

# ウェーブレット多重解像度成分のレジストレーションによる デジタルシネマから超高精細映像への空間超解像法の検討

## Image Super-resolution from Digital Cinema to Ultrahigh Definition Video using Registration of Wavelet Multi-scale Components

松尾 康孝<sup>†‡</sup>岩崎 真也<sup>†</sup>山村 勇太<sup>†</sup>甲藤 二郎<sup>†</sup>Yasutaka Matsuo<sup>†‡</sup>Shinya Iwasaki<sup>†</sup>Yuta Yamamura<sup>†</sup>Jiro Katto<sup>†</sup>

### 1. まえがき

近年、映画では、最大 4K の水平解像度を持つ高精細映像としてデジタルシネマ[1], [2]が実用化されている。またテレビジョンでは、8K の水平解像度を持つ超高精細映像としてスーパーハイビジョン[3], [4]が研究されている。ここでデジタルシネマからスーパーハイビジョンへの映像方式変換は、映画素材をスーパーハイビジョンで放送する場合などに必要である。このため我々は、これらの映像方式変換を目的とした研究を行う。

研究では、空間解像度、フレームレート、ビット深度、色などの変換の検討が必要である。このうち本論文では、空間解像度の超解像法を検討する。検討では、デジタルシネマのような高精細映像に適した超解像法の開発が課題となる。課題では、デジタルシネマ映像が持つフィルムの粒状性雑音や固体撮像素子の熱雑音が問題となる。問題となる雑音は、一般に超解像に悪影響を及ぼす。しかし映像表現として重要であるため[5]、単純に除去すれば良いものではない。

超解像の従来法としては、線形フィルタ処理である wavelet 超解像法[6]-[8]、非線形フィルタ処理である total variation 超解像法[9]、複数フレーム間のレジストレーションを用いた超解像法[10]などが知られている。このうち wavelet や total variation 超解像法は、フィルタ処理により空間高周波成分を推定して生成するため、必ずしも真の空間高周波成分を生成できないという課題がある。一方レジストレーションを用いた超解像法は、撮像位置が異なるオブジェクト間の位置合わせを行って画素密度を高めるため、レジストレーション精度と確度が高い領域では真に近い空間高周波成分を生成できる。しかし雑音等の影響により、レジストレーション精度や確度が低い領域では超解像画質が大きく劣化するという課題がある。また従来のデジタルシネマ映像は、4K の水平解像度に対してフレームレートが 24 フレーム/秒と低いため、動きぼやけ等によりフレーム間の画像相関が低く、複数フレーム間のレジストレーションが難しいという課題もある。

以上より本論文では、はじめにデジタルシネマ映像が持つ雑音成分の周波数解析を行う。そして解析結果を踏まえた超解像法を提案する。提案法の特徴を以下に示す。

(1) 原画像から信号成分と雑音成分を分離して、各々に適した超解像を行った後、合成する。

(2) 信号成分の超解像では、高精細映像の同一フレーム内での自己合同性や相似性の高さを利用して、信号成分とその多重解像度分解成分間でのレジストレーションにより空間高周波成分を生成する。

(3) 超解像した信号成分は、Double Interpolation PSNR (以下 DI PSNR) [11], [12]の考え方を基に、その縮小成分と元の信号成分の差分が最小となるように最適化する。

本論文の構成は以下のとおりである。2 章では、諸条件で撮影された高精細映像が持つ雑音成分の周波数解析結果を示す。3 章では、2 章の解析結果を踏まえて、信号と雑音成分を分離して各々最適な超解像を行うことを特徴とする方法を提案して、その詳細を説明する。4 章では、提案法のシミュレーション実験を行う。最後に 5 章では、本論文をまとめる。

### 2. 雑音成分の周波数解析

図 1 (a)-(f)に解析に用いる高精細映像を示す。同図(a), (b)は 4K の水平解像度を持つ CMOS 素子で撮像された画像である。同図(c), (d)はフィルムで撮像された画像を 4K の水平解像度でデジタル化した画像である。同図(e), (f)は高感度フィルムで撮像された画像を 4K の水平解像度でデジタル化した画像である。

解析では、はじめに図 1 (a)-(f)において信号成分がほぼ平坦な信号成分直流領域を選択する。そして信号成分直流領域を 1 次元離散フーリエ変換 (以下 1D DFT) で周波数解析する。なお 1D DFT では、hamming 窓を用いて 64 タップでフィルタ処理を行い、32 画素ずつシフトして水平、垂直方向に解析した結果の平均値を出力する。

