

動き補償と色信号間予測を併用したRGBカラー動画の可逆符号化 ～色信号間の動ベクトル共有による付加情報の削減～

Lossless Coding of RGB Color Video Using Motion Compensation and Inter-Color Prediction
～Reduction of Side-Information by Sharing Motion Vectors among Color Signals～

林 杏輔*
Kiyosuke HAYASHI

青森 久*
Hisashi AOMORI

松田 一朗*
Ichiro MATSUDA

伊東 晋*
Susumu ITOH

1. はじめに

先に我々は、RGB 4:4:4 フォーマットのカラー動画を対象として、動き補償と色信号間予測を同時に利用する可逆符号化方式を提案した [1]。この方式では、動き補償に必要な動ベクトルを色信号毎に独立に検出していた。しかし、一般にカラー動画では、同一時刻の各色信号の動き情報は共通であると考えられるため、それらを個別に伝送することは非効率的である。そこで本稿では、動ベクトルに関する情報を RGB 信号間で共有するものとし、それらを各色信号合計の符号量が最小となるように最適化する手法について検討を加えたので報告する。

2. RGB カラー動画のブロック適応線形予測 [1]

本稿では、RGB 4:4:4 フォーマットのカラー動画をフレーム毎に R G B の順序で符号化することとする。最初に符号化される R 信号については、モノクローム動画用に開発された、可変ブロックサイズの動き補償に基づいた可逆符号化方式 [2] を適用する。具体的には、フレーム毎に $M=28$ 種類の線形予測器を用意し、これらを 8×8 画素のブロック単位で切り換えながらラスタスキャン順に適用し、得られた予測誤差信号をコンテクストモデリングに基づいた適応算術符号器を用いて符号化する。各予測器は、GOP 内の先頭フレーム (I ピクチャ) を除き、当該フレームの符号化済み近傍画素だけではなく、動き補償された過去 2 枚の参照フレーム上の画素も同時に参照する 3 次元の構造を有しており、予測値はそれらの線形結合で算出される。また、2 枚の参照フレームのうち 1 枚は直前のフレームに固定されているが、もう 1 枚は過去 5 枚 ($\tau = 1, 2, \dots, 5$) の符号化済みフレームの中から適応選択できるようになっており、可変サイズのブロック単位で適切な参照フレームの番号と、それぞれについて検出された整数画素精度の動ベクトル v_1, v_2 の情報が必要となる。

一方、G, B 信号の符号化に際しては、上記の線形予測器を同一時刻の符号化済み色信号も参照するように拡張している。例えば、B 信号を符号化する場合、注目画素 p_0 の予測値 $\hat{B}_t(p_0)$ は図 1 のように符号化済みの R, G 信号を含めた最大 5 種類の信号の線形結合となる。ここで、時刻 t における画素 p の RGB 信号をそれぞれ $R_t(p), G_t(p), B_t(p)$ と表記すると、 $\hat{B}_t(p_0)$ の定義は次式で与えられる。

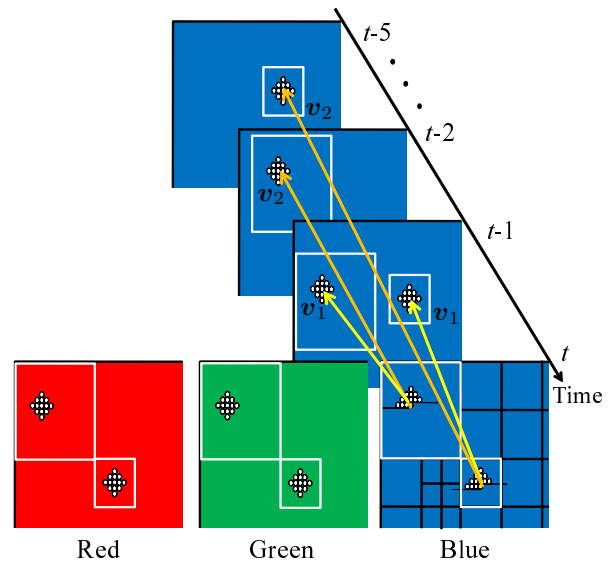
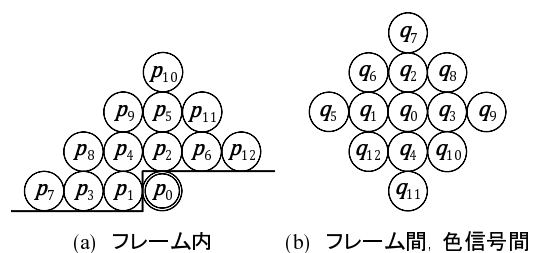


