

RGB 4:4:4 カラー動画像符号化のための 色信号間/フレーム間予測に関する一検討

A Study on Inter-Color/Inter-Frame Prediction for Video Coding in RGB Color Space

小林 雅典*

青森 久*

松田 一朗*

伊東 晋*

Masanori KOBAYASHI

Hisashi AOMORI

Ichiro MATSUDA

Susumu ITOH

1. はじめに

近年、動画像符号化において1/2, 1/4画素精度といった非整数画素精度の動ベクトルによる動き補償予測を用いることが一般的となっている。これに対し我々は、動ベクトルの精度を整数画素精度に限定し、代わりにフレーム毎に設計した複数の内挿フィルタを適応選択する方式を提案している [1]。本稿では、この方式をデジタルシネマなど色再現性を重視する分野で用いられる RGB 4:4:4 フォーマットのカラー動画を対象に拡張し、時間方向(フレーム間)だけでなく、色信号間の相関も同時に利用する手法を導入したので報告する。

2. 適応内挿フィルタを用いた動き補償予測 [1]

文献 [1] では、 16×16 画素のマクロブロック単位で整数画素精度の動ベクトル \mathbf{v} を求め、これに従って動き補償された前フレーム上の画素群に対して、フレーム毎に設計した M 組の内挿フィルタのいずれかを適用することで予測値を得ている。注目画素 \mathbf{p} が属するブロックに m 番目のフィルタ ($m = 1, 2, \dots, M$) が割り当てられているとき、この画素における予測値は次式により算出される。

$$\hat{s}_t(\mathbf{p}) = \sum_{k=1}^K a_m(k) \cdot \tilde{s}_{t-1}(\mathbf{p} + \mathbf{q}_k + \mathbf{v}) \quad (1)$$

ただし、 $a_m(k)$ および \mathbf{q}_k ($k = 1, 2, \dots, K$) はそれぞれ内挿フィルタの係数とその配置(マスク形状)を定義するベクトルであり、本稿では菱形のマスク形状に基づいた2次元非分離型の内挿フィルタを採用している。これらのフィルタ係数 ($K \times M$ 個) およびマクロブロック毎に割り当てられるフィルタの番号と動ベクトルは、以下に示すコスト関数 J が最小となるようにフレーム毎に繰り返し最適化され、最終的に付加情報として復号側に伝送される。

$$J = D + \lambda R \quad (2)$$

ただし、 D は予測誤差の2乗和、 R は上記の付加情報に要する符号量、 λ はラグランジュ乗数である。

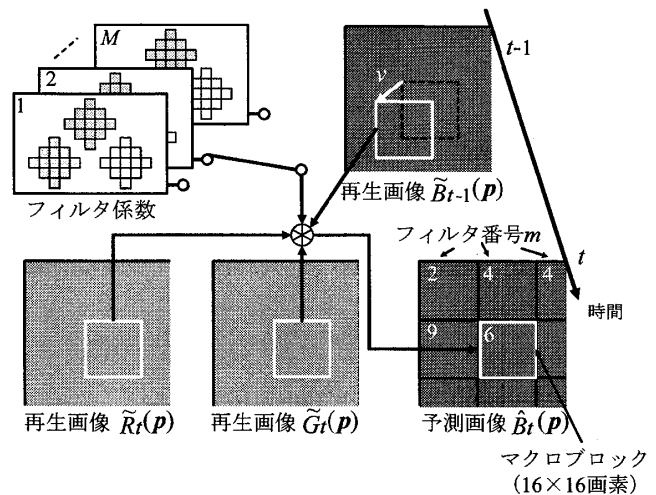


図1 色信号間/フレーム間予測の例 (B信号)

3. RGB 4:4:4 カラー動画像への拡張

カラー動画を構成する RGB 信号間には高い相関が存在するため、符号化に際してはこの相関の低減を狙って YUV 信号などへの色座標変換が施される場合が多い。しかし、高品質な画像を扱う場合、RGB \rightarrow YUV \rightarrow RGB の各変換処理における丸め誤差に起因した画質の劣化が問題となるため [2]、近年はカラー動画を RGB 信号のまま直接符号化する方式も検討されている [3]。そこで本稿では、RGB 4:4:4 フォーマットのカラー動画を符号化対象とし、適応内挿フィルタによるフレーム間の動き補償予測に加え、同一時刻の符号化済みの他の色信号からの予測も可能とする手法の導入を図る。すなわち、各フレームにおいて R \rightarrow G \rightarrow B の順序で符号化するものとしたとき、R 信号に対しては (1) 式と同様に前フレームの R 信号を用いたフレーム間予測を適用する。一方、G、B 信号においては、前フレームの同一色信号に加え、同一フレームで符号化の終了した色信号の参照も可能とする。したがって B 信号では、図1のように前フレームの B 信号の再生値 $\hat{B}_{t-1}(\mathbf{p})$ に同一時刻の R と G 信号の再生値 $\hat{R}_t(\mathbf{p})$ と $\hat{G}_t(\mathbf{p})$ を加えた3種の信号を使用し、次式に基づいて予測値を算出する。

*東京理科大学 理工学部 電気電子情報工学科

$$\hat{B}_t(\mathbf{p}) = \sum_{k=1}^{K_B} a_m(k) \cdot \tilde{B}_{t-1}(\mathbf{p} + \mathbf{q}_k + \mathbf{v}) + \sum_{k=1}^{K_G} a_m(K_B + k) \cdot \tilde{G}_t(\mathbf{p} + \mathbf{q}_k) + \sum_{k=1}^{K_R} a_m(K_B + K_G + k) \cdot \tilde{R}_t(\mathbf{p} + \mathbf{q}_k) \quad (3)$$

