

画像復元型符号化における折り返し歪の適応付加

Encoder-side Adaptive Aliasing Embedment for Reconstructive Video Coding

三須 俊枝†
Toshie Misu

松尾 康孝†
Yasutaka Matsuo

境田 慎一†
Shinichi Sakaida

鹿喰 善明†
Yoshiaki Shishikui

1. まえがき

現在、一般的に用いられている映像符号化方式は、MPEG-2, MPEG-4 AVC/H.264をはじめ、動き補償予測と直交変換(DCTなど)とをベースとした手法が主流である。しかし、スーパーハイビジョン[1]など超高精細映像を家庭に届けるためには1/100~1/1000オーダの超高圧縮率が求められ、従来型の映像符号化手法では絵柄によっては極端な画像劣化が生じることがある。

そこで我々は、図1に示すように、予め縮小変換により画像の標本点数を間引いてから従来型の映像符号化を行い、復号画像に対して超解像を行うことで元の解像度に復元する「画像復元型符号化」を提案している[2][3]。

画像復元型符号化では、符号化側において原画像が既知であるため、縮小変換と超解像の対を原画像の特性に応じてパラメータ調整することが可能である。我々は、符号化側で復号および超解像を試行し原画像と比較することでパラメータを最適化するローカルデコード型のアプローチ[3]と、符号化側で標本化定理などに基づいて一方的にパラメータを決定するフィードフォワード型のアプローチ[2]の両面から検討を進めている。

フィードフォワード型のアプローチについては、テクスチャ複雑度に応じた標本化間隔制御について既に[2]で報告したが、今回は動静に応じた折り返し歪制御を導入し、画質の向上を認めたので報告する。

2. 画像復元型符号化

画像復元型符号化方式の構成を図1に示す。本方式は、入力画像を画像縮小してから従来方式により符号化を行う。復号側においては、復号された縮小画像に対し超解像処理を行うことで元の解像度の画像を得る。

画像縮小によって符号化すべき標本点数が削減されることから、同一ビットレートにおいては、入力画像をそのまま従来符号化する場合よりも量子化ステップが細くなる。その結果、ブロック歪等の不自然な画像劣化を抑えることができる。

しかし、事前の画像縮小において高域情報が失われるため、復号画像を単純に拡大すると画像にぼやけが生じる。画像復元型符号化では画像拡大に超解像を用いることで高域を補い、鮮鋭な画像を得る。このとき、符号化側で入力画

像の特徴や超解像の最適なパラメータを抽出し、サイド情報として復号側へ伝送し、超解像復元を補助することで、一層の画質向上が可能となる。

2.1 折り返し歪と複数枚超解像

超解像技術は、中低域から高域を推定して補う手法、自己相似性・合同性に基づく手法、学習により得た低解像と高解像の画像パターンの対を用いる手法など、画像信号の特性をモデル化したいわゆる一枚超解像と、複数フレームの小数精度の位置合わせによって高域を復元するいわゆる複数枚超解像に大別できる。

