

## 複数のフレームを用いた適応3次元予測による動画像の可逆符号化

## Lossless Video Coding Based on Adaptive 3D Prediction Using Multiple Frames

石橋 寿則\*

青森 久\*

松田 一朗\*

伊東 晋\*

Toshinori ISHIBASHI

Hisashi AOMORI

Ichiro MATSUDA

Susumu ITOH

## 1. はじめに

先に我々は、ブロック適応3次元予測を用いた動画像の可逆符号化方式を提案した [1]. またこの可逆符号化方式に、複数の符号化済みフレームの中から2枚の参照フレームを適応選択するマルチフレームMCの手法を導入することで、符号化効率を改善できることも確認している [2]. 本稿では、上記の手法を拡張し、多数の参照フレームを同時に用いる予測手法について検討を加える.

## 2. ブロック適応3次元予測 [1]

ブロック適応3次元予測は、フレーム毎に  $M$  組の線形予測器を最適設計し、これらを  $8 \times 8$  画素のブロック単位で切り替えながら適用することで、動画像信号の冗長度を削減する手法である. GOP 内の先頭フレーム (Iピクチャ) を除き、各予測器は当該フレームの符号化済み近傍画素  $\{p_k\}$  だけではなく、直前のフレームの画素  $\{q_k\}$  も同時に参照する3次元構造を有している. また、参照画素  $\{q_k\}$  については、図1に示すようにMCブロックと呼ばれる可変サイズのブロック単位 ( $8 \times 8 \sim 32 \times 32$  画素) で検出した整数画素精度の動ベクトル  $v$  によって動き補償されている. ここで、注目画素  $p_0$  が属するブロックに  $m$  番目 ( $m = 1, 2, \dots, M$ ) の予測器が割り当てられているとき、この画素の予測値  $\hat{s}_t(p_0)$  は次式によって算出される.

$$\hat{s}_t(p_0) = \sum_{k=1}^{K_p} a_m(k) \cdot s_t(p_k) + \sum_{k=1}^{K_q} a_m(K_p + k) \cdot s_{t-1}(q_{k-1} + v) \quad (1)$$

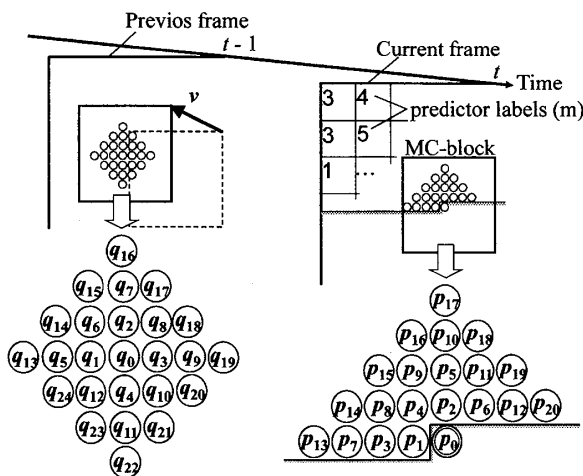


図1 ブロック適応3次元予測

ただし  $K_p$  および  $K_q$  は、それぞれ当該フレームおよび前フレームに配置させる参照画素の数、 $\{a_m(k) \mid k = 1, 2, \dots, K_p + K_q\}$  は  $m$  番目の予測器の係数 (予測係数)、 $s_t(\cdot)$  は時刻  $t$  のフレームの輝度値を表している. この適応3次元予測に必要な予測係数やブロック毎の予測器の割り当て、動ベクトルといったパラメータは、フレーム毎の符号量が最小となるように繰り返し最適化され、付加情報としてビットストリームに多重化される. 各画素で算出された予測誤差  $s_t(p_0) - \hat{s}_t(p_0)$  の値は、コンテキストモデリングに基づいた適応的な多値算術符号器を用いて、ラスタスキャン順に符号化される [3].

