

適応内挿フィルタを用いた動き補償予測

～ フレーム毎のフィルタ数の適応化に関する一検討 ～

Motion-Compensated Prediction Using Adaptive Interpolation Filters

～ A Study on Adaptation of the Number of Filters Used in Each Frame ～

海野 恭平*

青森 久*

松田 一朗*

伊東 晋*

Kyohei UNNO

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方式を提案している [1]。従来、フレーム毎に選択可能なフィルタの数は、符号化処理に先立って設定する必要があった。本稿では、画像サイズや絵柄、符号化レートなどの条件に応じて、内挿フィルタの数をフレーム毎に変更可とする手法について検討を加えたので報告する。

2. 適応内挿フィルタを用いた動き補償予測 [1], [2]

本方式では、 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を検出する。マクロブロック毎の予測画像は、図1に示すように動き補償された参照画像(前フレームの再生画像)に対して M 種類の内挿フィルタのいずれかを適用することで生成される。ここで、 m 番目の内挿フィルタ ($m = 1, 2, \dots, M$) 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。

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

ここで、 $\tilde{s}_{t-1}(p)$ は参照画像(前フレームの復号画像)の輝度値、 $a_m(k)$ は内挿フィルタの係数、 v は動ベクトルをそれぞれ表している。また、 $\{q_m(k) \mid k = 1, 2, \dots, K_m\}$ は2次元フィルタマスクの形状およびサイズ(タップ数 K_m) を定義するベクトルの集合である。本方式では、様々な形状のフィルタマスク(計64通り)を用意し、内挿フィルタ毎に後述のコスト関数に基づいて適切なマスク形状とフィルタ係数の組み合わせを決定している [2]。したがって、復号側で(1)式の予測処理を実行するためには、マクロブロック毎の動ベクトル(v)とフィルタ番号(m)に加え、 M 種類の内挿フィルタに関する2次元フィルタマスク $\{q_m(k)\}$ とその内部に配置される係数 $a_m(k)$ の情報が必要となる。これらのパラメータは、フレーム毎

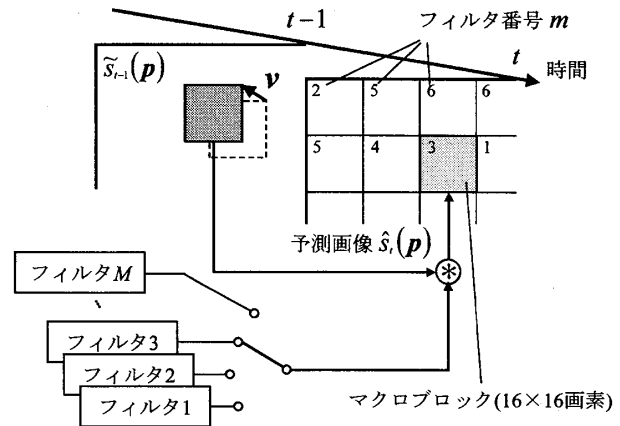


図1 適応内挿フィルタを用いたMC

に以下のコスト関数 J が最小となるよう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。

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

但し、 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。

3. フレーム毎のフィルタ数の適応化

文献 [1], [2] では、前フレームで使用した内挿フィルタの係数を、上記の最適化処理の初期値として利用していた。これは初期値との差分値を符号化対象とすることで、フィルタ係数に関する付加情報の符号量低減を狙ったものであるが、内挿フィルタの数 M は各フレームで共通であることが前提となっていた。また M の設定値は、符号化処理に先立って決定する必要があり、これまでは多数の画像を対象とした予備実験に基づいて決めた値を固定的に使用していた。しかし、 M の設定値と最終的な符号化効率の関係は画像サイズや絵柄によって異なり、また同一のシーケンスにおいてもその最適値はフレーム毎に変動するものと考えられる。そこで本稿では、各フレームにおいて十分な数の内挿フィルタを初期値として用意し、(2)式のコスト関数に基づいた最適化手順の中で不要なフィルタを徐々に削減する手法 [3], [4] の導入を図る。具体的な手順を以下に示す。

