

複数フレーム参照によるフレーム内挿法とその映像符号化への応用 A Multi-Frame based Frame Interpolation and the Application to Video Coding

三須 俊枝†

松尾 康孝†

岩村 俊輔†

井口 和久†

境田 慎一†

Toshie Misu

Yasutaka Matsuo

Shunsuke Iwamura

Kazuhisa Iguchi

Shinichi Sakaida

1. まえがき

映像の臨場感や実物感を高めるため、これまで空間解像度の高精細化を主軸に技術開発が進められてきた。一方、とくにスポーツコンテンツなど動きの激しい映像においては、高フレームレート化が映像品質向上に有効であることから、ハイビジョン以上の空間解像度と 120Hz 以上のフレームレートをサポートするカメラやディスプレイが世に出始めている。NHK では空間解像度 7680×4320 画素、フレームレート 120Hz (プログレッシブ)、ビット深度 12 ビットをフルスペックスーパーハイビジョンと定義して、撮像、符号化、伝送、表示等の研究・開発を進めている[1]。

しかし、映像の高品位化の過渡期にあっては、仮に映像ソースや表示デバイスが高フレーム化されていても、映像符号化装置の処理や伝送路の帯域が追い付かず、やむを得ずフレームを間引いて符号化伝送することも考慮する必要がある。

そこで本稿では、伝送されたフレームから間引かれたフレームを生成するのに好適な補間方法を送信側で求め、サイド情報として伝送することで、受信側での高精度な補間フレーム生成を補助する符号化伝送方式について提案する。

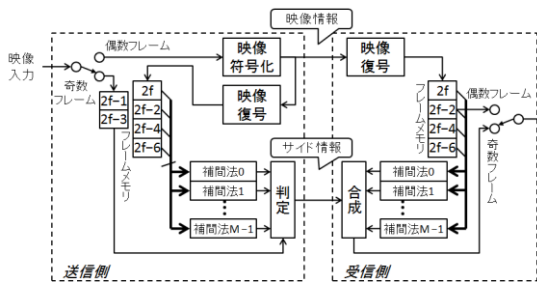


図 1. 映像符号化伝送システムの構成

2. 映像符号化システムの構成

図 1 に提案する映像符号化伝送システムの構成を示す。本システムでは、入力映像を所定フレーム間隔で間引き、低フレームレート化した映像を伝送する。本稿では、2 フレームにつき 1 フレームを間引くこととし、偶数フレームのみを伝送するものとして説明する。

一方、間引かれた奇数フレームは、受信側において復号された偶数フレームに基づいて補間する。本システムでは複数の異なるフレーム補間法を実装し、ブロック単位で適切な手法を組み合わせることで精度の高い補間結果を実現する。

送信側には、受信側と同じ複数のフレーム補間法を備

†日本放送協会, NHK

え、間引かれたフレームの原画を参照しつつ各補間法の予測誤差を評価することで、フレーム毎・ブロック毎にいずれの補間法の効果が高いかを求める。こうして求められた有効／無効のフラグをサイド情報として伝送し、受信側の補間法の組み合わせを制御することで高画質な補間フレームを得ることができる。

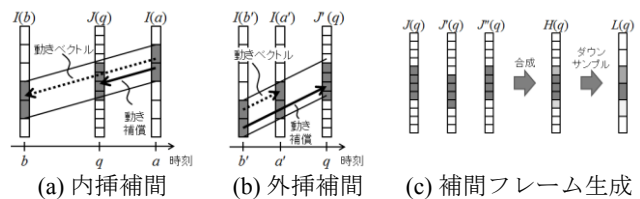


図 2. フレーム補間の原理

3. 複数フレーム参照によるフレーム補間法

本方式では、ブロック毎に複数のフレーム補間法を切り替えるのではなく、複数の補間法の結果を重ねることで、ぼやけやアンカバ領域の画質劣化の改善を図った。

図 2 は提案する画像重畳によるフレーム補間の原理である。本手法では、時刻 a における動き推定元フレーム $I(a)$ と、時刻 b における動き推定先フレーム $I(b)$ とに基づき、時刻 q における仮の補間結果 $J(q)$ を求める。このとき、時刻 q に対して時間の前後や時間間隔の異なるさまざまな時刻対 (a, b) (ただし、 $a \neq b$) を用いて、内挿補間 (図 2(a)) や外挿補間 (図 2(b)) を実行する。

