

空間階調削減と復元を用いた 超高精細映像の超解像復元型映像符号化の検討

Video Coding of 8K Ultra High Definition Television using Image Size and Bit Depth Reductions and Their Restorations

松尾 康孝†
Yasutaka Matsuo

三須 俊枝†
Toshie Misu

岩村 俊輔†
Shunsuke Iwamura

境田 慎一†
Shinichi Sakaida

1. まえがき

将来のテレビジョンとして、超高精細映像である 8K・スーパーハイビジョン (SHV) [1, 2] が研究開発されている。この SHV 映像は、水平 7,680×垂直 4,320 画素 (8K) の画素数、最大 120 フレーム/秒のフレームレート、最大 12 ビット/色の階調数、従来のハイビジョン色域を超える ITU-R BT.2020[3] で規定された色域を持つ。この膨大な情報量を持つ SHV 映像の高圧縮符号化を目的として、我々はこれまでに「超解像復元型映像符号化」という新しい符号化の枠組を提案してきた[4, 5]。この超解像復元型映像符号化は、以下のような特徴を持つ。

- ・高圧縮符号化時の符号化劣化軽減のため、符号化/復号の前段 (符号化側) で画像削減、後段 (復号側) で超解像復元を行う。
- ・復号側の超解像復元を高精度かつ高安定に実施するため、符号化側にて超解像復元パラメータを最適化して、復号側へサイド情報として伝送する。

この超解像復元型映像符号化の枠組において、我々はこれまでに空間[4]、時間、階調[5]、色方向の処理を用いた方法を検討してきた。本論文では、このうち空間縮小と階調削減 (空間階調削減) および空間超解像復元と階調復元 (空間階調復元) を用いた新しい映像符号化法を提案する。

本論文の構成を以下に示す。2 章では、空間階調削減と復元の概要を示す。3 章では、2 章を基にした提案法を示す。4 章では、提案法の実験を行い、提案法は従来法に比べて高い客観画質と主観画質が得られることを示す。5 章では、本論文をまとめる。

2. 空間階調削減と復元の概要

2.1 節では、空間階調削減における空間縮小と階調削減の従来法を示し、これを基にした空間階調削減の概要を説明する。この空間階調削減では雑音除去の工夫を導入するため、2.1.1 節では、SHV 映像の雑音周波数解析を行い、その特性を明らかにする。2.2 節では、空間階調復元における空間超解像復元と階調復元の従来法を示し、これを基にした空間階調復元の概要を説明する。

2.1 空間階調削減の概要

空間縮小の従来法としては、原画像に bi-cubic, Lanczos, wavelet などの任意係数を持つ低域通過型線形フィルタを適用後、画素間引する方法がある。空間縮小では、基本的

に急峻な遮断特性を持つ低域通過フィルタを用いることで、良好な空間縮小画像が得られる。しかし後段の空間超解像復元では、空間縮小時の低域通過フィルタの遮断特性により空間超解像画質を変化させるため、本論文では wavelet フィルタを可変とする空間縮小法を用いる。

階調削減の従来法としては、原画像の各画素値を除算して、剰余を切り捨てまたは四捨五入する方法がある。他にも原画像を Lloyd-Max 量子化する方法がある。階調削減では、Lloyd-Max 量子化を用いることで原画像との差分が最小となる階調削減画像が得られるため、本論文では Lloyd-Max 量子化を用いた階調削減法を用いる。ここで、通常の Lloyd-Max 量子化では、雑音除去は考慮されない。しかし SHV 画像は、撮像素子における 1 画素あたりの受光面積が小さく、雑音レベルが大きくなりがちである。そこで本論文では、SHV 映像の符号化効率と空間超解像復元効率を高めるために、Lloyd-Max 量子化に雑音除去の工夫を導入する。このため 2.1.1 節では、SHV 映像の雑音周波数解析を行い、その特性を明らかにする。