- (1) 16×16 画素のマクロブロック毎にブロックマッチング法を適用し、整数画素精度の動ベクトル v の初

*東京理科大学 理工学部 電気電子情報工学科

期値を決定する。

- (2) 各マクロブロックに対して前フレームで使用した M' 種類の内挿フィルタを順に適用し、2乗誤差 D を最小とするフィルタ番号 m を割り当てる。なお、前フレームがイントラフレームの場合は、(1) で求めたブロックマッチングの誤差を $M = 40$ レベルに量子化した値を仮の m とする。
- (3) 同一のフィルタ番号 m が割り当てられたマクロブロックの集合について、各マスク形状を前提とした最小2乗法を順に適用し、コスト関数 J を最小とするフィルタマスク $\{q_m(k)\}$ とフィルタ係数 $a_m(k)$ の組み合わせを決定する。
- (4) 前フレームで使用した内挿フィルタに、(3) で新たに設計した分を加えた $M = 2M'$ 種類を、当該フレームにおける内挿フィルタの初期値とする。但し、演算量の観点から M の上限は40とし、 $2M' > 40$ の場合は、割り当てられたマクロブロック数の多い順に新しく設計した $40 - M'$ 種類の内挿フィルタを優先して採用する。
- (5) 各マクロブロックについて、動ベクトル v の終点を ± 1 画素の範囲で変化させながら M 種類の内挿フィルタを順に適用し、コスト関数 J を最小とする動ベクトルとフィルタ番号 m の組み合わせを決定する。
- (6) (3) と同様の手順により、 M 種類の内挿フィルタを設計し直す。
- (7) m 番目の内挿フィルタを1つ削除すると仮定した際のコスト J_m を算出する。このとき、 m 番目の内挿フィルタを選択していたマクロブロックでは最適でない内挿フィルタが代用されるため、(2) 式において2乗誤差 D の項は一般に増加するが、削除した内挿フィルタの分だけ付加情報 R は減少する。
- (8) 上で算出した J_m ($m = 1 \sim M$) の最小値を求め、その値が内挿フィルタを削減する前に比べて減少する場合は実際にその内挿フィルタを削除し、 $M \leftarrow M - 1$ とおく。
- (9) (7) と (8) の処理を内挿フィルタが削除されなくなるまで繰り返し、更にこの状態を新たな初期値とみなして、(5) 以降の手順を反復的に実行する。

4. 特性評価とまとめ

CIF サイズ (352 × 288 画素) および SD サイズ (720 × 480 画素) のモノクローム動画 (30Hz, 30 フレーム分) を対象に符号化シミュレーションを実施した。但し、動き補償後の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。図2は $QP = 4$ の場合について、フレーム

毎に実際に使用された内挿フィルタの数 M を調査した結果である。これより、画像サイズや絵柄に応じてフレーム毎に適切なフィルタ数が自動的に設定されていると考えられる。また図3は、本方式において内挿フィルタの数を可変および $M = 16$ に固定した場合、更に 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を適用した際について符号化特性を比較したものである。同図より、3. で述べた手順によってフィルタ数の適応化を図った結果、SN比が最大で0.4 dB程度改善されている様子を確認できる。

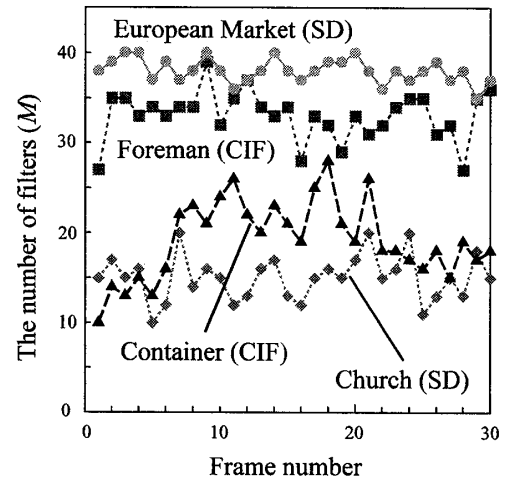


図2 フレーム毎のフィルタ数 M の推移 ($QP = 4$)

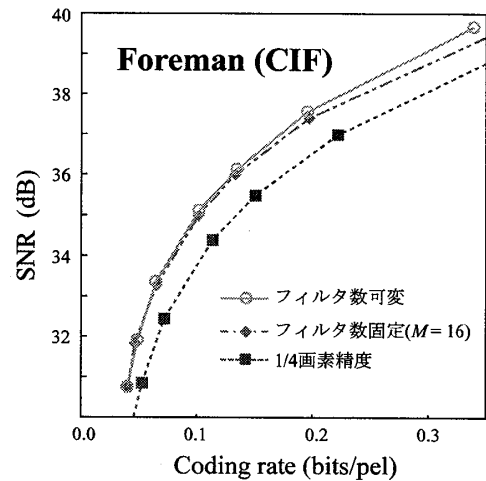


図3 符号化特性

【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] 柳原 他: “可変マスク形状適応内挿フィルタによる動き補償予測”, 映像学技報, Vol.30, No.41, pp.13-16 (July 2006)
- [3] 峯澤 他: “動き補償と適応3次元予測に基づいた動画の可逆符号化～各フレームに適した予測器数の決定に関する一検討～”, 第21回画像符号化シンポジウム (PCSJ2006), No.P-5.16, pp.117-119 (Nov. 2006)
- [4] 峯澤 他: “ブロック適応3次元予測に基づいた動画の可逆符号化のための予測器の設計法”, 信学論文誌, Vol.J90-D, No.7, pp.1639-1640 (July 2007)