こうして仮の補間結果を複数生成し、図 2(c) に示すようにこれら仮の補間結果を重ね合わせることで合成補間フレーム $H(q)$ を生成する。このとき、重ね合わせに採用すべき仮の補間結果の組み合わせをブロック毎に制御することで、信頼性のない仮の補間結果を排除する。

補間結果の組み合わせ方の制御信号は、符号化側からサイド情報にて伝送される。符号化側においては、各仮の補間結果 $J(q)$ が真の間引きフレーム $I(q)$ と符合するか否かをブロック毎に検証し、組み合わせるべき補間法を導出する。

なお、仮の補間フレーム $J(q)$ と合成補間フレーム $H(q)$ の解像度は、入力画像の解像度よりも高くとることにより、アンチエイリアシングや複数フレーム超解像の効果が期待できる。最後に、合成補間フレーム $H(q)$ を入力画像の解像度までダウンサンプルし、出力補間フレーム $L(q)$ として出力する。

3.1 送信側の動作

まず、送信側において、偶数フレームを符号化し、ローカル復号する。ローカル復号画像を I' とおく。

補間法 m の参照する動き推定元フレームの時刻を $a^{(m)}$ 、動き推定先フレームの時刻を $b^{(m)}$ とそれぞれおく。

動き推定元フレーム $I'(a^{(m)})$ をブロック分割し、 k 番目のブロック領域 B_k について類似領域を動き推定先フレーム $I'(b^{(m)})$ の中から探索 (探索領域 C) する。その結果得られる動きベクトルを $(u_k^{(m)}, v_k^{(m)})$ とおく。例えば、絶対値誤差和 (SAD) の最小化による場合には、下式により動きベクトルを求めることができる。

$$(u_k^{(m)}, v_k^{(m)}) = \arg \min_{(u,v) \in C} \sum_{(x,y) \in B_k} |I'(b^{(m)}; x+u, y+v) - I'(a^{(m)}; x, y)|$$

続いて、動きベクトル $(u_k^{(m)}, v_k^{(m)})$ を比例按分し、当該ブロック B_k のテクスチャを補間時刻 q (q は奇数) まで動き補償したときの動き補償誤差 $e(k, m)$ を、原画 $I(q)$ を参照しつつ下式により評価する。

$$e(k, m) = \sum_{(x,y) \in B_k} \left| I \left(q; x + \frac{q-a^{(m)}}{b^{(m)}-a^{(m)}} u_k^{(m)}, y + \frac{q-a^{(m)}}{b^{(m)}-a^{(m)}} v_k^{(m)} \right) - I'(a^{(m)}; x, y) \otimes h(x, y) \right|$$

ここに、 h は小数画素精度の動き補償を行うために畳み込む点拡がり関数であり、例えば標準偏差 0.5 画素の 2 次元ガウシアンカーネルを用いる。

動き補償誤差 $e(m, k)$ の大きなブロックはフレーム補間に用いるべきではないことから、動き補償誤差 $e(m, k)$ を閾値判定し、補間法 m 毎、ブロック B_k 毎の採否判定 $F(m, k)$ (採用を 1 とする) をサイド情報として受信側へ伝送する。

$$F(m, k) = \begin{cases} 1, & e(m, k) \leq \theta \\ 0, & e(m, k) > \theta \end{cases}$$

なお、 θ は採否判定の閾値 (例えば、動き推定元と動き推定先のブロックの平均画素値の 30%) である。

3.2 受信側の動作

受信側においても、送信側同様に復号された 2 フレーム $I'(a^{(m)})$ および $I'(b^{(m)})$ から動きベクトル $(u_k^{(m)}, v_k^{(m)})$ を求める。