2.1.1 雑音周波数解析

Figure 1 に、解析に用いる SHV 映像 (雑音が重畳された DC 値の低い画像と高い画像) を示す。解析では、128 点の 1 次元離散フーリエ変換 (Hamming 窓) を 64 点ずつずらしながら水平、垂直方向に適用したスペクトルの平均値 (直流成分を除く) を算出する。

Figure 2 に解析結果を示す。結果より SHV 映像は、直流値が低い場合と高い場合のいずれにおいても、ほぼ白色雑音を持つ。

2.2 空間階調復元の概要

空間超解像復元の従来法としては、単一フレーム画像に wavelet などの任意係数を持つ線形フィルタを適用する方法[6, 7]、total variation interpolation[8]などの非線形フィルタを適用する方法、または複数フレーム間のレジストレーションを用いる方法[9]などがある。ここで線形・非線形フィルタを用いる場合は、基本的に原画像のフィルタ処理により標準化周波数を超える空間高周波成分を生成するため、真の空間高周波成分を生成できない。一方、複数フレーム画像間のレジストレーションを用いる場合は、位置合わせを高精度に行えば、動領域において真の空間高周波成分を生成可能である。しかし SHV 映像はフレーム間動き量が大きい場合が多く[10]、処理量が膨大かつ動きばやけにより位置合わせが困難となるため、本論文では単一フレーム画像に wavelet フィルタを適用する空間超解像法を用いる。

† 日本放送協会 放送技術研究所

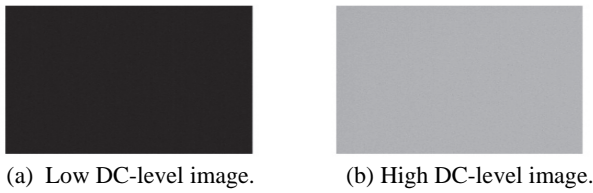
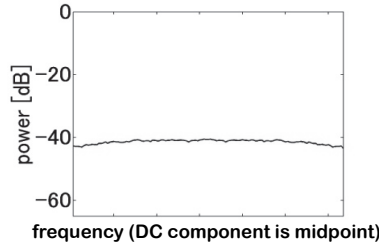
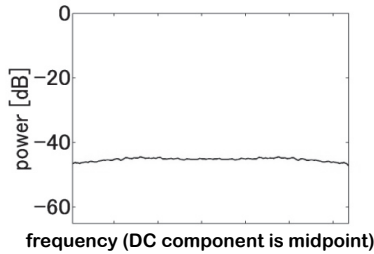


Figure 1: Test sequences for noise frequency analysis.



(a) Power spectrum of the low DC-level image in Figure 1 (a) without DC component.



(b) Power spectrum of the high DC-level image in Figure 1 (b) without DC component.

Figure 2: Averaged value of power spectrum of 1-dimensional discrete Fourier transform (1D-DFT) of the horizontal, and vertical directions. 1D-DFT is calculated by 128 DFT points, and DC component is midpoint.

階調復元の従来法としては、階調削減に Lloyd-Max 量子化を用いる場合は、その逆量子化テーブルを用いる方法がある。この階調復元では、基本的に階調復元画像の階調値ヒストグラムは櫛歯状となるため、何らかの階調補間処理が必要となる。また超解像復元パラメータの最適化を考慮すると、空間超解像復元と階調復元は一元的に処理することが望ましいため、本論文では wavelet フィルタによる空間超解像復元時に階調復元も行う方法を用いる。

3. 提案法

3章では、2章を基にした提案法を説明する。Figure 3に提案法の全体図を示す。以下に提案法の特徴を記載する。

- ・ 高圧縮符号化時の符号化劣化軽減のため、符号化/復号の前段(符号化側)で空間階調削減、後段(復号側)で空間階調復元を行う。
- ・ 符号化効率と超解像効率向上のため、空間階調削減では原画像を雑音除去しながら階調削減後、空間縮小する。階調削減は Lloyd-Max 量子化を用いて、かつ雑音除去の工夫を導入する。
- ・ 復号側の空間階調復元を原画像に忠実に実施するため、符号化側で原画像を基準に空間階調削減と復元パラメータを最適化して、復号側へサイド情報として伝送する。

