

グローバル動き補償符号化における条件付画素補填

秦泉寺久美 阪谷徹 八島由幸 小林直樹

NTTサイバースペース研究所

〒239-0847 神奈川県横須賀市光の丘 1-1

+81-468-59-3034

あらまし

MPEG-4にはグローバル動き補償やスプライト符号化といった新しい概念の映像符号化ツールが具備されている。特にスプライト符号化はある条件下では劇的な圧縮効率を達成することが報告されている。本稿では、グローバル動き補償符号化において、注目領域となりやすい前景動物体の画質の向上と、スプライト的な全体符号量削減を図り、条件付画素補填を適用するための前景背景分離方法を提案した。また、本アルゴリズムをカメラの動きのある画像に適用し、符号化実験を行った結果、前景動物体領域においては通常のGMCの場合に比べて0-2dB高いSNRを達成し、主観的品質も高まることが確認された。しかし、全体的な符号量の観点では、高い前景比率、画像によっては背景が過度に劣化にするなどにより、スプライト的な圧縮効果は期待できないことがわかった。

キーワード MPEG-4 条件付画素補填, グローバル動き補償, 映像圧縮符号化

Conditional Replenishment for Global Motion Compensation

Kumi Jinzenji, Toru Sakatani, Yoshiyuki Yashima, Naoki Kobayashi

NTT Cyber Space Labs.

1-1 Hikari-no-oka Yokosuka-Shi, Kanagawa 239-0847 JAPAN

+81-468-59-3034

Abstract

MPEG-4 include a several effective video compression tools, such as Global Motion Compensation and Sprite Coding. Specially, sprite coding is reported to achieve drastic coding efficiency compared to existing coding technologies under a certain condition. This paper proposed macro-block-based foreground/background separation algorithm for applying conditional replenishment to the global motion compensation. The goal of this study is; to improve the quality of moving objects, and to reduce the coding amount like sprite coding. We had experience on several video sequence using proposed algorithm. The SNR of foreground region is improved to 1-2dB compared to usual global motion compensation. However, it is found that drastic compression efficiency like sprite coding is not expected because of the too large foreground region and burred background image in some cases.

Keyword MPEG-4, Conditional Replenishment, Global Motion Compensation, Video Compression

1. はじめに

新しいオーディオビジュアル符号化標準 MPEG-4[1]では、従来の符号化方法 (H.26X, MPEG-1,2)と比較して強力な圧縮ツールが具備されている。その中にスプライト符号化やグローバル動き補償 (Global Motion Compensation: GMC) といった、大局的な動きをする部分を一つのグローバルな動きベクトル (Motion Vector: MV) で表現する方法がある。前者は筆者らが文献[2]にて、動画像を前景と背景に分離し、前景をオブジェクト符号化、背景をスプライト符号化することで、劇的な符号量削減をもたらすことを報告している。後者は、MVの全体の符号量に占める割合が比較的大きい低ビットレートでの符号化や、平行移動動きモデルが破綻するような動き (ズーム等) のあるシーケンスには大きな効果があることが知られている[2,3]。

一方で、従来の符号化方式は動き補償 (MC: Motion Compensation) と条件付画素補填 (Conditional Replenishment: CR) によって隣接フレーム間の情報を効率的に圧縮してきた。MC は隣接フレーム間において動きベクトル (Motion Vector: MV) を算出し、その差分の DCT 係数を符号化する。また、後者は、時間的に前後するフレームにおいて、その座標が同じマクロブロック (Macro Block: MB) 同士の差分がほとんど無い場合にマクロブロックの符号化をスキップするもので、テレビ電話など、固定カメラで撮影した映像の背景部分の情報圧縮に威力を発揮する。

また、これらの CR、スプライト符号化、GMC 符号化は、圧縮原理に共通する部分がある。GMC 符号化は直前フレームをスプライトとみなしたものであるし、スプライト符号化は一般的な CR をカメラの動きのある部分に拡張したものとも考えることができる。

本稿では、GMC 符号化に CR を導入し、人間の ROI に近い前景動物体により多くの符号量を割り当てて動物体の品質を向上させる方法を提案する。また、GMC に CR を適用した場合の