図 2 に解析結果を示す。同図(a), (b)より CMOS 素子撮像画像の雑音成分は、ほぼ白色雑音である。これは、CMOS 撮像素子の画素毎のアンブにおける熱雑音がほぼ白色雑音であるためだと考察される。また同図(c), (d)よりフィルムで撮影された画像の雑音成分も、ほぼ白色雑音である。これはフィルムの粒状性雑音はランダムであり、ほぼ白色雑音に近いめだと考察される。一方、同図(e), (f)より高感度フィルムで撮影された画像の雑音成分は、白色雑音ではない。これは高感度フィルムの粒状性が粗いため、空間高周波成分が少ないことが原因と考察される。

<sup>†</sup> 早稲田大学大学院 基幹理工学研究科,  
Department of Computer Science, Waseda University

<sup>‡</sup> NHK 放送技術研究所,  
NHK Science & Technology Research Laboratories



図1 実験画像

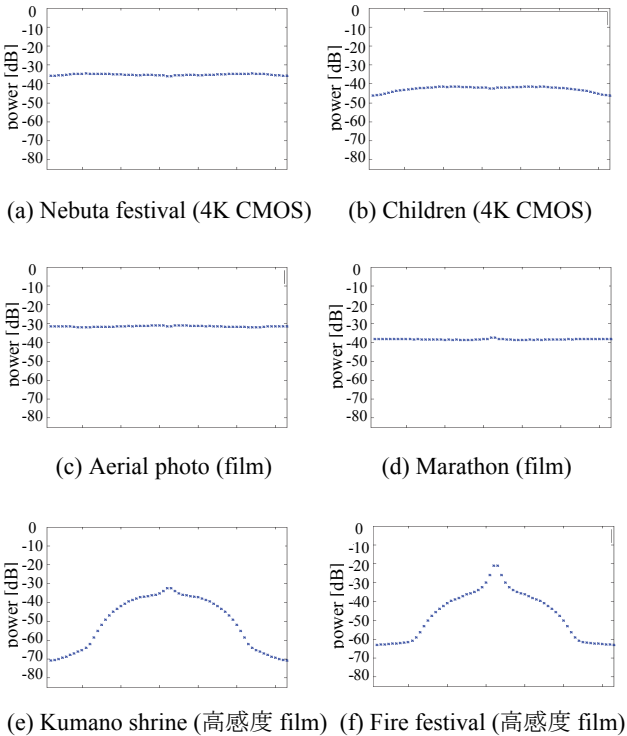


図2 周波数スペクトル (信号成分直流領域)  
(横軸は中央を直流成分とする周波数)

一般に雑音成分は、信号成分の超解像に悪影響を及ぼす。特に一般画像は空間低周波帯域に多くのパワーを持つため、空間高周波帯域では信号対雑音比が低く、雑音の影響を受けやすい。図 2(e), (f)の雑音成分は白色雑音ではないため、空間高周波帯域の雑音レベルが低く、従来超解像法でも雑音の影響を受けにくい。一方、同図(a)-(d)の雑音成分はほぼ白色雑音であるため、空間高周波帯域の雑音レベルが高く、従来超解像法では超解像画質が劣化しやすい。

以上を踏まえて、我々は原画像を信号成分と雑音成分に分離することを特徴とする超解像法を提案する。

### 3. 提案法

2章の雑音成分の周波数解析結果を踏まえて、原画像を信号成分と雑音成分に分離して各々適切な超解像を行い、合成する手法を提案する。図3に提案法の全体ブロック図を示す。以下では、同図を基に詳細を説明する。なお説明における超解像倍率は、水平、垂直各2倍とする。

#### 3.1. 信号雑音分離

信号雑音分離ブロックでは、原画像  $O$  から信号成分  $S$  と雑音成分  $N$  を分離する。提案法では雑音成分  $N$  として、図 2(a)-(d)のように空間高周波帯域の雑音レベルが高く超解像画質が劣化しやすい白色雑音を分離する。ここで白色雑音は無相関性が高いため時空間相関が低い。そこで時空間 wavelet 分解を用いて、時空間相関が低い時空間高周波帯域の孤立点要素を抽出することで白色雑音を分離する方法[13]を用いる。

図4に信号雑音分離の詳細ブロック図を示す。以下では、同図を基に詳細を説明する。

雑音位相抽出ブロックでは、白色雑音成分の位相位置を検出する。検出では、白色雑音成分は時空間相関が低いため、時空間高周波帯域の孤立点要素位置を調べる。処理では、はじめに原画像  $O$  をデシメーション無しで時間方向に1階かつ空間方向に1階 (以下(1, 1)階) wavelet 分解する。

