

可変マスク形状適応内挿フィルタを用いた両方向動き補償予測

Bi-Directional Motion-Compensated Prediction

Using Adaptive Interpolation Filters with Variable Shape Masks

高松 慶*

松田 一朗*

伊東 晋*

Kei TAKAMATSU

Ichiro MATSUDA

Susumu ITOH

1. はじめに

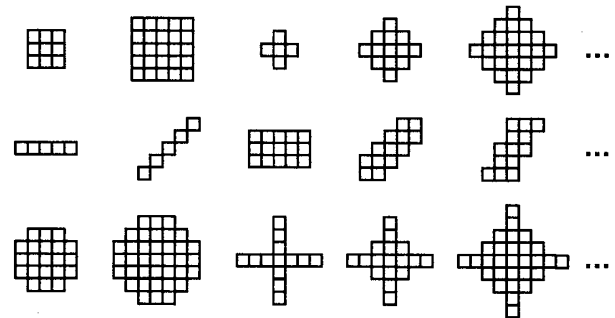
動画像符号化の国際標準方式では、一般に 1/2, 1/4 画素精度といった非整数画素精度で動ベクトルを検出することで動き補償予測の精度を高めている。これに対し、筆者らは動ベクトルを整数画素精度に限定し、代わりに複数の内挿フィルタを適応選択する方式を開発している [1]。この方式では、選択可能な内挿フィルタの組をフレーム毎に設計しているため、復号側に伝送すべき付加情報の増加が問題となる。このような観点から文献 [2] では、内挿フィルタの係数の配置 (マスク形状) を可変とすることで、予測精度の向上に寄与する係数のみを効率的に伝送する手法を提案した。本稿では、この手法を両方向予測を用いる B ピクチャにも適用できるように拡張し、符号化性能の評価を行ったので報告する。

2. 可変マスク形状適応内挿フィルタを用いた MC [2]

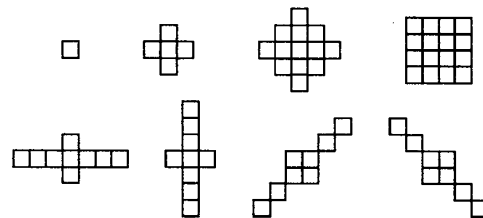
本方式では、 16×16 画素のマクロブロック毎に整数画素精度の動ベクトルを検出し、さらに動ベクトルが参照する画像に対し、16 種類の 2 次元内挿フィルタのいずれかを適用することで予測値を得る。前フレームを参照画像とする P ピクチャの場合、 m 番目の内挿フィルタ ($m = 1, 2, \dots, 16$) を選択したマクロブロックに属する画素 p の予測値は、次式の 2 次元畳み込み演算によって算出される。

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

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 $Q_m = \{q_m(k) \mid k = 1, 2, \dots, K_m\}$ は内挿フィルタ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り、本稿ではこの Q_m のことをフィルタマスクと呼ぶ。文献 [2] では、フィルタマスク Q_m を図 1 (a) およびその一部を回転/反転したものを含む 64 通りから選択できるものとし、それぞれの形状毎に 2 乗誤差を最小とするフィルタを設計していた。このとき、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、16 種類のフィルタそれぞれについてマスク Q_m とその内部に配置される係数 $a_m(k)$ が必要となる。これらの値は、フレーム毎に以下のコスト



(a) P ピクチャ用 (64 種類)



(b) B ピクチャ用 (8 種類 × 2 組)

図 1 フィルタマスクの例

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 [2]。

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

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. 両方向予測への拡張

上記方式を B ピクチャに適用する場合、前後フレームに対して個別に予測値を求めてから両者を適当な比率で合成する方法が考えられる。しかし、この方法では動ベクトルだけでなく、フィルタ番号や合成比に関する付加情報も増大してしまう。そこで文献 [3] では、前後のフレームに対するフィルタ番号を共通とし、それぞれのフィルタ係数を同時に最適化する手法を導入した。この方法では、予測値の算出は次式によって行われる。

$$\hat{s}_t(p) = \sum_{k=1}^{K_m} a_m(k) \cdot \tilde{s}_{t-1}(p + v + q_m(k)) + \sum_{k=1}^{K'_m} a_m(k + K_m) \cdot \tilde{s}_{t+1}(p + v' + q'_m(k)) \quad (3)$$