符号化効率や画質を文献[2,3]のスプライト符号化の符号化実験の結果も含めて考察を行う。以下、2章にて隣接フレームの情報をを用いた背景領域決定法を提案する。また3章では、GMC 符号化に CR を組み込んだ符号化アルゴリズムを提案し、符号化実験を行う。4章で結論を述べる。

2. 隣接フレーム間情報を用いた前景背景分離法の提案

CR の一番簡単な例は LME (Local Motion Estimation) を行った後、予測画像と対象画像の差分の小さい部分の差分を符号化しない方法である。大儀的には、LME 以外による予測誤差、例えば GME の予測誤差に関しても適用が可能である。差分の定義や符号化をするか否かの判断法についてはさまざまな方法が考えられる。映像をカメラの動きのみを反映する背景領域と、これ以外の動物体を反映する前景領域に分離できるとすると、背景領域を忠実に抜き出して CR を適用することが考えられる。さらに GMC や CR は MB 単位で行われるので、分離の精度は MB 単位でよいとする。また、GMC が隣接フレーム間の圧縮を強化するツールであるので、隣接フレーム間の情報のみを使って背景領域と前景領域を分離する方法を提案する。本節では、背景領域に GMC+CR を適用するための、隣接フレーム間情報を利用した MB 単位の前景背景分離法を提案する。

2.1 LMV と GMV の距離を用いた分離方法と問題点

GMV は画面全体の大局的な動きであるので、任意の MB における LMV と GMV の差 (大きさ、方向等) が小さい場合は、その MB が背景部分を反映していると予想される。しかし、LMV は MB 内の画素の輝度値に偏向がある場合 (例: 競技場の白線近辺) や、分散値が極端に小さい場合 (例: 空など) は誤抽出される傾向がある。よって、本来背景となるべき部分が前景と判断される場合がある。

2.2 GMC 後の差分を用いた分離方法と問題点

GMC を行った後の原画像の差分値を計算すると、背景部分には小さな差分、動物体には大きな差分が現れることが予想できる。しかし、細かいエッジがたつような部分（例：樹木では差分の値が大きくなる傾向にある。さらに、前景背景を切り分けるための閾値をどこに置くかは、画像によるところがあるという問題がある。

2.3 提案方法

本稿では、LMV と GMV の距離 D と、GMC による差分の双方を用いた前景背景分離法を提案する。MB 単位での分離であるので、GMC による差分には MB 毎に MAD (Mean Absolute Difference) 値を用いる。図 1 に、GMC+CR を行うための前景背景分離アルゴリズムを示す。

最初に、MB 単位に LMV を計算する。さらに、LMV の値から、文献[2,3,4]の手法によって GMV を算出する。この方法は従来法におけるトップダウン的アプローチ[5]とボトムアップ的アプローチ[6,7]の長所を生かした方法で、カメラの操作を代表する GMV が算出される。任意の MB における MV を (u,v) 、MB の中心座標を (x,y) 、そのときの文献[]における 4 パラメータからなる GMV を $\{a,b,c,d\}$ (各々、カメラのズーム、回転、パン、チルトを示す) とすると、以下の式 (1) が成り立つ。

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} a' & b \\ -b & a' \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} \quad (1)$$

$a = a' + 1$

GMV が既知であれば、式 (1) より各 MB における LMV に変換できる。この LMV を LMV' とする。座標 (x,y) における LMV' を (u',v') とする。ここで、GMV と LMV の距離 D を式 (2) のように定義する。