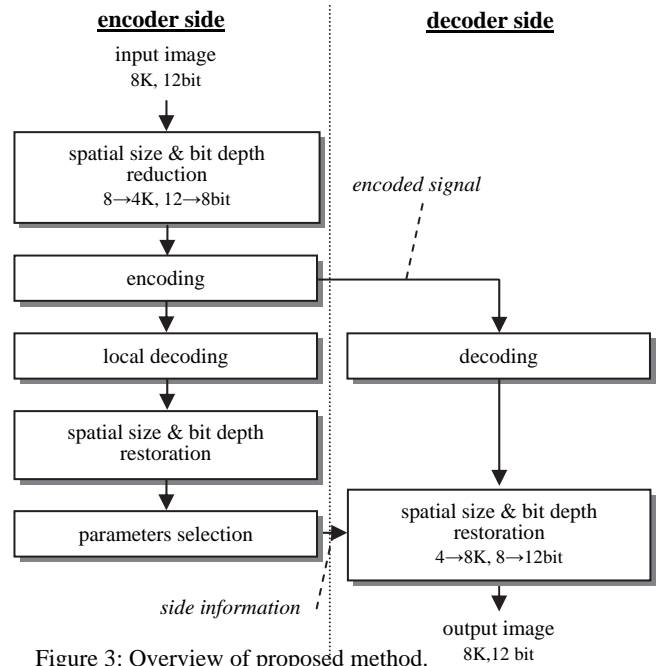


Figure 3: Overview of proposed method.

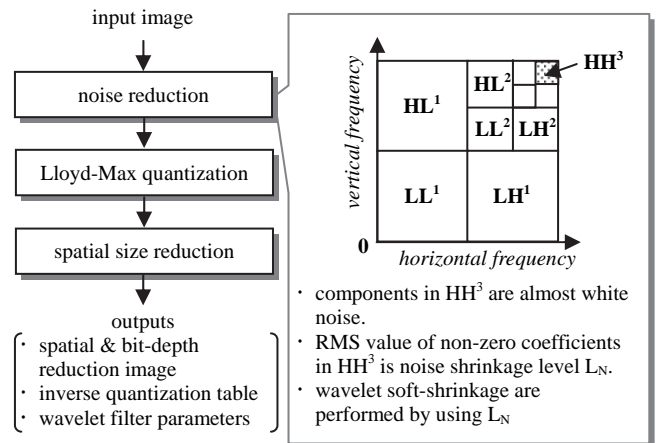


Figure 4: details of spatial size and bit depth reductions.

提案法の符号化側では、はじめに原画像を空間階調削減して符号化する。続いて復号側の空間階調復元を原画像に忠実に行うために、符号化信号を局部復号して空間階調復元後、原画像と空間階調復元画像の差分が最小となるような空間階調削減と復元パラメータを選択する。続いて選択した空間階調削減パラメータを用いて空間階調削減して符号化した符号化信号を復号側へ伝送する。さらに選択した空間階調復元パラメータをサイド情報として復号側へ伝送する。

復号側では、符号化信号を復号後、サイド情報を用いて符号化側と同様の最適な空間階調復元を行う。

以下、3.1節では空間階調削減、3.2節では空間階調復元、3.3節では最適パラメータ選択の詳細を説明する。

3.1 空間階調削減

Figure 4 に空間階調削減の詳細図を示す。空間階調削減では、符号化効率と超解像効率の向上のため、原画像の雑音除去をしながら階調削減後、空間縮小する。

階調削減では、はじめに原画像に wavelet soft-shrinkage [11] を適用して雑音除去画像を生成する。ここで [11] においては、原画像の輝度レベル分散値を雑音 shrinkage 閾値とする。一方、提案法では 2 章の結果より、白色雑音が重畳された SHV 映像の雑音除去をより高精度に行うため、原画像を空間高周波方向に n_H 階 wavelet 分解した空間斜め周波帯域 (HH^{n_H}) 成分はほぼ水平、垂直方向への相関が低い白色雑音であるとして、この帯域の非零成分の root mean square 値を雑音 shrinkage 閾値とする。なお提案法の wavelet 分解および再構成は、雑音レベル検出や超解像処理に用いるため、画像パワーの積分が周波数パワーの積分と等しくなるようにノルムの正規化を行う。続いて雑音除去画像を training set として Lloyd-Max 量子化を行い、階調削減画像とその逆量子化テーブルを出力する。