図 1 B 信号符号化時の予測手法

$$\hat{B}_t(p_0) = \sum_{k=1}^{K_1} a_m(k) \cdot B_t(p_k) + \sum_{k=0}^{K_2-1} \left\{ b_m(k) \cdot B_{t-1}(q_k + v_1) + b_m(K_2+k) \cdot B_{t-\tau}(q_k + v_2) \right\} + \sum_{k=0}^{K_3-1} \left\{ c_m(k) \cdot G_t(q_k) + c_m(K_3+k) \cdot R_t(q_k) \right\} \quad (1)$$

但し K_1, K_2, K_3 は、それぞれフレーム内、フレーム間、色信号間の予測に用いる参照画素数に関するパラメータである。図 2 にフレーム内及びそれ以外の場合の参照画素の配置を示す。また、 $a_m(k), b_m(k), c_m(k)$ は m 番目 ($m = 1, 2, \dots, M$) の予測器における各色信号に対する重み (予測係数) である。



(a) フレーム内 (b) フレーム間、色信号間

図 2 参照画素の配置

*東京理科大学 理工学部 電気電子情報工学科

3. 動ベクトルに関する情報の共有

文献 [1] では、動ベクトル v_1, v_2 を色信号毎に独立に検出していた為、それらに関する付加情報も各色信号個別に伝送する必要があった。ここで、動ベクトルに関する付加情報とは、2本の動ベクトル v_1, v_2 に加えて、それらを検出したブロックのサイズ(四分木情報)及び v_2 が参照するフレームの番号を意味している。以下では、これらの情報を RGB 信号間で共有することで付加情報の削減を図るものとし、予測器や算術符号器に関する他のパラメータと共に、フレーム毎に最適化する手順について述べる。

- (1) 画面を3段階のサイズのブロック ($8 \times 8, 16 \times 16, 32 \times 32$) に一様分割し、それぞれのサイズのブロックについて過去5フレームとの間でRGB信号の2乗誤差和に基づいたブロックマッチングを実行する。ここで得られた5本の動ベクトルをそれぞれのサイズ毎に記録しておくと共に、前フレーム及び5フレーム全体について最適であったベクトルをそれぞれ v_1, v_2 の候補とする。
- (2) 実際の動ベクトルが割り当てられるブロック(以後、MCブロックと呼ぶ)のサイズを 16×16 画素と仮定し、(1)で求めた動ベクトル v_1, v_2 を前提に、各色信号それぞれ M 種類の予測器の割り当てを決定する。ここで用いる予測器は原則として前フレームの符号化に使用したものであるが、それらが存在しないときは、重み付き2乗誤差を最小化する [3] の手法により新たに設計する。
- (3) 文献 [2] と同様な手順により、予測係数や算術符号化に用いる確率モデルとコンテキストモデリングに関するパラメータの修正を行うと共に、更にそれを前提として 8×8 画素毎の予測器の割り当てを更新する。これらの処理はRGB各信号独立に実行可能である。
- (4) 32×32 画素のMCブロックを 8×8 画素になるまで再帰的に分割しながら、それぞれのサイズについて記録されている動ベクトルの終点を周囲8画素の範囲で修正する。更に、各分割状況における修正後のコストを記録しおき、最終的に 32×32 画素の領域についてコストを最小とするMCブロックのサイズ及び動ベクトル v_1, v_2 の組み合わせを決定する。
- (5) 付加情報も含めたRGB信号全体の符号量をコストとして(3)と(4)の処理を交互に繰り返す。反復打ち切り条件は、コストが減少しなくなるか、予め設定した反復回数に達したときである。