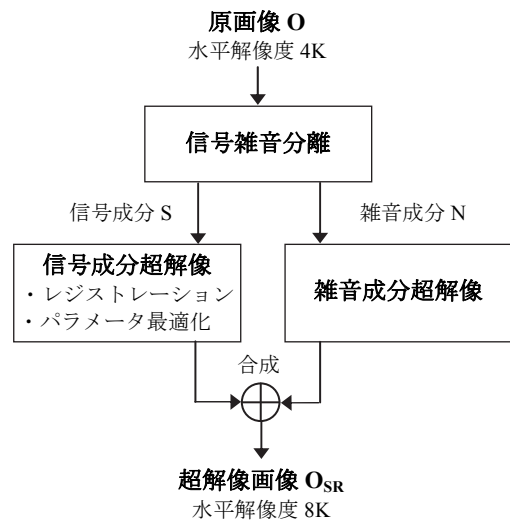


図3 全体ブロック図

次にその(1, 1)階の時空間高周波帯域  $H^1HH^1$  内の全要素の絶対値を 2 値化して  $bH^1HH^1$  とする. その閾値は  $H^1HH^1$  内の非零要素絶対値の中央値とする. 最後に  $bH^1HH^1$  内の孤立点要素だけ 1 を残した雑音位相情報  $P$  を出力する. 孤立点要素は,  $bH^1HH^1$  内の 1 の要素を中心とした  $3 \times 3$  領域の合計値 (1 から 9 の値をとる) が閾値  $Th$  以下 (例えば  $Th = 1$ ) の要素とする.

雑音レベル抽出ブロックでは, 原画像  $O$  をデシメーション無しで時間方向に  $M_t$  階かつ空間方向に  $M_s$  階 (以下  $(M_t, M_s)$  階) wavelet 分解する. そしてその任意の  $(m_t, m_s)$  階の帯域内で  $P=1$  の位相位置にある全要素の絶対値の中央

値を雑音レベル  $LV_N (*^{m_t} *^{m_s})$  として出力する. 例えば  $H^1HH^1$  の雑音レベル  $LV_N (H^1HH^1)$  は, 中央値演算関数を  $median$  とすると,  $median(H^1HH^1 \cap P)$  で計算される.

信号雑音成分分離ブロックでは, はじめに原画像  $O$  をデシメーション無しで  $(M_t, M_s)$  階 wavelet 分解する. 次に, 時空間低周波帯域  $L^{M_t}LL^{M_s}$  以外の全帯域から雑音成分を抽出する. 抽出は, 任意帯域  $*^{m_t} *^{m_s}$  の  $P=1$  の位相位置における要素の絶対値と, その帯域の雑音レベル  $LV_N (*^{m_t} *^{m_s})$  のうち, 大きい方の値を抽出して任意帯域の雑音要素  $*^{m_t} *^{m_s}_N$  とする. またその残りを信号要素  $*^{m_t} *^{m_s}_S$  とする. 最後に  $*^{m_t} *^{m_s}_S$  と  $*^{m_t} *^{m_s}_N$  を各々デシメーション無しで  $(M_t, M_s)$  階 wavelet 再構成して, 信号成分  $S$  と雑音成分  $N$  を出力する.

### 3.2. 信号成分超解像

高精細画像では, 空間解像度が高く同一フレーム内に多くのオブジェクトが含まれるため, 自己合同性や相似性の高さを利用した超解像[14]が有効である. 例えば, 向日葵やビル街などの映像では, 同一オブジェクトまたは大小様々な異なるオブジェクト内に, 自己合同性や相似性の高い箇所が存在する可能性が高い. そこで提案法では, 信号成分  $S$  を  $n$  階 wavelet 分解した空間低周波成分  $LL_R^n$  から  $S$  への位置合わせを行う. そして位置合わせされた箇所では空間高周波成分も似ていると仮定して,  $S$  を  $n$  階 wavelet 分解した空間高周波成分を  $S$  の標準化周波数を超える空間高周波帯域に割当て wavelet 再構成して, 信号成分超解像画像を生成する.

図 5 に信号成分超解像の詳細ブロック図を示す. 以下では, 同図を基に詳細を説明する.

空間 wavelet 分解ブロックでは, 信号成分  $S$  をデシメーション無しで空間方向に 1 階 wavelet 分解する. そして空間低周波成分  $LL_S^1$ , および水平, 垂直, 斜め高周波成分  $LH_S^1, HL_S^1, HH_S^1$  を抽出する.