空間縮小は、階調削減画像を空間 1 階 wavelet 分解した空間低周波帯域成分を整数化した空間階調削減画像と、用いた wavelet フィルタパラメータ F_w を出力する。

3.2 空間階調復元

空間階調復元では、符号化画像を復号後、逆量子化テーブルを用いて逆量子化する。このとき逆量子化画像の輝度値ヒストグラムは楕歯状となる。続いて逆量子化画像を空間階調復元する。この際、符号化側では原画像と空間階調復元画像の差分が最小となるように空間階調削減と復元のパラメータを選択して、復号側へサイド情報として伝送する。また復号側ではサイド情報を用いて、符号化側と同様の方法で最適な空間階調復元を行う。

逆量子化画像の空間階調復元では、wavelet 超解像と階調補間を同時に行う。wavelet 超解像では、はじめに逆量子化画像をデシメーション無しで空間 1 階 wavelet 分解して、空間低周波帯域 (LL)、垂直、水平、斜め高周波帯域 (LH, HL, HH) 成分を抽出する。次に wavelet 超解像の波形整形と階調補間のため、LH, HL, HH 内成分を $B_s \times B_s$ ブロック領域毎に平滑化処理する。平滑化処理では、画像のエッジ領域以外は階調が滑らかに変化すると仮定して、分散値 σ の bilateral フィルタを用いる。そして空間低周波帯域成分を逆量子化画像、水平、垂直、斜め高周波帯域成分を bilateral フィルタ適用後の LH, HL, HH 成分として空間 1 階 wavelet 再構成して、空間階調復元画像を得る。

3.3 最適パラメータ選択

最適パラメータ選択では、空間縮小で用いる wavelet フィルタパラメータ F_w と、空間階調復元で用いる LH, HL, HH 内の $B_s \times B_s$ ブロック領域毎の bilateral フィルタの分散値 σ について、原画像と空間階調復元画像の差分が最小となる値を選択する。

4. 実験

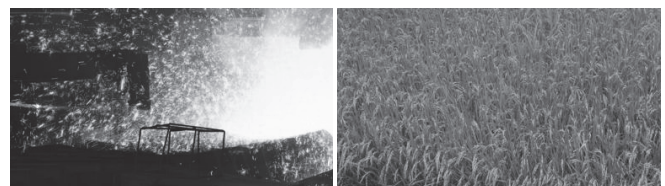
3 章の提案法の評価実験を行い、提案法の有効性を示す。Table 1 に実験で用いる 4 種類の符号化法を示す。また Figure 5 に実験で用いる 2 種類の原画像を示す。実験で用いる原画像は、8K の SHV 画像を $3,840 \times 2,160$ 画素に切り出して使用する。この原画像は、60 フレーム/秒、12 bit YUV4:0:0 信号、16 フレーム長の製鉄所で火花が飛び散る映像 (Spark) と稲穂が揺れる映像 (Rice field) である。提案法では原画像を $1,920 \times 1,080$ 画素、8bit 深度に空間階調削減する。符号化/復号は、AVC/H.264 (JM18.0 [12]) を用いる。1GOP は 16、B スライス は 1 フレーム毎に挿入する。そして予備実験により、 F_w は直交性と線形位相性を満たす Haar と線形位相性を満たす Cohen-Daubechies-Feauveau(CDF) 5/3 および 9/7、視覚の周波数チャンネル理論 [13] より n_H は 3、bilateral フィルタタップ数は 5、 σ は 0.1, 2.1, 4.1, 6.1, B_s は 64 とした。なおサイド情報はフレーム毎に伝送し、ZIP 可逆圧縮を行うことで約 70[kbps] の情報量を総符号量に含める。