複数枚超解像は、入力画像に折り返し歪が含まれ、かつ画像間に小数画素のずれがあることが前提である。図2

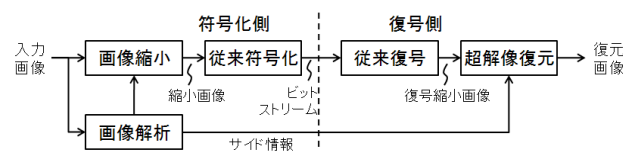


図1. 画像復元型符号化の構成

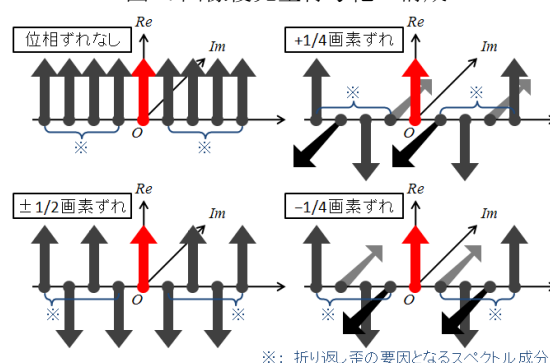


図2. 標本化格子のスペクトル

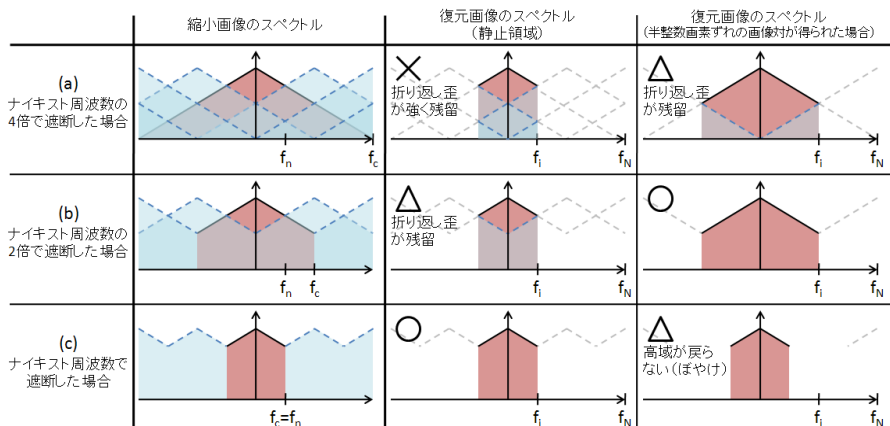


図3. 画像縮小における折り返し歪と超解像結果

†日本放送協会, NHK

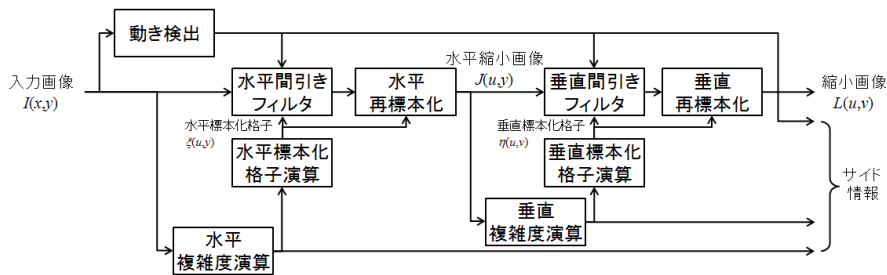
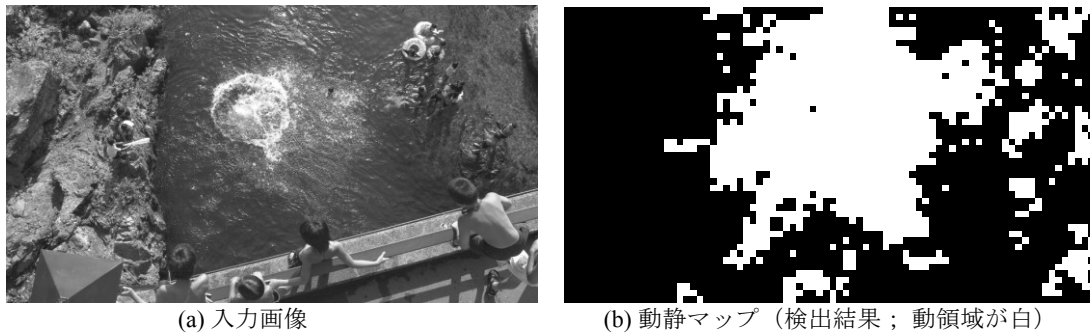


図4. 画像縮小の手順



(a) 入力画像

(b) 動静マップ (検出結果; 動領域が白)

図5. 動き検出の結果

に示すように、基準となる画像の標本化格子に対し、参照画像の標本点が小数画素ずれている場合、そのずれ量に応じて折り返し歪の要因となる標本化格子のスペクトル成分(※印)の位相だけが回転する。ずれた格子で標本化した画像間ではベースバンド信号は不変だが、折り返し歪は互いに位相がずれるため、これらの画像を重ね合わせることで、折り返し歪を相殺し、高域成分を復元することができる。

2.2 画像復元型符号化における折り返し歪制御

通常の超解像技術では与えられた縮小画像のみから高解像画像を推定するのに対し、画像復元型符号化では超解像しやすいように縮小画像の作り方を制御することやその制御情報を既知として利用することが可能である。