4. 特性評価とまとめ

ITE/ARIB 標準動画 version 2 からコンバートしたRGB 4:4:4 フォーマットのカラー動画(CIFサイズ, 30Hz, RGB信号各8bit)を対象として符号化シミュレーションを実施した。表1に25フレーム分の符号化レートを比較した結果を示す。ここで、「従来方式」とは、色信号毎に独立に動ベクトルを検出する文献 [1] の方式を指す。但し、「本方式」「従来方式」共に、参照画素数に関するパラメータは $K_1 = 12, K_2 = 13, K_3 = 13$ と設定している。また、「FRExt」はH.264のHigh 4:4:4 プロファイルに基づいた可逆符号化方式(JM10.1, 参照フレーム5枚, CABAC符号化)[4]であり、GOP構造はIPBBB...と設定した。表1より、「本方式」は「従来方式」と比べて符号化レートを約0.04bits/pel削減している様子を確認できる。また、「FRExt」に対しては22.8%低い符号化レートを達成しており、RGB 4:4:4 フォーマットのカラー動画の可逆符号化方式として、高い水準の性能を有していると言える。

を比較した結果を示す。ここで、「従来方式」とは、色信号毎に独立に動ベクトルを検出する文献 [1] の方式を指す。但し、「本方式」「従来方式」共に、参照画素数に関するパラメータは $K_1 = 12, K_2 = 13, K_3 = 13$ と設定している。また、「FRExt」はH.264のHigh 4:4:4 プロファイルに基づいた可逆符号化方式(JM10.1, 参照フレーム5枚, CABAC符号化)[4]であり、GOP構造はIPBBB...と設定した。表1より、「本方式」は「従来方式」と比べて符号化レートを約0.04bits/pel削減している様子を確認できる。また、「FRExt」に対しては22.8%低い符号化レートを達成しており、RGB 4:4:4 フォーマットのカラー動画の可逆符号化方式として、高い水準の性能を有していると言える。

表1 符号化レートの比較 (bits/pel)

Sequence	本方式	従来方式	FRExt
Truck train	8.570	8.627	10.408
Red leaves(pan up)	10.944	10.968	12.964
Fountain(follow)	9.269	9.301	13.436
Studio concert	11.716	11.797	13.971
Woman at harbor	8.378	8.382	9.658
Colorful world A	11.241	11.290	13.391
Average	10.020	10.061	12.305

【参考文献】

- [1] 林 杏輔, 青森 久, 松田 一朗, 伊東 晋: “動き補償と色信号間予測を併用したRGBカラー動画の可逆符号化”, 電子情報通信学会総合大会, No.D-11-1, Mar. 2011.
- [2] H. Maeda, A. Minezawa, I. Matsuda and S. Itoh: “Lossless Video Coding Using Multi-Frame MC and 3D Bi-Prediction Optimized for Each Frame”, Proc. 14th European Signal Processing Conf. (EUSIPCO-2006), Thu. 3.2, Sep. 2006.
- [3] 松田一朗, 森弘史, 前田譲治, 伊東晋, “可逆符号化のためのレートを最小とする予測器の設計法と評価”, 電子情報通信学会論文誌 Vol.J85.D.II No.3, pp.448-456, Mar. 2002.
- [4] G. J. Sullivan, P. Topiwala and A. Luthra: “The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions”, Proc. SPIE Conf. on Applications of Digital Image Processing XXVII, Vol.5558, pp.53-74, Aug. 2004.