$$D = \sqrt{(u - u')^2 + (v - v')^2} \quad (2)$$

距離 D がある閾値 $Th1$ より小さい場合、その MB を背景候補 MB とする。この背景候補 MB 群を Map1 とする。さらに、各々の MB に

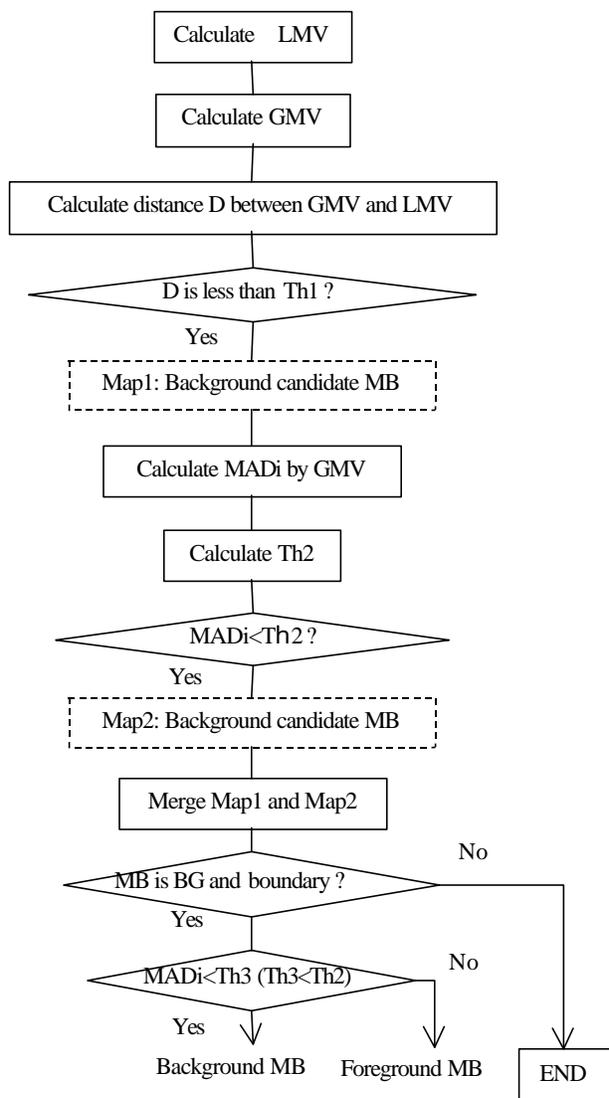


Fig.1. Overflow of MB-based foreground/background separation algorithm.

て GMC 後の予測画像と元の画像との MAD 値を計算する。ここで、Map1 群に属する背景候補 MB、その MAD が小さい方から順に全体の 90%程度のところあたりの MAD 値を $Th2$ とする。これで、MAD による差分閾値を自動設定したことになる。すべての MB にて、差分閾値 $Th2$ よりも小さい MB を背景候補 MB として Map2 にする。さらに Map1 と Map2 の OR を算出し、背景候補領域とする。ここで、背景候補領域と前景候補領域の境界に位置する背景候補 MB は微小に前景領域がある場合が考えられる。よっ

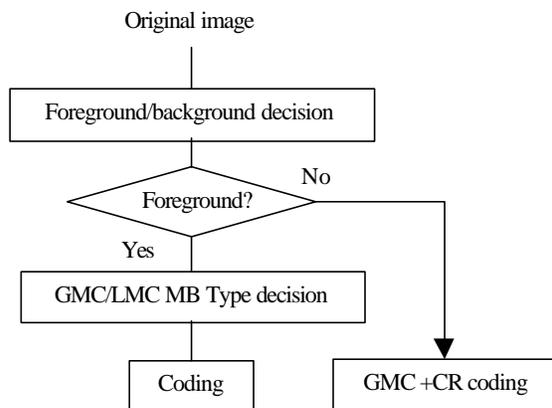


Fig 2. Flowchart of coding.

て、差分閾値 $Th3(Th3 < Th2)$ にて境界 MB の MAD 値を前景と背景に分離する．このようにこのようにして MB 単位の前景背景分離を隣接フレーム間情報だけで達成する．

3．符号化実験ならびに考察

図 2 に GMC 符号化に CR を適用するためのフローを記す．前景背景の分離を行った後に、背景 MB と判定された部分に GMC かつ予測誤差を符号化しない CR を行う．前景 MB と判定された部分は、MPEG-4 VM.15.0[8] による GMC/LMC のマクロブロック判定を行い、GMC もしくは LMC 後に予測誤差を符号化する．