レジストレーションブロックでは, はじめに初期値として  $LH_S^1, HL_S^1, HH_S^1$  を信号成分  $S$  の標準化周波数を超える水平, 垂直, 斜め高周波帯域の要素としてコピーする. 次に  $S$  を空間方向に  $n$  階 wavelet 分解する. そして  $n$  階空間低周波成分  $LL_R^n$  を多重解像度画像として抽出する. 次に  $LL_R^n$  を  $4 \times 4$  の小ブロックに分割して, その全ブロックにおいて  $LL_R^n$  を基準フレーム,  $S$  を参照フレームとして小数画素精度の全探索ブロックマッチングを行う. そしてその SSD (Sum of Square Difference) 値が最小となる位置合わせ情報を出力する. 次に全ブロックにおいて, 位置合わせ情報を用いて  $LH_R^n, HL_R^n, HH_R^n$  の当該ブロック位置の要素を  $LH_S^1, HL_S^1, HH_S^1$  に割り当てる. 次に割り当てた要素において点拡がり関数を用いた MAP(Maximum A Posteriori)再構成[15]を行う. 最後に MAP 再構成後の  $LH_S^1, HL_S^1, HH_S^1$  の波形整形のために, 分散値  $\sigma_{LH}, \sigma_{HL}, \sigma_{HH}$  のガウシアンによる bilateral フィルタ[16]を適用する.

空間 wavelet 再構成ブロックでは,  $S$  を空間低周波成分, bilateral フィルタ適用後の  $LH_S^1, HL_S^1, HH_S^1$  を空間高周波成分としてデシメーション無しで空間方向に 1 階 wavelet 再構成して, 超解像信号成分  $S_{SR}$  を出力する.

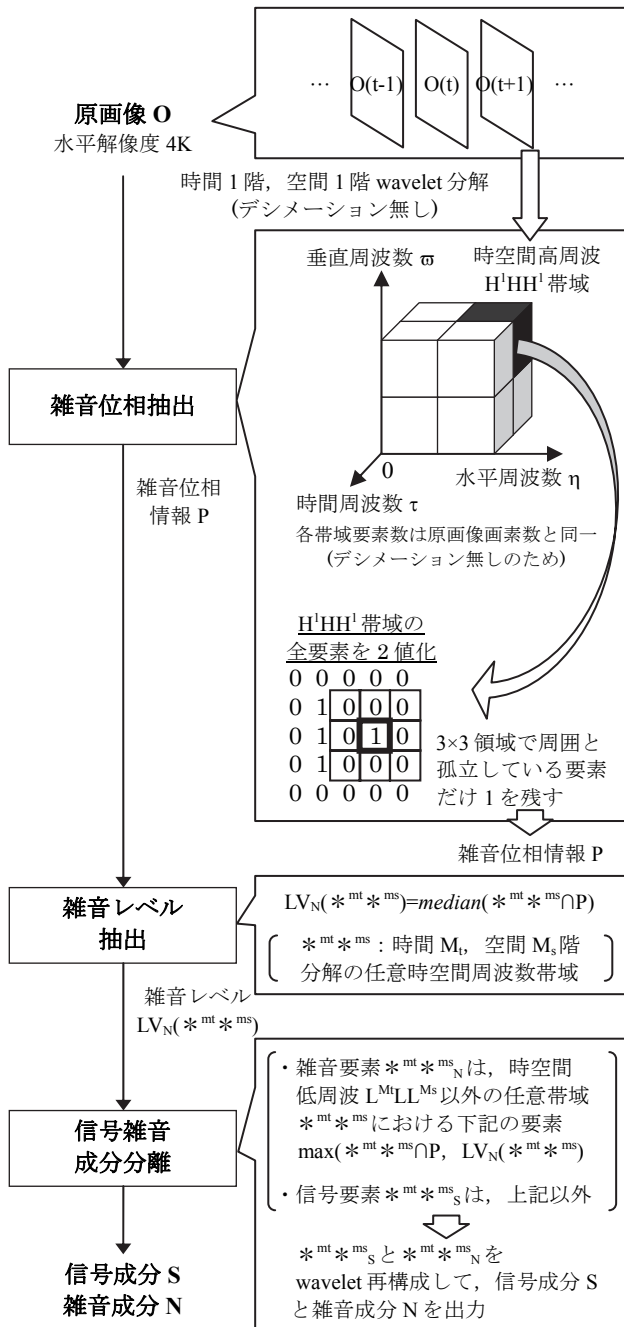


図 4 信号雑音分離の詳細ブロック図

最適化ブロックでは、DI PSNR 最大化による分散値 $\sigma_{LH}$ ,  $\sigma_{HL}$ ,  $\sigma_{HH}$ の最適化を行う。ここで DI PSNR は、もし画像縮小と超解像が理想的に行えるならば原画像とその超解像画像を水平、垂直、斜め方向に 1 画素ずつ位相位置をずらして縮小した 4 枚の画像の差分は最小になることより、原画像を基準とする 4 枚の縮小画像の PSNR (Peak Signal to Noise Ratio) 値を計算し、それらの和を平均して求める。ゆえに DI PSNR が大きいほど、超解像手法として有効であることが見込まれる。最適化では、はじめに  $S_{SR}$  をデシメ

