

# 高フレームレート映像信号に対する 時間方向ダウンサンプルフィルタリングの基礎検討

坂東幸浩 高村誠之 上倉一人 八島由幸

日本電信電話株式会社 NTT サイバースペース研究所

あらまし：近年、撮像系のフレームレートの増加は目覚しく、1000 Hz を超える高フレームレート映像が撮影可能となっている。しかし、現行ディスプレイに表示する場合、こうした高フレームレート映像は、フレームレートをダウンサンプリングする必要がある。本稿では、高フレームレート映像信号に対する効率的なダウンサンプリング方法について検証する。提案するダウンサンプリング方法は、フレーム間予測誤差を最小化する観点からダウンサンプリングフィルタを設計し、シーケンス全体での動き補償予測誤差の最小化を実現する方法である。平均値フィルタと比べて、提案手法はフレーム間予測画像の PSNR を 0.16 [dB] 向上させることが確認できた。

## 1 はじめに

近年、撮像系のフレームレートの増加は目覚しく、1000 Hz を超える高フレームレート映像が撮影可能となっている。一方、現行のディスプレイの上限は 120 Hz 程度であり、映像の入力・出力システムのフレームレートの上限は非対称である。1000 Hz 超の高フレームレート映像をスローモーション再生での観察等に用いる場合には、こうした入出力システムにおけるフレームレートの非対称性は問題にならない。一方、高速度カメラで撮影された映像を実時間再生で表示する場合には、フレームレートのダウンサンプリングが必要となる。

ダウンサンプリングの際、時間方向のフィルタリング [1], [2], [3] が施される。従来の 30 Hz, 60 Hz のフレームレートの映像の場合、複数フレームにまたがる時間軸方向のフィルタリングは、フィルタリング後のフレームに動きボケが発生するため、フィルタリングのための十分なサンプル（即ち、フレーム）を確保することが困難であった。例えば、60 Hz の映像信号をフィルタリングして 30 Hz の映像信号を生成する場合、フィルタリングの対象フレームに重複を許さない条件下では、フィルタリングの対象となるフレームは 2 フレームに限定される。一方、高フレームレート映像の場合、フィルタ設計の自由度は拡張される。例えば、1000 Hz の映像信号から 62.5 Hz の映像信号を生成する場合、フィルタリングの対象フレームに重複を許さない条件下であっても、16 フレームをフィルタリングの対象とすることができる。したがっ

て、高フレームレート映像を入力してフィルタリングにより低フレームレート映像を得る場合、フィルタリング設計の自由度が高まる。

そこで、本稿では、動き補償予測誤差を最小化するように、フィルタリング係数を設定する方法を提案する。これは、シーケンス全体での動き補償予測誤差の最小化を保証した方法であり、高フレームレート映像信号に対して、ダウンサンプリング後の低フレームレート映像信号の符号化効率を考慮した形で、低フレームレート映像信号を決定するフレームレート変換方法である。

## 2 適応ダウンサンプリング

フレーム間隔を  $\delta_t$  として、時刻  $t = j\delta_t$  ( $j = 0, 1, \dots$ ) のフレームにおける位置  $x$  の画素値を  $f(x, t)$  ( $x = 0, \dots, X-1$ ) と表す。この画素信号  $f(x, t)$  をダウンサンプリングにより、フレーム数を  $1/M$  に変換する場合を考える。つまり、フレームレートを  $\frac{1}{\delta_t}$  から  $\frac{1}{M\delta_t}$  へ変換する場合である。なお、以下では簡単のため、一次元信号を例にとり説明する。

タップ長  $2\Delta + 1$  のフィルタ係数  $\mathbf{W}_i = (w_i[-\Delta], \dots, w_i[\Delta])$  を用いたフィルタリングによりダウンサンプリング後のフレームを得ることとする。

$$\hat{f}(x, iM\delta_t, \mathbf{W}_i) = \sum_{j=-\Delta}^{\Delta} w_i[j] f(x, (iM + j)\delta_t)$$

なお、 $\sum_{j=-\Delta}^{\Delta} w_i[j] = 1$  を満たすものとする。さらに、フィルタ係数  $\mathbf{W}_i = (w_i[-\Delta], \dots, w_i[\Delta])$  が取り得る係数集合として、 $N$  通りの係数  $\Psi_n = (\psi_n[-\Delta], \dots, \psi_n[\Delta])$  ( $n = 0, \dots, N-1$ ) を考える。以下では、 $\hat{\Psi}_n = (\Psi_0, \dots, \Psi_{N-1})$  とおく。各フレームでは、この  $N$  通りの中から適切なフィルタ係数を選択するものとする。

以下、フィルタ係数の選択方法について、詳述する。信号  $\hat{f}(x, iM\delta_t, \mathbf{W}_i)$  に対して、サイズ  $S$  の区間  $B[k]$  ( $k = 0, 1, \dots, K-1$ ) に分割し、各区間  $B[k]$  を単位として動き補償（推定変位置量  $d_i = (d_i[0], \dots, d_i[K-1])$  を行った場合、同区間内の動き補償後の予測誤差を次式とする。

$$\sigma_i^2(\mathbf{W}_i, \mathbf{W}_{i-1}) = \sum_{k=0}^{K-1} \sum_{x \in B[k]} |\hat{f}(x, iM\delta_t, \mathbf{W}_i) - \hat{f}(x - d_i[k], (i-1)M\delta_t, \mathbf{W}_{i-1})|^2$$

