

動き補償された複数のフレームを用いる適応内挿フィルタの検討

A Study on Adaptive Interpolation Filters for Motion-Compensated Multiple Frames

中島 勝己*

高松 慶*

松田 一朗*

伊東 晋*

Masaki NAKAJIMA

Kei TAKAMATSU

Ichiro MATSUDA

Susumu ITOH

1. はじめに

近年、動画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が一般的に用いられている。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手法について検討している[1]。また、この手法を2枚の参照フレームを用いるBピクチャに導入することで、画素間の輝度値の内挿処理だけでなく、フレーム毎に得られた予測値の合成処理についても同様に最適化できることを示した[2]。本稿では、更に多数の参照フレームに対して上記手法を導入する場合について検討を加える。

2. 複数のフレームを用いる適応内挿フィルタ

本方式では、フレーム毎に16種類の内挿フィルタを用意し、これらを16×16画素のマクロブロック単位で適応選択している。予測値を生成する際は、図1のように連続するN枚の参照フレームを、マクロブロック毎に与えられた整数画素精度の動ベクトル $v_1 \sim v_N$ に従って動き補償すると共に、それぞれのフレームから $K_n (n = 1, 2, \dots, N)$ 個ずつ抽出した画素群を内挿フィルタに基づいて加重平均する。ここで、 m 番目の内挿フィルタ($m = 1, 2, \dots, 16$)が選択されたブロックに属する画素 p の予測値 $\hat{s}_t(p)$ は次式により算出される。

$$\hat{s}_t(p) = \sum_{n=1}^N \sum_{k=1}^{K_n} a_m \left(\sum_{i=1}^{n-1} K_i + k \right) \cdot \tilde{s}_{t-n}(p + v_n + q_k) \quad (1)$$

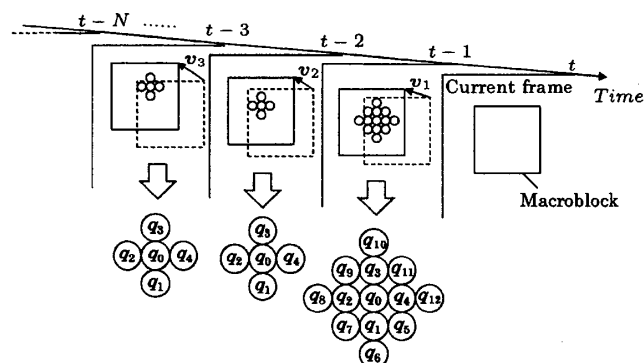


図1 複数フレームを用いた予測値の生成

但し、 $a_m(k)$ は m 番目の内挿フィルタの係数、 $\{q_k\}$ は図1に示すように $q_0 = (0, 0)$ を中心になんらかの配置した画素群の相対座標を表すベクトル、 $\tilde{s}_{t-n}(p)$ は符号化対象フレーム(時刻 t)から n 枚離れた参照フレームの輝度値(再生値)をそれぞれ意味している。このように、本方式の内挿フィルタは空間方向だけでなく時間方向にも係数 $a_m(k)$ を配置した3次元構造を有している。

3. 参照フレーム同士のブロックマッチングに基づいた動ベクトルの検出

上述のように、 N 枚の参照フレームに対して動き補償を適用するには、マクロブロック毎に N 本の動ベクトル $v_1 \sim v_N$ が必要となる。このため、フレーム数 N の増加に伴って動ベクトルに関する付加情報の増大が問題となる。この問題を解決するため、本方式では直前のフレームに対する動ベクトル v_1 のみを符号化し、残りの動ベクトル $v_2 \sim v_N$ を復号側に設置した動き検出器を用いて推定する手法[3]を導入する。具体的には、図2のように直前の参照フレーム(時刻 $t-1$)上に当該マクロブロックを v_1 だけ平行移動させたブロック B を設定し、このブロックと $n \geq 2$ 枚目の参照フレームとの間でブロックマッチングを行うことにより動ベクトル v'_n を検出する。

$$v'_n = \operatorname{argmin} \left\{ \sum_{p \in B} (\tilde{s}_{t-1}(p) - \tilde{s}_{t-n}(p + v'_n))^2 \right\} \quad (2)$$

但し、上記ブロックマッチングの探索範囲を広く設定すると演算量が膨大となるため、予め各参照フレームに存在する符号化済み動ベクトルを用いて v'_n の探索中心 $\hat{v}'_n$ を求め、その±2画素の範囲に限定してブロックマッチングを実行する。ここで $\hat{v}'_2$ は、図3のようにブロッ

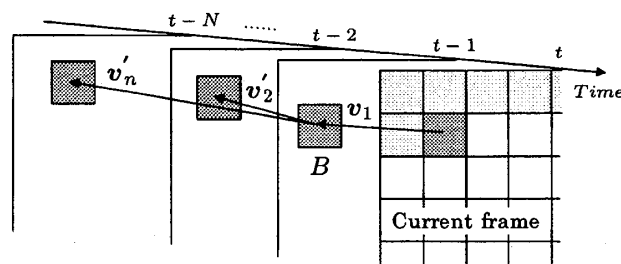


図2 参照フレーム同士のブロックマッチング

*東京理科大学 理工学部 電気電子情報工学科

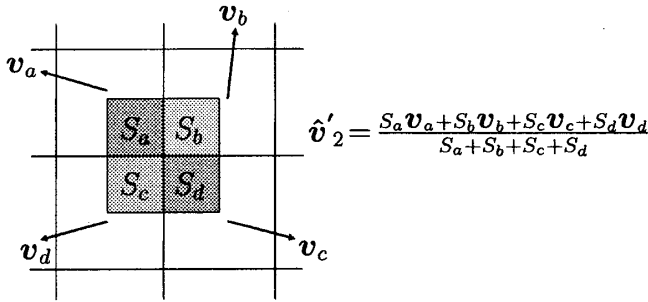


図3 符号化済み動ベクトルと探索中心の関係

ク B とオーバーラップする 4 つのマクロブロックの符号化済み動ベクトル $v_a \sim v_d$ を、それぞれオーバーラップした領域の面積 $S_a \sim S_d$ で加重平均したものである [4]。同様に $n \geq 3$ 枚目についても、 $v_1 + v'_{n-1}$ だけ平行移動したブロック B と各マクロブロックとのオーバーラップ領域の面積比を算出することで、時刻 $t - n + 1$ で符号化した動ベクトルから $\hat{v}'_n - v'_{n-1}$ を求めることができる。