入力画像よりも高解像な 2 つのバッファ $S_k^{(m)}$ および $W_k^{(m)}$ を確保し、 $S_k^{(m)}$ にはブロック B_k に関する動き補償結果に加重を施したものの和を、 $W_k^{(m)}$ には加重値自体の和を、下式により漸次的に求める。この際、採否判定 $F(m, k)$ が 1 のもののみを重ね合わせることにする。

$$\begin{aligned} S_k^{(m)}(q; x + \lambda^{(m)} u_k^{(m)}, y + \lambda^{(m)} v_k^{(m)}) \\ = S_{k-1}^{(m)}(q; x + \lambda^{(m)} u_k^{(m)}, y + \lambda^{(m)} v_k^{(m)}) + \mu^{(m)} \cdot I'(a^{(m)}; x, y) \otimes h(x, y) \end{aligned} \quad (1)$$

$$\begin{aligned} W_k^{(m)}(q; x + \lambda^{(m)} u_k^{(m)}, y + \lambda^{(m)} v_k^{(m)}) \\ = W_{k-1}^{(m)}(q; x + \lambda^{(m)} u_k^{(m)}, y + \lambda^{(m)} v_k^{(m)}) + \mu^{(m)} \cdot h(x, y) \end{aligned} \quad (2)$$

$$\begin{aligned} \text{for } \forall (x, y) \in B_k \quad \text{iff } F(m, k) = 1 \\ \text{for } \forall k \in \{1, 2, \dots, K\} \end{aligned}$$

$$\text{where } \lambda^{(m)} = \frac{q - a^{(m)}}{b^{(m)} - a^{(m)}}, \quad \mu^{(m)} = \frac{1}{\sqrt{(q - a^{(m)})(b^{(m)} - a^{(m)})}}$$

$$S_0^{(m)}(q; x, y) = W_0^{(m)}(q; x, y) = 0$$

なお、 $\lambda^{(m)}$ は按分の比率、 $\mu^{(m)}$ は動き推定および動き補償の時間間隔が開くほど信頼度が下がることを反映した重みづけである。

ここで、仮の補間フレーム $J^{(m)}$ は、

$$J^{(m)}(q; x, y) = S_K^{(m)}(q; x, y) / W_K^{(m)}(q; x, y)$$

により求められるが、今回の実装では仮の補間フレーム $J^{(m)}$ を陽には求めず、以下の手順で合成補間フレーム H を導出する。

$$H(q; x, y) = \sum_{m=1}^M S_K^{(m)}(q; x, y) / \sum_{m=1}^M W_K^{(m)}(q; x, y)$$

なお、いずれの補間法によっても補うことのできなかった画素 ($W_K^{(m)}(q; x, y) = 0$ なる画素) については、前後に隣接する時刻の 2 伝送フレームに対して線形補間 (すなわち画素ごとの画素値の平均) を行った結果を補間フレームの画素値として採用することとする。

最後に合成補間フレームを入力映像の解像度にダウンサンプルし、出力補間フレーム $L(q)$ を得る。伝送されてくる偶数フレームについては復号された当該フレーム I' に、伝送されてこない奇数フレームについては補間フレーム L に、交互にフレームを切り替えつつ時間多重を行うことで、元の映像のフレームレートの映像を得ることができる。

表 1. 実験に用いた補間法 (f は伝送フレームの連番)

補間法番号 m	動き推定元 a	動き推定先 b	補間時刻 q	意味	図 4 のビット位置
0	$2f$	$2f-2$	$2f-3$	3:1 前方外挿	01h
1	$2f-6$	$2f-4$	$2f-3$	3:1 後方外挿	02h
2	$2f$	$2f-4$	$2f-3$	3:1 前方内挿	04h
3	$2f-4$	$2f$	$2f-3$	1:3 後方内挿	08h
4	$2f-2$	$2f-6$	$2f-3$	1:3 前方内挿	10h
5	$2f-6$	$2f-2$	$2f-3$	3:1 後方内挿	20h
6	$2f-2$	$2f-4$	$2f-3$	1:1 前方内挿	40h
7	$2f-4$	$2f-2$	$2f-3$	1:1 後方内挿	80h