本稿では、復号側で複数枚超解像を行うことを前提とし、符号化側で動きに応じた折り返し歪付加の制御を導入する。

図3に原画像を1/4に縮小(再標本化)する場合について、画像縮小時の折り返し歪量と復元画像の関係を周波数軸上で例示する。(a)は間引きフィルタを適用せずに縮小した場合、(b)はナイキスト周波数 f_n の2倍を遮断周波数 f_c とした場合、(c)はナイキスト周波数を遮断周波数とした場合の例である。

図3の中列と右列は、動きが全くない場合と、ちょうど半整数画素の動きがあった場合それぞれについて、理想的な復元が行われたときのスペクトルを対比している。

動きが全くなかった場合には、再標本化時に生じたナイキスト周波数 f_n を超え遮断周波数 f_c 未満のスペクトルの折り返し歪は、複数枚超解像によって取り除くことができないため、復元画像には折り返し歪がそのまま残存する。

動きがある場合には、複数枚超解像により、動き量、動き補償の確度、利用できる画像枚数に応じて、折り返し歪の一部乃至全部を相殺することができる。

すなわち、折り返し歪の残留を避けつつ、できる限り精細な復元画像を得るためには、

- 静領域の遮断周波数はナイキスト周波数ちょうど
- 動領域の遮断周波数はナイキスト周波数超

に設定するのが望ましいといえる。

3. 動静に応じた折り返し歪制御

折り返し歪の制御方法を図4に示す。まず、ブロックごとに画像動静を判定する。続いて、水平方向と垂直方向について縦続的に画像縮小する。水平縮小の際には、ブロックごとに水平方向の複雑度を求め、複雑な領域ほど細かく再標本点を配置する。動静に応じて間引きフィルタの帯域を制御しつつ水平方向の再標本化を行う。垂直方向についても、垂直方向の複雑度に基づき同様の処理を行う。

3.1 動き検出

まず、入力画像をブロック $A_{i,j}$ に分割し、ブロックごとに動きの有無 $Q(t; i, j)$ を判定する。

本稿では、ブロック内のフレーム間差分の絶対値の総和に対する閾値処理により動静を判定した。

$$Q(t; i, j) = \begin{cases} 1, & \left[\sum_{x \in A_{i,j}} |I(t, x, y) - I(t-1; x, y)| \right] / |A_{i,j}| > \theta \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

図5に動き検出の結果(動静マップ)を示す。図5は、 3840×2160 画素の画像を 48×48 画素のブロックで分割し、閾値を $\theta = 10$ とした場合の動き検出結果である。中央付近の水しぶきと、下側の人物が動領域として検出されている。

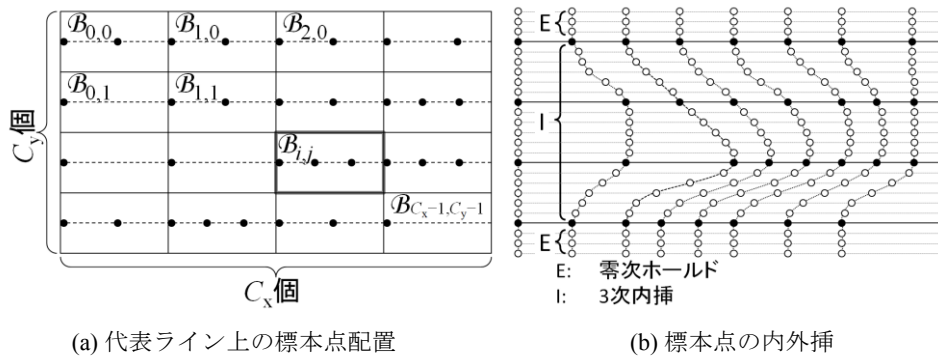


図 6. 水平標本化格子の生成

3.2 標本化格子

再標本化は不均一間隔の標本化格子により、水平、垂直独立に順次実行する。以下では水平方向の処理に関して説明する。

まず、ブロック $B_{i,j}$ (動き検出のブロック $A_{i,j}$ とは異なるサイズとして構わない) ごとの水平方向のトータルバリエーションノルムを水平複雑度 $M_H(i,j)$ として定義する。

$$M_H(i,j) = \sum_{x \in B_{i,j}} |I(t,x+1,y) - I(t,x,y)| \quad (2)$$

続いて、行方向に並んだブロック群について、それらの中央を貫く水平ライン (代表ラインと呼ぶ; 図 6(a)の点線) 上に再標本化用の標本点群を配置する。このとき、各ブロックの水平重要度 $M_H(i,j)$ に比例する密度で標本点を配置した結果を図 6(a)の黒丸で示す。