最終的に復号側で (1) 式に基づいた予測値生成処理を実行するためには、16 種類の内挿フィルタの係数群 $\{a_m(k)\}$ に加え、マクロブロック毎に選択されたフィルタの番号 m と動ベクトル v_1 を付加情報として与える必要がある。これらの値は、上記の手順で $v_2 \sim v_N$ が決定されることを前提として評価したコスト関数 $J = D + \lambda R$ が最小となるよう、フレーム毎に繰り返し最適化される。但し、 D は原画像と予測画像 $\hat{s}_t(p)$ との 2 乗誤差和、 R は上述した付加情報の符号量、 λ はラグランジュ乗数である。なお、 $\{a_m(k)\}$ を符号化する際は、前フレームで使用した係数との差分値を符号化対象とすることで付加情報の増加を抑えている [1]。

4. 特性評価とまとめ

4 種類のモノクローム標準動画 (CIF, 30Hz, 30 フレーム分) を対象として符号化シミュレーションを実施した。但し、予測誤差に対しては H.263 に準拠した DCT 符号化 (量子化ステップサイズ Q 一定) を適用し、上述したコスト関数のラグランジュ乗数は $\lambda = 0.85 \cdot (Q/2)^2$ と設定した。図 4 は参照フレーム数 N を変化させた際の $N=1$ に対する符号化レートの削減率 $\Delta \text{ bitrate}$ [5] を示している。このとき、各フレームに配置する係数の数 K_n については、それぞれの N に対して平均的に優れた性能が得られるよう、予備実験に基づいて表 1 のように設定した。これより、 $N=5$ のときにほぼ最適な結果が得られていることがわかる。図 5 は、画像 Mobile & Calendar について $N=5$ と $N=1$ 、および H.264 と同様な 6 タップフィルタに基づいた 1/4 画素精度の動き補償予測 ($N=1$) を適用した際のレート・SNR 特性を示している。同図より、予測に用いる参照フレーム数を $N=5$ とすることで、従来方式 ($N=1$) に比べて同一レートで

0.3~0.8dB 程度高い SNR を達成している様子を確認できる。

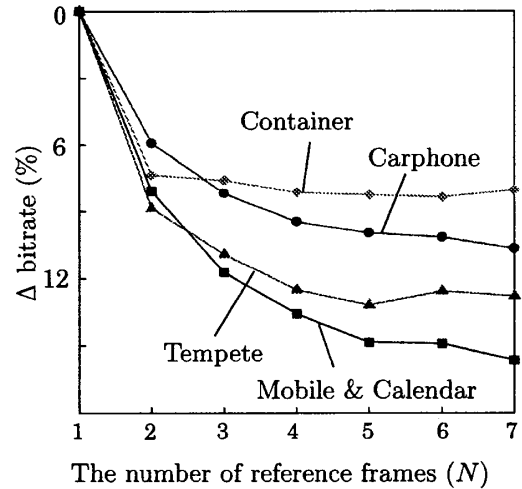
図4 参照フレームの枚数と $\Delta \text{ bitrate}$ の関係

表 1 各フレームに配置する係数の数

N	1	2	3	4	5~7
K_1	25	13	13	13	13
K_2	—	5	5	5	5
K_3	—	—	5	5	5
K_4	—	—	—	5	5
$K_5 \sim K_7$	—	—	—	—	1

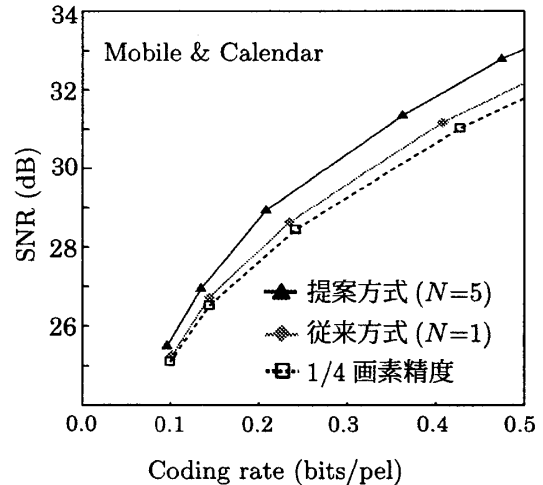


図5 符号化特性

【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] 長島 他: “適応内挿フィルタを用いた両方向動き補償予測の性能評価”, 第 28 回情報理論とその応用シンポジウム (SITA2005), Vol. 2, pp. 853-868, Nov. 2005.
- [3] Y. Suzuki: “Video Encoding Scheme Employing Intra and Inter Prediction Based on Averaged Template Matching Predictors”, IEICE trans. Vol. E91-D, pp. 1127-1134, Apr. 2008.
- [4] 清水 他: “参照フレーム変更によるビットストリームスケールング”, 信学論, Vol. J83-D-2, No. 3, pp. 967-976, Mar. 2000.
- [5] G. Bjontegaard: “Calculation of average PSNR differences between RDCurves”, Doc. VCEG-M33, Mar. 2001.