Figure 6 に原画像を基準とした PSNR の 16 フレーム平均値を示す。提案法 (Proposal) は AVC/H.264 に対して Spark で約 1.7[dB]、Rice field で約 1.1[dB] の改善が見られた。なお提案法で σ と F_w のパラメータを σ は 1.0, F_w は CDF5/3 に固定した場合 (Proposal w/o PO) は Spark で約 1.3[dB]、Rice field で約 0.8[dB]、提案法で雑音除去とパラメータ最適化を行わない場合 (Proposal w/o NR and PO) は Spark で約 1.2[dB]、Rice field で約 0.7[dB] の改善が得られた。結果より Spark の方が Rice field より AVC/H.264 に対する提案法の PSNR 改善効果が高いが、これは Spark の方がヒストグラムの偏りや雑音レベルが大きい画像であり、雑音除去を考慮した Lloyd-Max の効果が高いためと考えられる。

Figure 7 に Proposal と AVC/H.264 の 20Mbps での符号化画像を、輝度値を 4bit 右シフトして 8bit 深度画像として示す。同図、10bit 深度の有機 EL (SONY BVM-F250) マスタモニタによる目視、および結果データの数値解析により、提案法はいずれの画像でも AVC/H.264 に比べてぼやけやブロック歪などが少ないことを確認した。

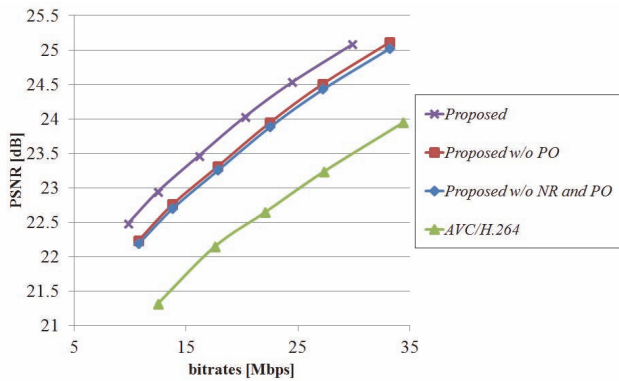
Table 1: Experimental methods.

AVC/H.264	Encoding images with 8K, 12bit
Proposal w/o NR & PO	Proposed method without noise reduction (NR) and parameters optimization (PO)
Proposal w/o PO	Proposed method without parameters optimization (PO)
Proposal	Proposed method in Section 3

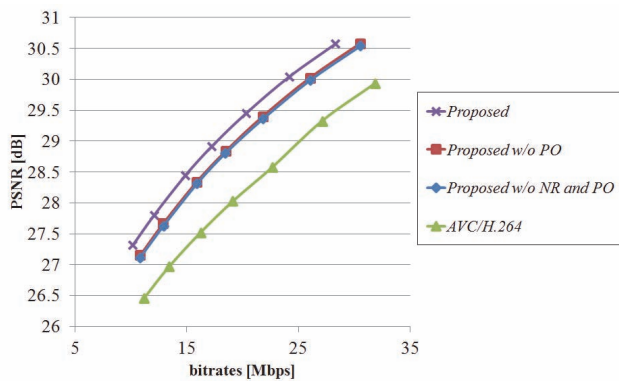


(a) Spark (b) Rice field

Figure 5: Test sequences for experiments.

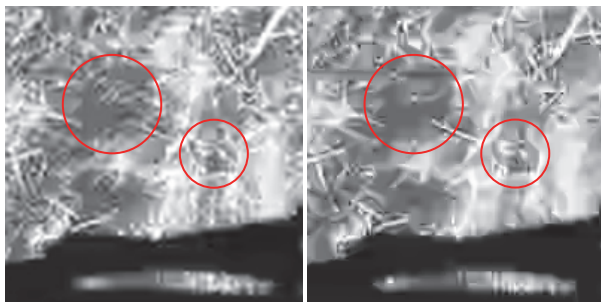


(a) Spark



(b) Rice field

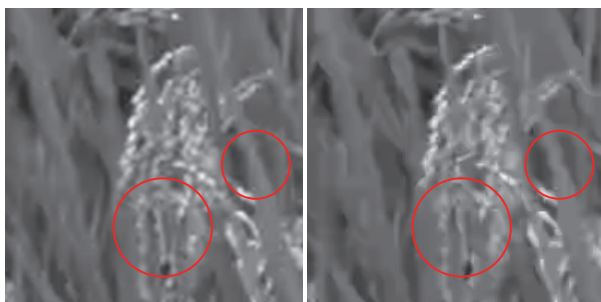
Figure 6: PSNR results.