代表ライン以外の水平ラインに対しては、上下の代表ライン上の標本点位置を三次内挿 (上下端の代表ラインの外側については零次ホールド) することで、標本点群を配置する (図 6(b))。

3.3 間引きフィルタ

3.1 で検出した動きと、3.2 で求めた標本化格子の粗密とに応じ、水平間引きフィルタのタップ係数を決定する。

以降では、間引きフィルタの遮断周波数 f_c を、 $f_c = \alpha f_n$ のようにナイキスト周波数 f_n の倍率 α で表わすこととする。

倍率 α は、再標本化の画素ごとに変化させる。3.1 で求めた動きの有無 $Q(t; i, j)$ に応じて、ブロック $A_{i,j}$ の中心点に対しては、

$$\alpha_{i,j} = \begin{cases} 1, & L(t; i, j) = 0 \\ \beta, & L(t; i, j) = 1 \end{cases} \quad (2)$$

により倍率 $\alpha_{i,j}$ を定める ($\beta > 1$ とする)。これは、静領域に対してはナイキスト周波数ちょうどで遮断し、動領域に対してはナイキスト周波数の β 倍で遮断することを意味する。ブロック $A_{i,j}$ の中心点以外の位置にある再標本化点に対しては、近傍ブロック中心点における遮断周波数を双一次内挿することで定める。

続いて、不均一間隔の標本化格子に対して間引きフィルタを適用する。以降では、入力画像を I 、水平再標本化後の水平縮小画像を J とおく。また、入力画像 I の元々の標本点を原標本点、再標本化後の標本点を再標本点と称することとする。

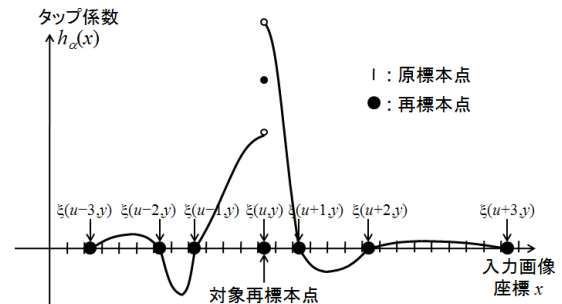


図 7. 間引きフィルタのタップ係数例 ($K=3$ の場合)

入力画像 I の注目ライン (垂直座標 y) 上における、左から u 番目の再標本点の水平座標を $\xi(u,y)$ とおく。水平縮小画像の標本値 $J(u,y)$ は、入力画像 I 内の線分 $\{(x,y) | x \in [\xi(u-K,y), \xi(u+K,y)]\}$ 上の原標本点列に対して次式のタップ係数 $h(x)$ の積和をとることで決定する。なお、 K はフィルタカーネルの大きさを決める定数 (自然数) であり、以降では $K=3$ とする。

$$h_k(x) = \begin{cases} g_\alpha \left(\frac{x - \xi(u+k,y)}{\xi(u+k+1,y) - \xi(u+k,y)} + k \right), & \xi(u+k,y) < x < \xi(u+k+1,y) \\ \frac{1}{2} \left(\frac{1}{\xi(u,y) - \xi(u-1,y)} + \frac{1}{\xi(u+1,y) - \xi(u,y)} \right), & k = x = 0 \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

$$h(x) = \frac{\sum_{k=-K}^{K-1} h_k(x)}{\sum_{x=\xi(u-K,y)}^{\xi(u+K,y)} \sum_{k=-K}^{K-1} h_k(x)} \quad (4)$$

ここに、 $g_\alpha(x)$ は、水平方向 $1/\alpha$ 倍にスケールした Sinc 関数

$$g_\alpha(x) = \begin{cases} 1 & (x = 0) \\ \sin(\alpha\pi x) / (\alpha\pi x) & (x \neq 0) \end{cases} \quad (5)$$