表 2. 実験条件

解像度	1920×1080 画素
ビット深度	8 ビット, グレースケール
原画フレームレート	60Hz, プログレッシブ
原画フレーム数	10 フレーム
伝送映像フレームレート	30Hz, プログレッシブ
伝送映像フレーム数	5 フレーム
ブロックサイズ	32×32 画素
ブロック数	2040 ブロック/フレーム
映像符号化方式	MPEG-4 AVC H.264 (JM16.0)
GOP 構造	IBBPBP

4. 実験

提案手法の動き補償補間法として、表 1 に示す補間法番号 1~8 の 8 種類を実装し、映像符号化の実験を行った。表 1 の 8 種類には、動き推定の因果の向き、補間時刻と 2 参照時刻が等間隔か否か、内挿か外挿かで相違するものを取り交ぜている。なお、本稿では、動き推定が未来から過去へ向かう因果を「前方」と称している。

実験条件を表 2 に示す。映像は、第 0, 2, ..., 8 フレームを伝送し、第 1, 3, ..., 9 フレームを間引くこととした。

実験結果 (第 5 フレーム) を図 3 に示す。(a) は原画像、(b) は線形内挿の結果、(c) は直後フレームから直前フレームへ求めた動きベクトルに基づき動き補償補間を行った結果 (但し、アンカバー領域には線形内挿結果を使用)、(d) は提案する補間法の結果である。各画像の右には部分拡大画像を載せた。なお、映像符号化においては、サイド情報を含むビットレートがほぼ同一となるように量子化パラメータを設定している。