表 1 に示す 3 つの SIF (353x240 画素)の画像に対し、符号化実験を行った．

表 1 実験に用いた画像

Video sequence	Frames	Contents
Mobile calendar	1-150	Pan, zoom
Horsrace	150frames	Pan, zoom
Skateboard	150frames	Pan, tilt

比較対象は以下の 3 つの符号化方法である．

- (1) P: GMC 符号化を用いない符号化
- (2) S: CR を用いない通常の GMC 符号化
- (3) CR: CR を用いる GMC 符号化(提案方法)

符号化における GOV 構造は、最初のフレームが I ピクチャ、残りのフレームはすべて P ピクチャとする．符号化制御法は MPEG-4 Verification Model Ver.15[8]を用いた．表 2 に符号



Fig 3. Example of foreground/background separation. White MB stands background (GMC+CR). UPPER image is original.

化パラメータを示す．

表 2 符号化パラメータ

Bitrate [kbps]	128, 384, 512
Frame rate [fps]	10, 15,

図 3 に、提案アルゴリズムによる分前景背景分離結果の例を示す．白色部分が背景と判断され、GMC+CR で符号化される部分である．おもむね良好に分離されているが、やや前景が多めに判定されている．これは、隣接フレーム間差分を用いているので、オクルージョンによって消失、出現した領域の部分が前景として分離されるからと考えられる．また、画像の端にかかった場合に、GMV による参照領域が画像の外側にかかった場合に CR を行わない制御をしているので、この部分が強制的に前景領域に分類される．この 2 つのことが見た目の動き領域よりも広範な部分を前景部分として判断されたと考えられる．

Video		[fps]	PSNR [dB]			
			Total	LMC	GMC	CR
mbcl	CR	9.8	18.43	22.4	24.5	17.2
	S	9.6	21.22	21	21.5	0
	P	7.6	21.28	21.3	0	0
horse	CR	10	27.52	27.6	28.5	26.1
	S	10	27.68	27.1	28.2	0
	P	8.6	27.13	27.1	0	0
skate	CR	10	23.69	24.8	26.2	23.3
	S	9.8	24.77	24.1	25.8	0
	P	9.2	24.71	24.7	0	0

Video		[fps]	PSNR [dB]			
			Total	LMC	GMC	CR
mbcl	CR	10	18.72	28.8	30.3	17.1
	S	10	24.74	24.4	25.3	0
	P	10	24.47	24.5	0	0
horse	CR	10	30.62	33.9	33.9	27.5
	S	10	32.74	32.3	33.5	0
	P	10	32.6	32.5	0	0
skate	CR	10	25.66	30.5	32	24
	S	10	28.88	28.3	30.3	0
	P	10	28.85	28.8	0	0

Video		[fps]	PSNR [dB]			
			Total	LMC	GMC	CR
mbcl	CR	15	17.94	27.3	28.6	16.5
	S	15	23.25	23	23.5	0
	P	15	23.01	23	0	0
horse	CR	15	29.85	32.4	33.3	28.5
	S	15	31.58	31	32.3	0
	P	15	31.2	31.2	0	0
skate	CR	15	24.77	29.3	30.7	23.5
	S	15	27.62	27.1	28.7	0
	P	15	27.54	27.6	0	0

Video		[fps]	PSNR [dB]			
			Total	LMC	GMC	CR
mbcl	CR	15	17.99	29.2	30.5	16.5
	S	15	24.34	24.1	24.7	0
	P	15	24.13	24.2	0	0
horse	CR	15	30.45	34.2	35.3	28.8
	S	15	33.05	32.6	33.8	0
	P	15	32.71	32.7	0	0
skate	CR	15	25.25	31.1	32.5	23.7
	S	15	29.01	28.4	30.1	0
	P	15	28.94	28.9	0	0

Fig 4. Coding result by SNR .