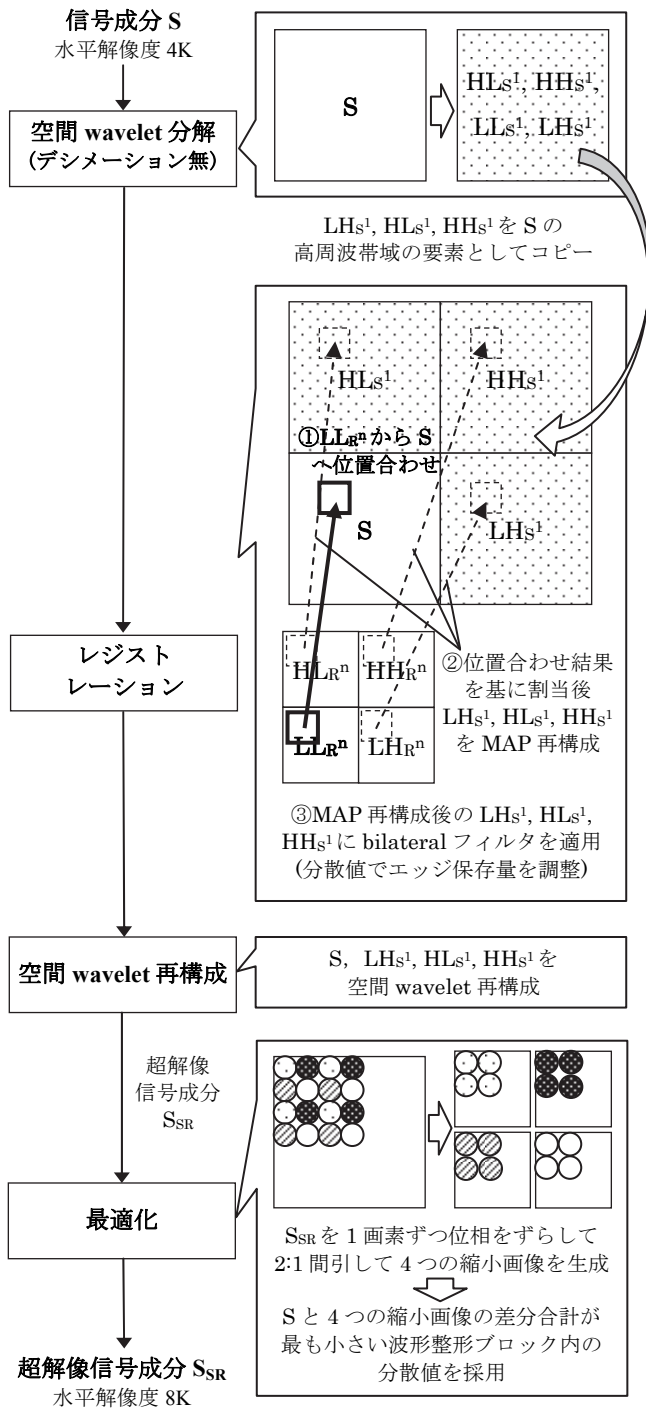


図 5 信号成分超解像の詳細ブロック図

ーション無しで空間方向 1 階 wavelet 分解する。次にその空間低周波成分を水平、垂直、斜め方向に 1 画素ずつ位相位置をずらして 2:1 画素間引きを行って 4 枚の縮小画像を生成し、各々  $S$  を基準とする PSNR 値を計算し、それらの和を平均して DI PSNR を計算する。以上の処理を波形整形ブロック内で $\sigma_{LH}$ ,  $\sigma_{HL}$ ,  $\sigma_{HH}$ を各々変化させて作成した  $S_{SR}$  で行い、DI PSNR が最大となる  $S_{SR}$  を出力する。

### 3.3. 雑音成分超解像と超解像信号成分への付加

雑音成分超解像では、白色雑音成分が時空間低相関であることを利用した超解像を行う。

図 6 に雑音成分超解像の詳細ブロック図を示す。以下では、同図を基に雑音成分超解像処理の詳細を説明する。

雑音高周波成分生成ブロックでは、白色雑音成分は時空間方向の相関が低く、空間低周波成分と空間高周波成分が同じ位相位置に存在しやすいことを利用して、雑音成分  $N$  を  $N$  を超える高周波帯域にコピーすることで各雑音要素のパワーを調整する。そしてこれを空間方向に 1 階 wavelet 再構成して、超解像雑音成分  $N_{SR}$  を出力する。

