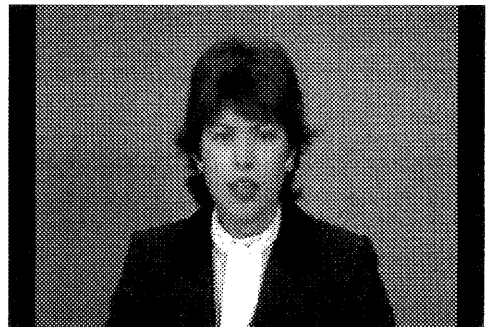


図 6(b) VLMVQ による最大劣化の再生画像
(SNR = 39.38dB)

対する処理を行っていないことから、エッジ量の多い画像に対してはコードブックの数がオブジェクトに比較して多く設置される可能性があり、この点について今後検証する予定である。

参考文献

- 1) Chee Sun Won: "A Block-Based MAP Segmentation for Image Compressions", IEEE Trans. on Circuits and Systems for Video Technology, Vol. 8, No. 5, pp592-601, 1998.
- 2) Roberto Castagno, Touradj Ebrahimi, Murat Kunt: "Video Segmentation Based on Multiple Features for Interactive Multimedia Applications", IEEE Trans. on Circuits and Systems for Video Technology, Vol. 8, No. 5, pp562-571, 1998.
- 3) 小野, 浜辺: "学習動きベクトル量子化におけるコードブックの性能評価", PCSJ'99, P-1. 14, pp. 29-30, 1999.