図4に各シーケンスの平均 SNR, 前景領域の SNR を示す. 図5に, それぞれの符号化方法になった MB の全体に対する割合を示す. LMC と GMC が前景と判定された部分である. 前景は全体の 20-40% 程度の割合となっている. 図6に GMC に CR を適用した符号化画像の例を示す. GMC+CR を行った場合は, 前景部分にて 0-5dB 程度の SNR の向上が見られた. また, 低レートでは GMC+CR の場合にフレーム落ちがほとんど無かった. しかし, 背景部分に関しては, ゆっくりとしたカメラ操作を伴い高いフレームレートにて符号化した場合に顕著な劣化が見られた(“Mobile calendar”等). これは CR が適用される回数が多いほど GMC によるアフィン変換の誤差が蓄積されていくからであると考え

られる. また, スプライト的な劇的な符号量削減にはいかなかった. これは, 文献[2,3]によれば, スプライトモードにおいて, 前景比率が 10-15% 程度のところで従来方法の半分以下の符号量で圧縮が可能である. しかし, 前景比率が 30% を超えてしまうとスプライト符号化の優位性が無くなることが報告されている. 本実験の前景比率(LMC 符号化された割合)に着目していると, 要求される前景比率よりも大きな値となっている. よって, 劇的な圧縮効果は得られないことと結論づけられる.

4. おわりに

本稿では, GMC 符号化に CR をする方法を提案し, 符号化シミュレーションを行った. その

Video		rate		
		LMC	GMC	CR
mbcl	10fps	0.26	0.07	0.67
	15fps	0.22	0.08	0.70
horse	10fps	0.41	0.09	0.50
	15fps	0.35	0.09	0.56
skate	10fps	0.41	0.03	0.56
	15fps	0.35	0.03	0.62

Fig.5 Percentage of each coding method.



Fig 6. Coding images.

Upper:GMC+CR, Foreground toys are clearly displayed, but background is significantly blurred. Lower: normal GMC.

結果，背景領域の符号量を削減する分を注目度の高い前景領域に配分することができるので，前景領域にたくさんの符号量が割り当てられ，より細かい量子化ステップで符号化でき，視覚的にも良好な結果が与えられることがわかった。しかし，前景比率がかなり大きな値となり，ほとんどスプライト的な効果は期待できないことがわかった。

今後の課題としては，さらに高効率性をたかめるために，コントラストの大きなエッジ部分などの，本来は背景となるべきが，前景として抽出されている部分の対処法，Map1 を作成す時の閾値Th1の自動決定方法の考案が今後の課

題である。

謝辞

本研究におよびシミュレーションにご協力いただいた(株)NTTソフトウェアの米原紀子氏に感謝します。

参考文献

- [1] ISO/IEC 14496-2, "Information Technology – Coding of audio-visual objects– Part 2: Visual" (1999).
- [2] K. Jinzenji, S. Okada, H. Watanabe, and N. Kobayashi, "Automatic Two-layer Video Object Plane Generation Scheme And Its Application to MPEG-4 Video Coding," IEEE ISCAS2000, pp.606-609, May 2000.
- [3] 秦泉寺久美, 渡辺裕, 岡田重樹, 小林直樹, "MPEG-4 スプライト符号化を用いた超高压縮映像符号化," 信学論 Vol. J84-DII No.5, pp.758-768, 2001 .
- [4] K. Jinzenji, H. Watanabe, and N. Kobayashi, "Global Motion Estimation for Static Sprite Production and Its Application to Video Coding," IEEE ISPACS '98, pp.328-332, November 1998.
- [5] M. Irani, S. Hsu, and P. Anandan, "Video Compression Using Mosaic Representation," Signal Processing: Image Communication, Vol. 7, Nos.4-6, pp. 529-552, November 1995.
- [6] J. Wang and E. Adelsen, "Representing Moving Images with Layers," IEEE Trans. on IP, Vol. 3, No. 5, pp.625-638, September 1994.
- [7] H. Jozawa, K. Kamikura, A. Sagata, H. Kotera, and H. Watanabe: "Two-stage motion compensation using adaptive global MC and local affine MC," IEEE Trans. Circuit & Systems for Video Tech., Vol. 7, No. 1, pp. 75-85, January 1997.
- [8] "MPEG-4 Video Verification Model Version 15.0," ISO/IEC JTC 1/SC 29/WG 11 MPEG98/N3093.