但し、 v , v' は図 2 に示すように、それぞれ前後方向の参照画像について求めた動ベクトルである。文献 [3] で

*東京理科大学 理工学部 電気電子情報工学科

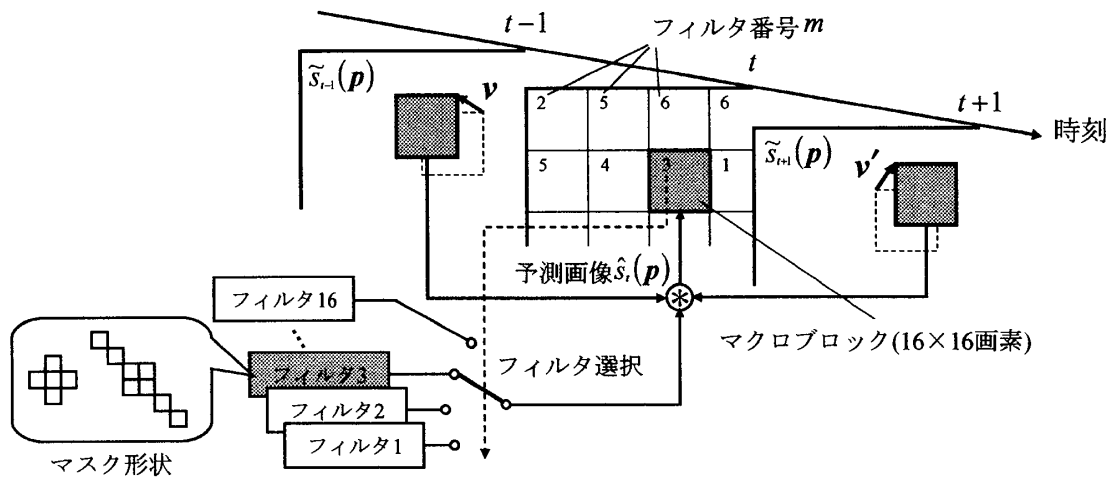


図2 Bピクチャにおける予測画像生成法

は、前後それぞれのフレームに設定するフィルタマスク $Q_m = \{q_m(k) \mid k = 1, 2, \dots, K_m\}$ および $Q'_m = \{q'_m(k) \mid k = 1, 2, \dots, K'_m\}$ を全て共通にするという条件の下でその適切な形状について調査し、13タップのひし形(図1(b)の上段左から3つ目)が平均的に良い性能を示すことを確認した。本稿では、このマスク形状を[2]と同様な手法を用いて可変形状とすることで、更なる性能改善を図る。但し、Bピクチャでは前後それぞれのフィルタマスク Q_m, Q'_m を異なる形状に設定可能であるため、図1(a)に示す64通りの形状をそのまま利用にすると、フィルタ当たりの形状の組み合わせが 64×64 と膨大となり、最適化に要する演算量および付加情報の点で実現的ではない。そこで、可変形状を前提としてBピクチャに適したマスク形状の組み合わせについて改めて調査を実施し、図1(b)に示す8通りを前後フレームに関して独立に選択できるものとした。

4. 特性評価とまとめ

標準動画 Carphone (CIF, 30Hz, 30フレーム分) のY信号を対象として符号化シミュレーションを実施した結果を図3に示す。ここで、「半画素精度」は 16×16 画素のブロック単位で検出した半画素精度の動ベクトルと共一次内挿に基づいた一般的なMCによる方式を表している。いずれの方式も予測誤差信号に対してはH.263[4]に準拠したDCT符号化(GOP長15, 量子化パラメータQP一定)を適用し、(1)式のラグランジュ係数は $\lambda = 0.85 \cdot (QP)^2$ とした。但し、Bピクチャにおいては量子化パラメータQPをPピクチャのそれに対して常に+4大きな値に設定している。図3より、提案方式においてBピクチャの導入を図ることで、SN比が0.5dB~0.8dB向上していることがわかる。ここで、Bピクチャにおいてマスク形状を可変にした効果を調べるために、GOPをIBPB...とした際のBピクチャのみの符号化特性を図4に示す。同図において「固定マスク」は、I, Pピクチャは提案方式と同一とし、Bピクチャのみマスク形状を文献[2]と同様に13タップのひし形に固定した際の特性である。これより、Bピクチャのマスク形状を可変とすることで0.4dB程度のSN比の向上が達成されている様子を確認できる。

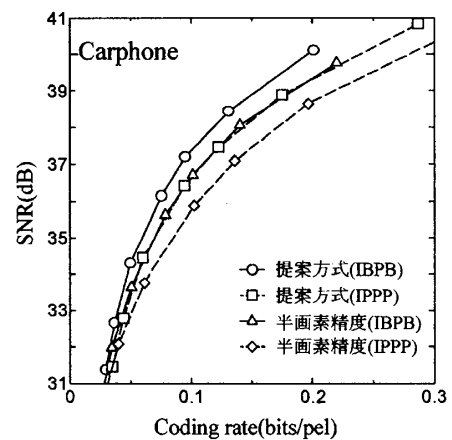


図3 符号化特性

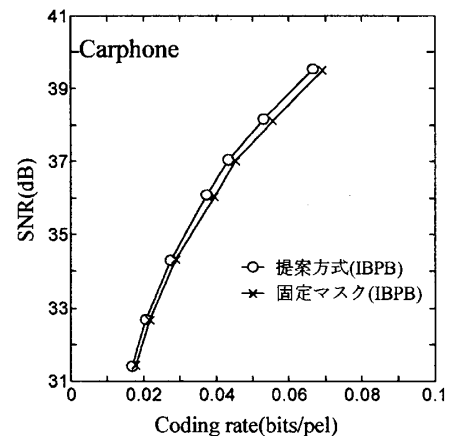


図4 Bピクチャのみの符号化特性

【参考文献】

- [1] 長島新一, 松田一朗, 伊東晋: “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] 柳原一晴, 松田一朗, 伊東晋: “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] 長島新一, 松田一朗, 森岡一幸, 伊東晋: “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”, 情報理論とその応用シンポジウム, Vol.II, pp.865-868, Nov. 2005.
- [4] ITU-T Rec. H.263, “Video Coding for Low Bitrate Communication”, Nov. 1996.