とする。このタップ係数 $h(x)$ は、図 7 に例示するように、Sinc 関数 $g_\alpha(x)$ を、 x が整数の点で分割し、分割結果の各セグメントの端が再標本点にフィットするように水平に伸縮し、振幅を水平伸縮率の逆数で正規化したものに相当す

る。なお、再標本化が均一間隔の場合には、 $g_{\alpha}(x)$ は遮断周波数 αf_n の Truncated Sinc 型低域通過型フィルタとなる。

3.4 再標本化

3.2で求めた再標本点ごとに、3.3の間引きフィルタの遮断周波数 $f_c = \alpha f_n$ を変化させつつ再標本化を行う。その結果を水平縮小画像 $J(u,v)$ とする。

3.5 垂直方向の処理

水平縮小画像 $J(u,v)$ に対し、3.2から3.4までで述べたものと同様の処理を垂直方向に適用することで、水平および垂直に不均一再標本化された縮小画像 $L(u,v)$ を得る。

3.6 サイド情報

超解像復元を補助するため符号化側から復号側へ伝送される情報である。サイド情報は、動き検出結果 $Q(t; i, j)$ 、水平複雑度 $M_H(i, j)$ および垂直複雑度 $M_V(i, j)$ で構成される。動き検出結果 $Q(t; i, j)$ は毎フレーム伝送する必要があるが、水平複雑度 $M_H(i, j)$ および垂直複雑度 $M_V(i, j)$ は間欠的に伝送しても構わない。

水平複雑度 $M_H(i, j)$ および垂直複雑度 $M_V(i, j)$ を更新すると縮小画像 $L(u,v)$ はその都度非線形の変形を受ける。その結果、従来符号化の動き補償が効きにくくなる恐れがある。このため、水平複雑度 $M_H(i, j)$ および垂直複雑度 $M_V(i, j)$ の更新はイントラピクチャ（または、イントラスライス）のタイミングで行うことが望ましい。

4. 超解像復元

縮小画像は従来符号化方式により符号化して伝送され、復号される。復号後の縮小画像を元の解像度に戻す画像復元処理には超解像技術を用いる。

縮小過程で折り返し歪や再標本化格子が空間的に不均一に設定されることから、周波数軸で中低域から高域を推定する超解像法や自己相似性を用いる超解像法は適さない。このような問題を解くためには、縮小画像生成過程の逆問題を解く再構成型の超解像技術を用いることが適当である。

我々は、縮小過程の逆問題を逐次モンテカルロ法で解く逐次モンテカルロ再構成型超解像 (SMC 超解像) を提案している[4]。本手法では、不良設定問題を解くための正則化 (トータルバリエーションノルムを使用) と、複数フレームのレジストレーションとを確率過程としてモデリングしている。これにより、動領域に対しては折り返し歪から高域が復元する複数枚超解像を、静領域に対してはノルム正則化の一枚超解像を、それぞれシームレスに適用することができる。

SMC 超解像の構成を図8に示す。

4.1 超解像結果の仮説生成

SMC 超解像では、超解像結果の候補である複数の画像を逐次更新することによって最適化を行う。この候補を仮説と称する。

仮説は、提案分布という確率密度分布からのサンプル抽出により生成する。本稿では、

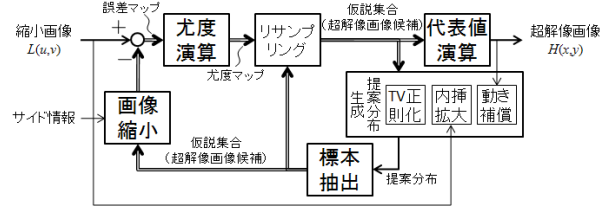


図8. 逐次モンテカルロ再構成型超解像 (SMC 超解像)

- 前ステップまでの仮説に対して TV ノルムを過剰に増加させない雑音を付加することで高域を付加した新仮説
- 前ステップまでの超解像画像と、現ステップにおける入力画像との動き補償によって折り返し歪からの高域を復元した新仮説
- 現ステップの入力画像を Lanczos-3 補間により拡大した新仮説