$W_i, W_{i-1}$  が与えられた場合、 $\sigma_i^2[W_i, W_{i-1}]$  を最小化するように推定変位量  $d[k]$  ( $k = 0, 1, \dots, K-1$ ) を設定し、 $d[k, W_i, W_{i-1}]$  ( $k = 0, 1, \dots, K-1$ ) とおく。

求めるべきは、次式を満たす  $J/M$  組のフィルタ係数である。

$$(W_0^*, \dots, W_{J/M-1}^*) \\ = \min \arg W_0^*, \dots, W_{J/M-1}^* \sum_{i=1}^{J/M-1} \sigma_i^2[W_i, W_{i-1}] \quad (1)$$

各  $W_i$  が  $\Psi_n$  に含まれる  $N$  通りの候補から選択を行うため、 $J/M$  組のフィルタ係数 ( $W_0, \dots, W_{J/M-1}$ ) の取り得る組み合わせは、 $N^{J/M}$  通りであり、この中から最適な組み合わせ ( $W_0^*, \dots, W_{J/M-1}^*$ ) を総当りで探索するのは、計算量的に困難である。

そこで、 $\sigma_i^2[W_i, W_{i-1}]$  が直前フレームの結果に依存することに着目する。まず、 $W_i$  として取り得る全ての係数集合  $\Psi_n$  に対して、最適な  $W_{i-1}, \dots, W_0$  を用いた場合の  $\sum_{i=1}^i \sigma_i^2[W_i, W_{i-1}]$  を  $S_i(W_i)$  として定義する。ここで、 $\sigma_i^2[W_i, W_{i-1}]$  が直前フレームの結果のみに依存することに着目すると、 $S_i(W_i)$  は  $S_{i-1}(W_{i-1})$  を用いて、次式のように表わせる。

$$S_i(W_i) = \min_{W_{i-1}} \{\sigma_i^2[W_i, W_{i-1}] + S_{i-1}(W_{i-1})\} \quad (2)$$

式 (2) の漸化関係を用いることで、式 (1) の最小化問題は、次式のように表せる。

$$\min_{W_{J/M-1}} S_{J/M-1}(W_{J/M-1}) \quad (3)$$

上記のように定式化することで、式 (1) と等価な最適化を  $O(N^2)$  の計算量で算出することが可能となる。

### 3 実験

フレーム間引きの基準となる映像は、高速度カメラで撮影した RGB カラー映像であり、フレームレートは 1000[フレーム/秒]、総フレーム数は 480[フレーム]、解像度は  $640 \times 480$ [画素] である。また、映像素材はミニカートの走行シーンである。なお、動き補償時のブロックサイズは  $16 \times 16$ [画素] とした。また、フレーム間予測は片方向予測とし、参照フレームは直前フレームとした。動き推定アルゴリズムとしては、SSD 最小化に基づく全探索を行った。

$M = 32, \Delta = 1$  とした場合の予測画像の PSNR の比較結果を表 1 に示す。ここで、カッコ内は平均二乗誤差の値である。平均値フィルタは、フィルタ係数を  $(1/3, 1/3, 1/3)$  としたフィルタである。提案手法については、 $N$  の値と

Table 1: 予測画像の平均 PSNR[dB] (括弧内の数値は MSE の値)

平均値フィルタ	29.067 (80.61)
提案法 ( $N=16$ )	29.220 (77.82)
提案法 ( $N=64$ )	29.224 (77.75)

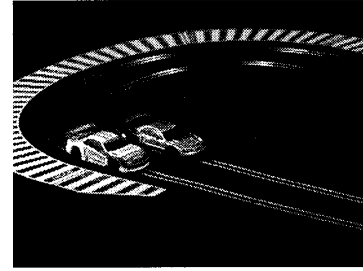


Figure 1: 実験に用いたシーケンス

して  $N = 16, 64$  の 2 通りを設定した。この結果、提案手法は、平均値フィルタに対して 0.16 [dB] の PSNR 向上を実現していることが確認できる。

### 4 おわりに

本稿では、高フレームレート映像に対する適応的なダウンサンプリング方法を提案した。実験を通して、単一の平均値フィルタと比較して、0.16 [dB] の PSNR 向上が確認できた。

### 参考文献

- [1] J. Ohm, "Advances in Scalable Video Coding", *Proceedings of the IEEE*, Vol. 93, No. 1, pp. 42–56, 2005
- [2] A. Golwelkar and J. Woods, "Motion-Compensated Temporal Filtering and Motion Vector Coding Using Biorthogonal Filters" *IEEE Trans. on CSVT*, Vol. 17, No. 4, pp. 417–428, 2007
- [3] H. Schwarz, D. Marpe, and T. Wiegand, "Hierarchical B Pictures", Joint Video Team, Doc. JVT-P014, Poznan (Poland), July 2005.

"A study of temporal down-sample filtering for high frame-rate video",  
Yukihiko Bandoh, Seishi Takamura, Kazuto Kamikura, and Yoshiyuki Yashima,  
NTT Cyber Space Laboratories, NTT Corporation