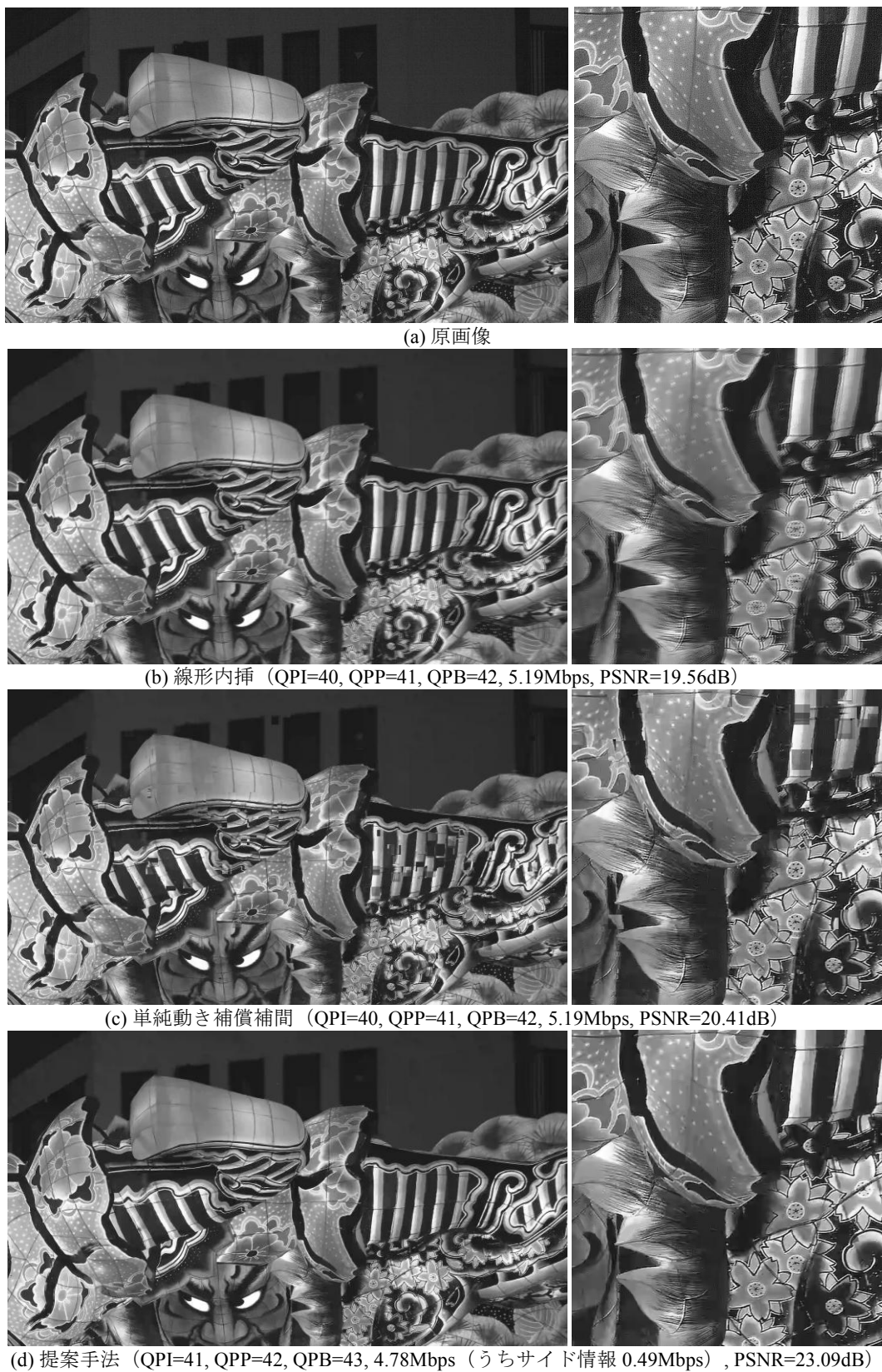


図 3. 実験結果

線形内挿(b)では動領域は二重にぶれた結果になっている。単純動き補償補間(c)ではぶれは改善するもののアンカバー領域や補間エラーがブロック歪みとして現れてい

る。提案手法(d)ではブロック歪みが削減されシャープな補間画像が得られている。客観評価においても提案手法が最も高い PSNR 値を得ている。



図 4. 採否フラグ (0.49Mbps)

図 5. 提案手法においてローカル復号を省略した場合
(QPI=41, QPP=42, QPB=43, 4.78Mbps (うちサイド情報 0.49Mbps), PSNR=22.99dB)

図 4 は、サイド情報として伝送される採否フラグ情報である。各ブロックに記した 16 進数がブロック毎の採否フラグを表している。ここでは、各補間法の採否を表 1 に示すビット位置にフラグを立てることで 8 ビット/ブロックの採否フラグとして表現している。多くの領域では全手法が採用 (フラグ値 FFh) されているが、楕円で示すようなアンカー領域では補間法の取捨が行われている (フラグ値 FFh 以外) ことがわかる。なお、この例では、ブロック当たり 8 ビットのフラグを、フレームあたり 2040 個のブロック、30 フレーム/秒で伝送するため、サイド情報は非圧縮で 489.60kbps となる。

本手法では、図 1 に示したように送信側にローカル復号を行っている。このようにすることで符号化劣化を含めた補間法の最適化が可能となる反面、復号で生じる遅延分だけ総遅延量が増加する。映像符号化における画像劣化が無視できる場合にはローカル復号を省略し、符号化前の伝送フレームを参照して補間法を最適化しよう構成してもよい。図 5 は、ローカル復号を省略した場合の補間フレームである。図 3(d) に比べてブロック歪みの増加および PSNR の低下が確認できるが、図 3(c) の単純動き補償補間に比べれば良好な画質が得られている。

4. まとめ

複数のフレーム補間法を組み合わせることで、高品質な補間フレームを生成する手法を提案した。また、本手法の応用として、いずれの補間法が有効であるかを送信側で最適化しサイド情報として伝送する映像符号化システムを提案し、その有効性を実験により示した。今後、画像縮小・超解像を利用する映像符号化方式 [2] と、本稿の映像符号化方式とを組み合わせ、フルスペックスーパーハイビジョンなど高解像・高フレームレート映像向けの映像符号化方式を開発していく。

参考文献

- [1] Y. Nishida, "Present state of Super Hi-Vision," KOBA 2012.
- [2] T. Misu, Y. Matsuo, S. Sakaida, Y. Shishikui, and E. Nakasu, "Novel video coding paradigm with reduction/restoration processes," *Proceedings of PCS 2010*, S2-2, pp. 466–469, December 2010.