を生成する提案分布からのサンプル抽出により、これら3種類の仮説を織り交ぜて生成する[2]。これにより、一枚超解像の適する静領域に対しては TV ノルム正則化法が、複数枚超解像の適する動領域に対しては画像位置合わせによる高域復元が、それぞれシームレスに適用される。また、これら超解像手法が不安定になる領域に対しては、Lanczos-3 補間がバックアップし、最適化結果が不安定化することを防ぐことができる。

4.2 画像縮小と尤度演算

仮説に対して符号化側における画像縮小と同様の処理を模擬し、得られた縮小画像と、入力された縮小画像 L との比較を行うことで、仮説の妥当性 (尤度) を評価する。

このため、SMC 超解像の画像縮小ブロックは、符号化側の画像縮小と挙動が一致するようサイド情報を通じて制御される。

このアーキテクチャでは、原理的には任意の種類 of 画像縮小処理の逆処理を実行することができる。このため、例えば、折り返し歪の付加量や標本化格子を空間的に変化させるなどの非線形操作に対しても、復元処理が可能である。

4.3 リサンプリング

逐次モンテカルロ法では仮説毎に重みが定義され、尤度に比例して重みを乗法的に更新する。尤度の高かった (観測と矛盾のなかった) 仮説の重みは重く、尤度の低かった仮説の重みは軽くなる。

今回は、画像を低解像画像の1画素に相当する超解像画像の部分領域ごとに独立に尤度を判定し、該尤度に比例して前記部分領域を複製/削除を行うことで仮説の再編成を行った。詳細は文献[4]を参照されたい。

4.4 代表値演算

以上の演算を各フレーム所定回数繰り返す。収束演算を行う。最後に、すべての仮説の期待値を求め、これを超解像画像として出力する。

表1. ビットレート配分

手法	映像 [Mbps]	重要度 マップ [Mbps]	動静 マップ [Mbps]	合計 [Mbps]
(a) AVC/H.264	28.62	—	—	28.62
(b) 折り返し歪 固定 $\alpha=2$	30.15	0.03	—	30.18
(c) 折り返し歪 制御 $\alpha=1-2$	28.00	0.03	0.22	28.25

5. 実験

従来の MPEG-4 AVC/H.264, 文献[2]の手法(折り返し歪適応制御なし), および提案手法(折り返し歪適応制御あり)の各手法により, 符号化・復号の実験を行った.

実験には 3840 x 2160 画素, 60Hz の 8 ビットグレースケールの動画を 15 フレーム用いた. この動画は, 非圧縮時には約 3.98Gbps に相当する. ビットレートはいずれも約 28Mbps とし, GOP 構造は IBPPBP... とした.

実験結果を図 9 (部分切り出し) に示す. (a) の従来方式では, 頭髪部分などなだらかなテクスチャを有する箇所に著しいブロック歪が見られるほか, 水面の泡のようなランダムに移動する細かいテクスチャも失われている.

折り返し歪量を画面全体で一定とした(b)では, 水面や頭髪などのブロック歪が抑制され, 画質が改善している. しかし, 橋の端の一部分において折り返し歪の残留が目

立っている.

提案手法(c)では, 水面や頭髪部分のブロック歪を抑制しつつ, 橋の端の折り返し歪も目立たなくすることができている.

図 9 の各結果のビットレートとその配分を表 1 に示す. サイド情報は, 標本化格子を決定するための重要度マップと, 折り返し歪を制御するための動静マップである. 表 1 には, これらサイド情報を毎フレーム非圧縮伝送した場合のビットレートを記している. 重要度マップを時間的に間引くことや, サイド情報の予測やエントロピー符号化を適用することにより, サイド情報の割合はさらに削減できると考えられる.

図 10 は, 本手法をカラー画像に適用した場合の結果である. 提案手法の標本化と, 従来符号化部分における色差標本化との干渉を避けるため, 今回は GBR4:4:4 の形式 (GBR の色空間を用い, 各色のサンプリング構造をすべて同一とする形式) によって処理・符号化を行った. また, 色ずれを防ぐため, 折り返し歪制御と標本化格子は全色共通とした (グリーン of の画像に基づき, 重要度マップや動静マップを生成し, これらをブルーやレッドにも流用した). ビットレートは, 図 10(a)~(d) が約 110Mbps, (e) および (f) が約 20Mbps である.