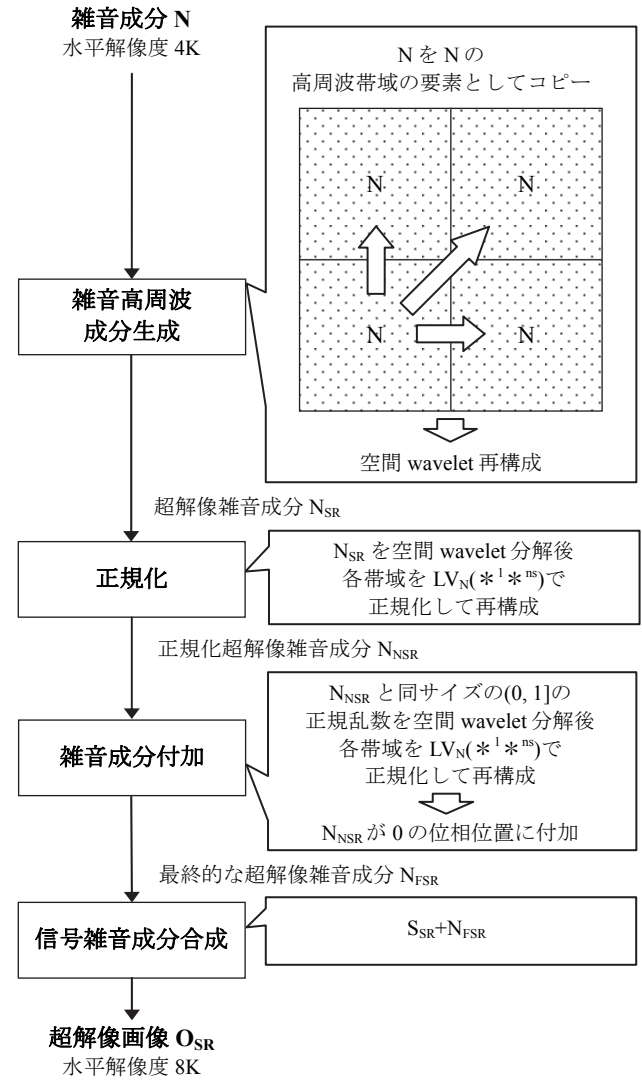


図 6 雑音成分超解像の詳細ブロック図

正規化ブロックでは、超解像雑音成分  $N_{SR}$  の各帯域全体のパワーを正規化する。正規化処理では、 $N_{SR}$  を空間方向に  $n_s$  階 wavelet 分解して各帯域成分を  $LV_N(*^1 *^{n_s})$  で正規化後、空間方向に  $n_s$  階 wavelet 再構成して正規化超解像雑音成分  $N_{NSR}$  を出力する。

雑音成分付加ブロックでは、白色雑音として  $N_{NSR}$  と同じ空間サイズの  $(0,1]$  の正規乱数を生成して、これを空間方向に  $n_s$  階 wavelet 分解して各帯域成分を  $LV_N(*^1 *^{n_s})$  で正規化後、空間方向に  $n_s$  階 wavelet 再構成して印加雑音成分を生成する。そしてこれを  $N_{NSR}$  の要素値が 0 の位相位置に加えることで  $N_{NSR}$  で再現できなかった位相位置の白色雑音要素を付加して、最終的な超解像雑音成分  $N_{FSR}$  を出力する。

最後に信号雑音成分合成ブロックでは、 $S_{SR} + N_{FSR}$  より超解像画像  $O_{SR}$  を生成する。

#### 4. 実験

表 1 に実験条件を示す。実験画像は、4K の水平解像度を持つ Nebuta festival (図 1(a)), Aerial photo (図 1(c)), Kumano shrine (図 1(e)) である。比較する方法は、提案法、単純 wavelet 拡大法 (原画像  $O$  に 3.2 節の信号成分超解像 (図 5) をレジストレーションブロックと最適化ブロック無しで適用, 波形整形ブロックの  $\sigma_{LH}$ ,  $\sigma_{HL}$ ,  $\sigma_{HH}$  は 1.0 を使用), total variation 拡大法 ([9]の方法), Lanczos-3 法, bi-cubic 法である。使用する wavelet は Haar, 時間方向の分解階数  $N_t$  と空間方向の分解階数  $N_s$  は、視覚の周波数チャンネル特性[17]より  $N_t=1$ ,  $N_s=4$  とした。波形整形ブロックで用いる bilateral フィルタのタップ長は 5, 分散値の範囲は実験的に 0.1~5.1 で 1 刻みとした。

図 7 に実験結果の切り出し画像を示す。同図より、提案法は高い主観画質を持つ。また図 5 の最適化ブロック内での処理により分散値を最適化することで、最適化しない場合に比べて DI PSNR が Nebuta で 1.4[dB], Aerial photo で 1.1[dB], Kumano shrine で 0.6[dB] 向上した。