## 3. 参照フレーム数の拡大

上記の予測手法は、直前に符号化した参照フレームを一枚のみ使用するという意味で、MPEGなどの非可逆符号化の標準方式におけるPピクチャに対応しているとみなせる. 一方、最新の国際標準方式であるH.264/AVCでは、2枚の参照フレームを使用するBピクチャが拡張され、過去に符号化した  $R$  枚の参照フレームから任意の2枚を適応選択する技術 (マルチフレームMC+双予測) が利用可能である. 筆者らは、同様な考え方をブロック適応3次元予測に基づいた可逆符号化方式に導入し、各フレームの符号化順序を変更することなく (すなわちGOP構造はIPBBB...となる) 符号化効率を改善できることを示した [2]. ただし、各3次元予測器において、同時に利用可能な参照フレーム数はH.264/AVCと同様に  $N = 2$  枚に固定していた. 本稿ではこの制約をなくし、図2のように  $R$  枚の参照フレームから同時に  $N > 2$  枚を参照可能とした場合の特性について調査した. ここで  $n$  枚目の参照フレームの時刻を  $\tau_n (\tau_n \in \{t-1, t-2, \dots, t-R\})$  と表記したとき、予測値の算出式は以下の通りとなる.

$$\hat{s}_t(p_0) = \sum_{k=1}^{K_p} a_m(k) \cdot s_t(p_k) + \sum_{n=1}^N \sum_{k=1}^{K_q} a_m(K_p + (n-1)K_q + k) \cdot s_{\tau_n}(q_{k-1} + v_n) \quad (2)$$

本方式では、MCブロック毎に  $N$  本の動ベクトル  $v_n$  と共に選択した参照フレームの番号  $\tau_n (n = 1, 2, \dots, N)$  も符号化する必要がある. 本稿では、 $\tau_1 = t-1, \tau_2 = t-2$  に固定すると共に、残りの  $N-2$  枚の参照フレーム番号のみ適応選択の対象とし、近傍ブロックの値との相関を利用したMTF法 [3] に基づいて符号化することで付加情報の増加を抑えている. また、予測器毎に必要なとなる  $K_p + NK_q$  個の予測係数  $\{a_m(k)\}$  については、参照画素配置 (非零の予測係数が割り当てられる画素の位置)

\*東京理科大学 理工学部 電気電子情報工学科

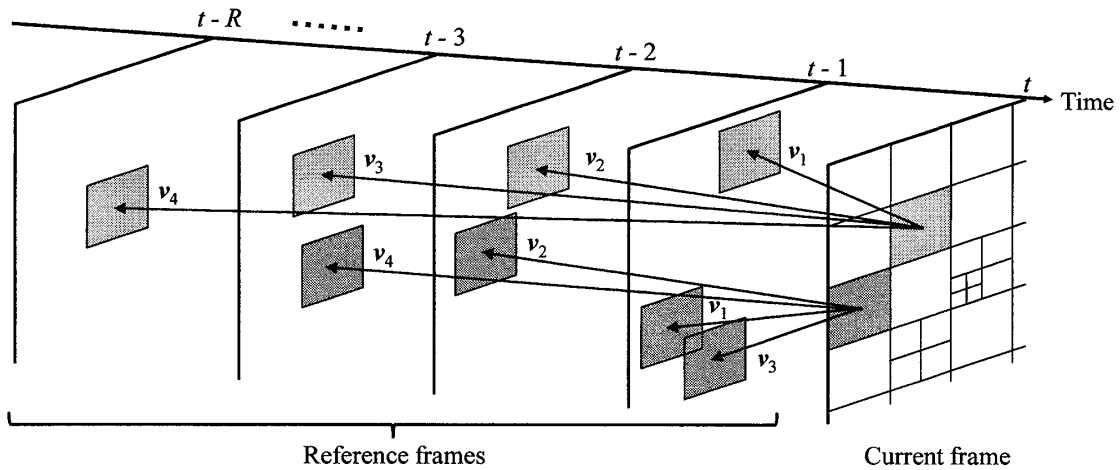


図2 参照フレーム数の拡大 ( $N = 4$ )