いずれの場合においても, 従来方式では平坦な絵柄のテクスチャがほとんど失われ, 著しいブロック歪が生じているのに対し, 画像復元型符号化方式ではこうした劣化が抑えられている.



(a) 従来方式 MPEG-4 AVC/H.264

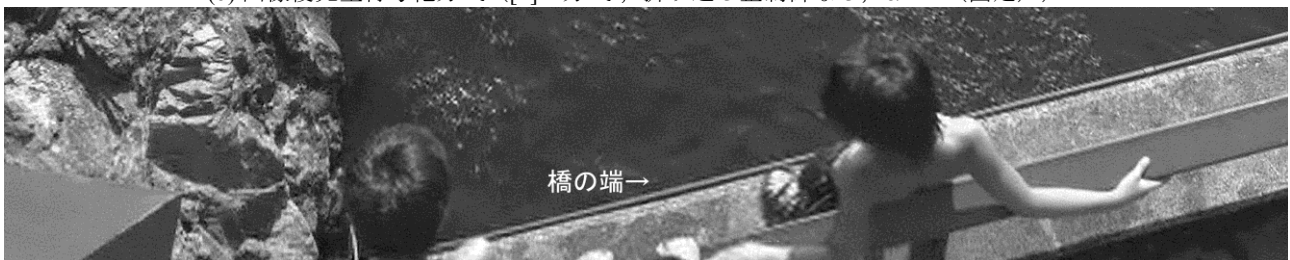
(b) 画像復元型符号化方式 ([2]の方式; 折り返し歪制御なし, $\alpha=2$ (固定))(c) 画像復元型符号化方式 (提案方式; 折り返し歪制御あり, $\alpha=1$ (静領域) ~2 (動領域))

図 9. 実験結果

5. むすび

従来の映像符号化の前後に解像度を変換する処理を行うことで画質改善を図る画像復元型符号化に関し、前処理での画像縮小時の折り返し歪を動静に応じて制御することを提案した。実験の結果、折り返し歪を固定的に付加した場合に問題となる静領域のジャギーが改善された。

参考文献

[1] Y. Shishikui, Y. Fujita, and K. Kubota: “Super Hi-Vision –

the star of the show!” EBU Technical Review, pp. 4–16 (Jan. 2009).

[2] T. Misu, Y. Matsuo, S. Sakaida, Y. Shishikui, and E. Nakasu: “Novel video coding paradigm with reduction/restoration processes,” PCS 2010, S2-2, pp.466–469 (2010).

[3] 松尾康孝, 三須俊枝, 境田慎一, 鹿喰善明: “ウェブレット超解像処理を用いた超高精細映像符号化の検討,” 2010年映情学年次大, 14-10 (2010).

[4] T. Misu, Y. Matsuo, S. Sakaida, and Y. Shishikui: “Novel framework for single/multi-frame super-resolution using sequential Monte Carlo method,” ACM Multimedia 2010, pp.771–774 (2010).



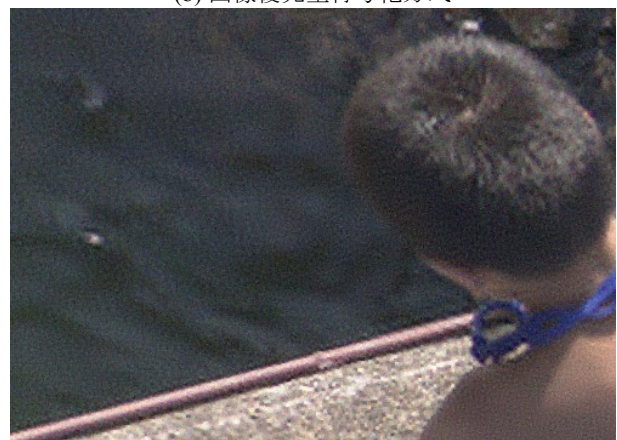
(a) 従来方式 MPEG-4 AVC/H.264



(b) 画像復元型符号化方式



(c) 従来方式 MPEG-4 AVC/H.264



(d) 画像復元型符号化方式



(e) 従来方式 MPEG-4 AVC/H.264



(f) 画像復元型符号化方式

図 10. 実験結果