図 8 に実験結果画像の 1D DFT による周波数スペクトルを示す。同図より、提案法は従来法よりも原画像に近い周波数特性が得られている。

表 1 実験条件

実験画像	<ul style="list-style-type: none"> <li>• Nebuta festival (図 1(a))</li> <li>• Aerial photo (図 1(c))</li> <li>• Kumano shrine (図 1(e))</li> </ul>
比較法	<ul style="list-style-type: none"> <li>• 提案法</li> <li>• 単純 wavelet 拡大法</li> <li>• total variation 拡大法</li> <li>• Lanczos-3 法</li> <li>• bi-cubic 法</li> </ul>
wavelet	<ul style="list-style-type: none"> <li>• Haar (時間分解数 <math>N_t=1</math>, 空間分解数 <math>N_s=4</math>)</li> </ul>
波形整形	<ul style="list-style-type: none"> <li>• bilateral フィルタ (タップ長 5, 分散値 0.1~5.1 で 1 刻み)</li> </ul>

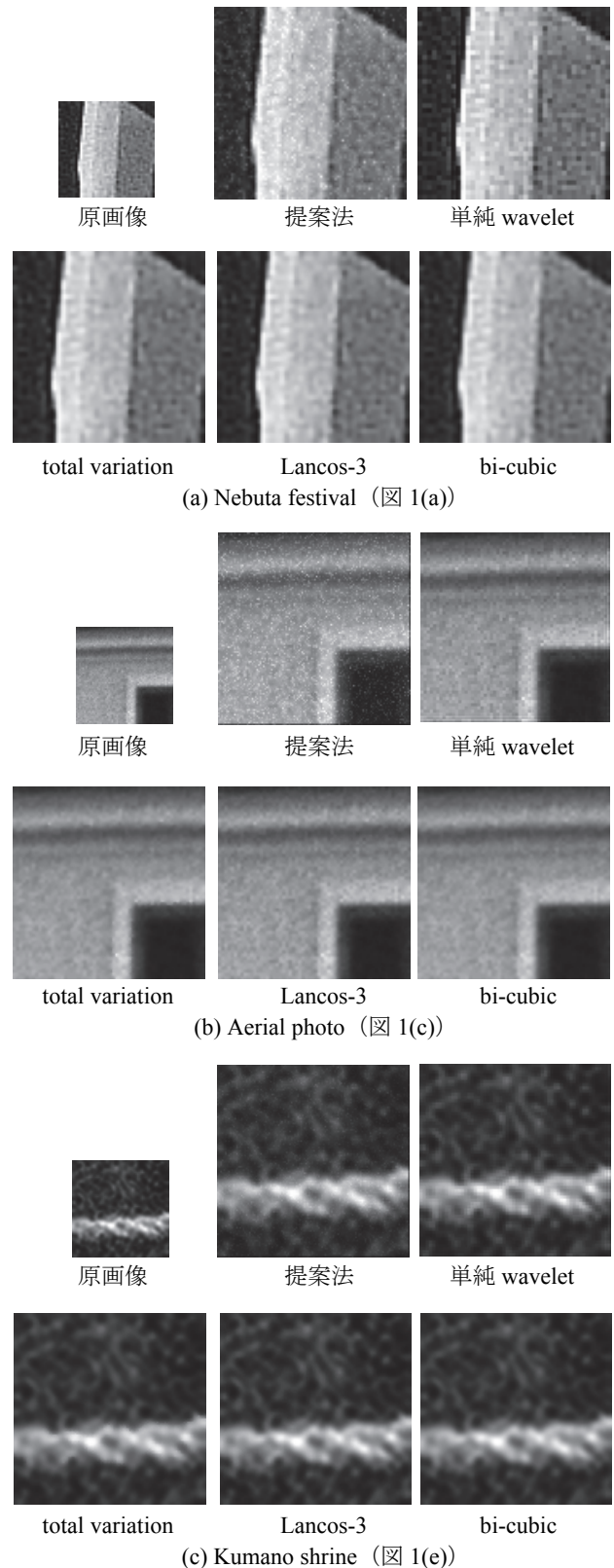


図 7 実験結果 (切り出し画像)