(i) Proposal

(ii) AVC/H.264

(a) Spark



(i) Proposal

(ii) AVC/H.264

(b) Rice field

Figure 7: Encoded images for Proposal and AVC/H.264 methods with 20Mbps.

5. まとめ

本稿では、AVC/H.264 符号化/復号の前段で空間階調削減、後段で空間階調復元を用いる超解像復元型映像符号化を提案した。符号化効率と超解像効率向上のため、空間階調削減では原画像を階調削減後、空間縮小した。階調削減では、Lloyd-Max 量子化において原画像の雑音除去画像を training set として用いることで、雑音除去を考慮した階調削減を行った。そして雑音除去は、原画像を wavelet 分解して雑音レベル検出を行い、これを雑音 shrinkage 閾値として wavelet soft-shrinkage にて行った。空間縮小は、原画像を wavelet 分解して、その空間低周波成分を整数化して行った。原画像に忠実な空間階調復元を行うため、空間階調復元では符号化側で原画像を基準に最適なパラメータを選択して、復号側にサイド情報として伝送した。実験より、提案法は従来法に比べ高い客観画質と主観画質が得られることを確認した。

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] M. Sugawara, and K. Masaoka, "UHDTV Image Format for Better Visual Experience," Proceedings of the IEEE, vol. 101, no. 1, pp. 8-17, Jan. 2013.
- [3] Rec. ITU-R BT.2020, "Parameter values for ultra-high definition television systems for production and international programme exchange," 2012.
- [4] Y. Matsuo, T. Misu, S. Sakaida, and Y. Shishikui, "Video Coding with Wavelet Image Size Reduction and Wavelet Super Resolution Reconstruction," Proceedings of IEEE ICIP, MP.PG.6, pp. 1181-1184, 2011.
- [5] Y. Matsuo, T. Misu, S. Iwamura, and S. Sakaida, "Ultra High-definition Video Coding using Bit-depth Reduction with Noise Reduction and Pseudo Contour Prevention," Proceedings of IEEE VCIP, FP 489, Nov. 2013.
- [6] M. Nguyen, and P. Milanfar, "An Efficient Wavelet-based Algorithm for Image Superresolution," Proceedings of IEEE ICIP, vol. 2, pp. 351-354, Sep. 2000.
- [7] T. Temizel, and T. Vlachos, "Image Resolution Upscaling in the Wavelet Domain using Directional Cycle Spanning," Journal of Electronics Imaging, vol. 14, no. 4, 2005.
- [8] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Total Variation Super Resolution using a Variational Approach," Proceedings of IEEE ICIP, pp. 641-644, Oct. 2008.
- [9] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution Image Reconstruction: a Technical Overview," IEEE Signal Processing Magazine, vol. 20, no. 3, pp. 21-36, 2003.
- [10] 松尾, 三須, 境田, 鹿喰, "超高精細映像の高確度レジストレーション手法の一検討," 電子情報通信学会誌, vol. J93-D, no. 9, pp. 1675-1678, 2010.
- [11] D. L. Donoho, "De-noising by soft-thresholding," IEEE Transactions on Information Theory, vol. 41, no. 3, pp. 613-627, May. 1995.
- [12] H.264/AVC JM reference software, "http://iphome.hhi.de/suehring/tml/"
- [13] K. T. Mullen, "The Contrast Sensitivity of Human Vision to Red Green and Blue-yellow Chromatic Gratings," Journal of Physiology, vol. 359, pp. 381-409, Feb. 1985.