なお、カラー動画信号に対してフレーム間予測のみを適用する場合、色信号間でフィルタ係数や動ベクトルを共有することで付加情報を削減することも可能であるが [4]、本手法では色信号毎に予測式が異なるため、フィルタ係数も色信号毎に独立した M 組を求める必要がある。しかし、(3) 式に基づいた予測処理を適応的に実行することにより、フレーム間予測のみを適用する R 信号以外では、固定の色座標変換よりも強力な冗長さの削減が期待できる。

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, 30 フレーム分, RGB 信号各 8bit) を対象に符号化シミュレーションを実施した。ただし、予測誤差に対しては H.263 に準拠した DCT 符号化 (量子化パラメータ QP 一定) を適用し、(2) 式に示すコスト関数のラグランジュ乗数は $\lambda = 0.85 \cdot (QP)^2$ と設定した。また、比較のために YUV 4:4:4 信号に対して本手法を適用した場合の符号化特性についても調査した。ただし、いずれの場合も SNR は RGB 色空間で測定するものとし、YUV 信号を符号化対象とする際はこの SNR が最大となるように QP の値に重み付けを行っている [5]。更に、色信号毎に用意するフィルタの数 M およびタップ数に関するパラメータ $K_R, K_G, K_B(K_Y, K_U, K_V)$ は、予備実験の結果に基づいて表 1 に示す値に設定した。図 2 に示すレート・SNR 特性より、RGB 信号を符号化対象とした際に、色信号間予測の導入によって大幅に符号化効率が向上していることがわかる。この傾向は、画像 Fountain における水しぶきのように、動き補償予測が困難な絵柄において顕著であった。また、高ビットレートにおいては、

表 1 各パラメータの設定条件

信号	フィルタの数 (M)	タップ数		
		K_R/K_Y	K_G/K_U	K_B/K_V
R	32	41	-	-
G	48	13	25	-
B	48	13	13	13
Y	32	41	-	-
U	32	13	13	-
V	32	13	5	13

YUV 信号に色座標変換してから符号化するより、RGB 信号をそのまま符号化の方が高い SNR が得られることも確認できる。

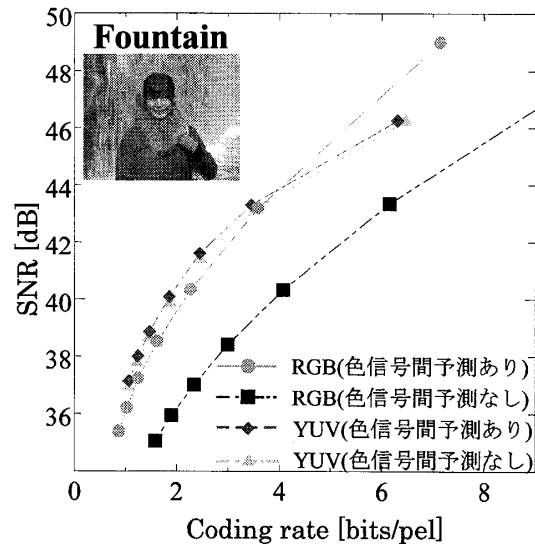
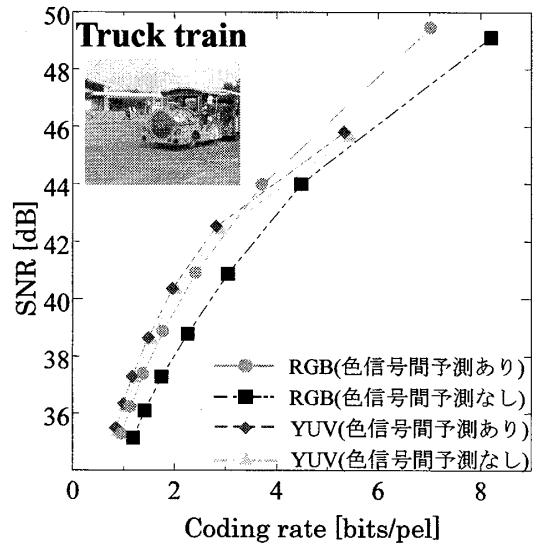


図 2 符号化特性

【参考文献】

- [1] 長島 他: “適応内挿フィルタを用いた非整数画素精度動き補償方式”, 信学技報, Vol.104, No.752, pp.43-48 (Mar. 2005)
- [2] G. Sullivan: “Approximate Theoretical Analysis of RGB to YCbCr to RGB Conversion Error”, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6, Document: JVT-1017 (Sep. 2003)
- [3] G. J. Sullivan et al.: “New Standardized Extensions of MPEG4-AVC/H.264 for Professional-Quality Video Applications”, Proc. of ICIP 2007, Vol.I, pp.13-16 (Sep. 2007)
- [4] 青木 他: “適応内挿フィルタを用いたカラー動画の動き補償予測”, 信学総大, No.D-11-11 (Mar. 2006)
- [5] 松田 他: “カラー画像の適応的可変ブロック形状 KL 変換符号化”, テレビ誌, Vol.47, No.10, pp.1277-1285 (Oct. 1993)