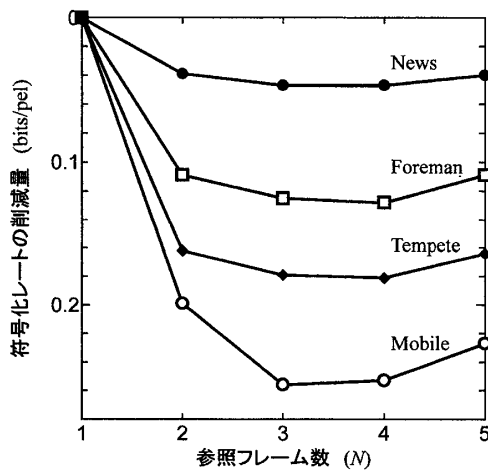


図3 参照フレームの数と符号化レートの削減量

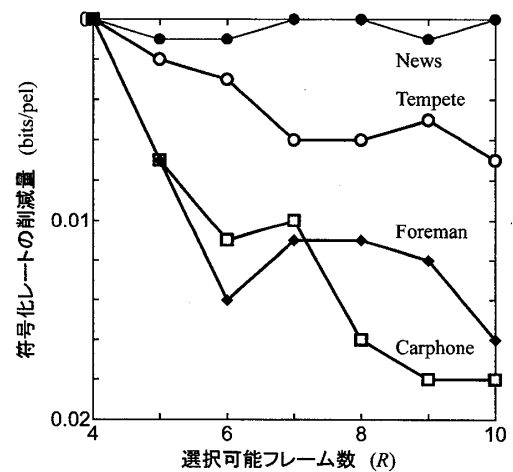


図4 選択可能なフレームの数と符号化レートの削減量

を自動設定する手法 [4] を導入し、符号化すべき予測係数の削減を図っている。

4. 特性評価とまとめ

CIFサイズの動画像 (1~25 フレーム, 30Hz, Y 信号のみ) を対象として符号化シミュレーションを実施した。図3は、選択可能なフレームの数を  $R = 10$  枚に固定した上で、同時に参照するフレーム数  $N$  を変化させた際の符号化レートの削減量 ( $N = 1$  基準) を示している。ただし、 $N \leq 2$  のときは、 $\tau_1 = t - 1$  のみを固定としている。これより、 $N = 3$  または  $4$  のときにほぼ最適な結果が得られていることがわかる。また図4は、同時に参照するフレーム数を  $N = 4$  枚に固定した上で、選択可能なフレーム数  $R$  を変化させた際の符号化レートの削減量 ( $R = 4$  基準) を示している。これより、 $R \leq 6$  の範囲では選択可能なフレーム数  $R$  を増加させることにより符号化レートが減少するものの、 $R > 6$  の範囲ではその効果が画像によって異なっている様子を確認できる。最後に、実際の符号化レートを比較した結果を表2に示す。表中の「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2枚, 選択可能なフレーム6枚, CABAC符号化)[5]を用いて符号化した結果である。本方式において、同時に参照するフレーム数を文献 [2] と同様な  $N = 2$  から  $N = 4$  へと増加させることにより、全ての画像において符号量を削減することができた。

表1 符号化レートの比較 (bits/pel)

	本方式		FRExt ( $N = 2, R = 6$ )
	( $N = 4, R = 6$ )	( $N = 2, R = 6$ )	
Carphone	<b>2.612</b>	2.618	3.072
Container	<b>2.219</b>	2.236	2.778
Foreman	<b>2.663</b>	2.684	3.165
Mobile	<b>3.394</b>	3.443	4.147
News	<b>1.290</b>	1.299	1.652
Tempete	<b>3.317</b>	3.338	3.935
Average	<b>2.583</b>	2.603	3.125

【参考文献】

[1] 塩寺 他: 映像学誌, Vol.60, No.7, pp.1051-1058, July 2006.  
 [2] H. Maeda et al.: Proc. EUSIPCO-2006, Thu. 3.2, Sep. 2006.  
 [3] 松田 他: 信学論, Vol.J88-D-II, No.9, pp.1798-1807, Sep. 2005.  
 [4] 峯澤 他: 信学技報, Vol.107, No.40, pp.29-34, May 2007.  
 [5] G.J.Sullivan et al.: Proc. of SPIE Conf., Vol.5558, pp.53-74, Aug. 2004.