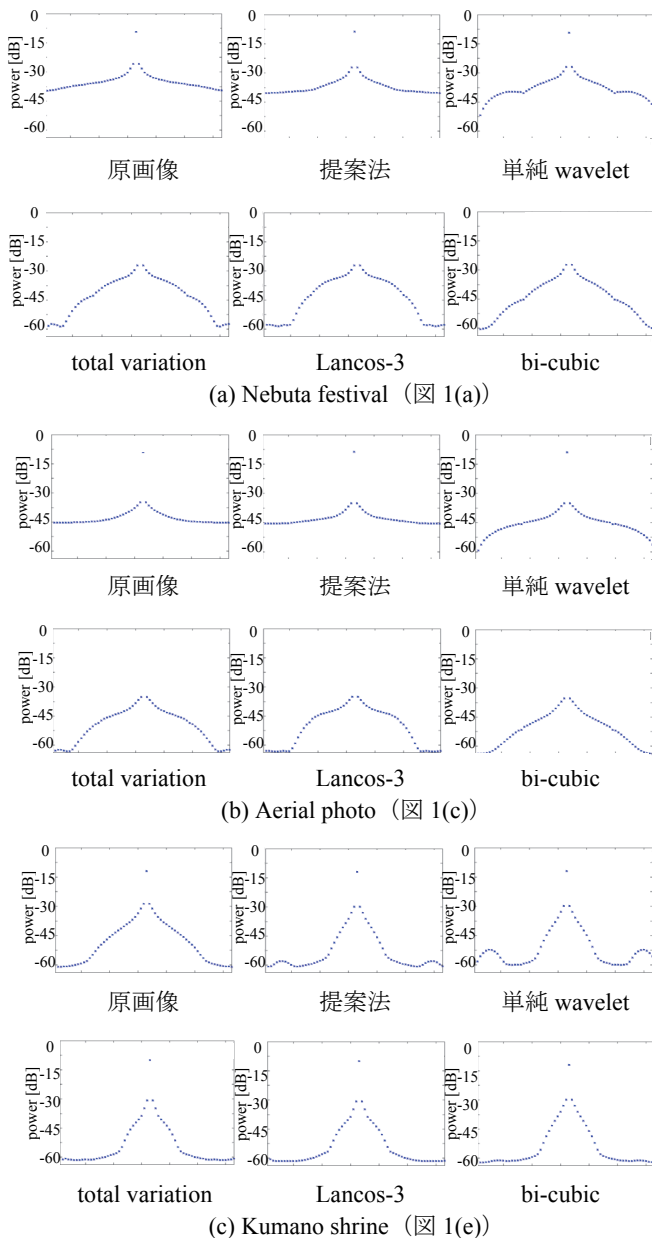


図8 実験結果画像の周波数スペクトル  
(横軸は中央を直流成分とする周波数)

## 5. まとめ

本研究では、原画像を信号成分と雑音成分に分離して、信号成分はその多重解像度成分とのレジストレーションを用いた超解像と、DI PSNR の考え方を基にした超解像最適化を行い、超解像信号成分と超解像雑音成分を合成することを特徴とする超解像法を提案した。結果の主観画質や空間周波数スペクトルより、提案法は従来法に比べて高い画質が得られていることを示した。

## 参考文献

- [1] Digital Cinema Initiatives, LLC, "Digital Cinema System Specification, Version 1.2," Mar. 07 2008.
- [2] SMPTE 428-1-2006, "Digital Cinema Distribution Master (DCDM)", 2006.
- [3] Y. Shishikui, Y. Fujita, and K. Kubota, "Super Hi-Vision – the star of the show!," *EBU Technical Review*, pp. 4-16, Jan. 2009.
- [4] SMPTE 2036-3-2010, "Ultra High Definition Television – Image Parameter Values for Program Production", 2010.
- [5] M. Schlockrman, S. Wittmann, T. Wedi, S. Kadono, "Film grain coding in H.264/AVC," *JVT of ISO/IEC MPEG & ITU-T VCEG, JVT-1034*, Sep. 2003.
- [6] C. V. Jiji, M. V. Joshi, and S. Chaudhuri, "Single-frame image super-resolution using learned wavelet coefficients," *International Journal of Image Systems and Technology*, vol. 14, no. 3, pp. 105-112, 2004.
- [7] M. Nguyen and P. Milanfar, "An efficient wavelet-based algorithm for image superresolution," *IEEE Proceedings on Image Processing*, vol. 2, pp. 351–354, 2000.
- [8] D. Capel, "Image Mosaicing and Super-resolution," Springer, U.S.A., 2004.
- [9] S. D. Babacan, R. Molina, and A. K. Katsaggelos, "Total variation super resolution using a variational approach," *Proceedings of IEEE ICIP*, pp.641-644, 2008.
- [10] 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.
- [11] A. Heinrich, G. de Haan, and C. N. Cordes, "A novel performance measure for picture rate conversion methods," *Digest of Technical Papers of ICCE*, pp.1-2, Jan. 2008.
- [12] T. Saito, K. Ishikawa, and T. Komatsu, "Super-resolution interpolation with a quasi blur-hypothesis," *Proceeding of IEEE ICIP*, pp.1169-1173, 2011.
- [13] Y. Matsuo, T. Nishida, S. Gohshi, and S. Sakaida, "Reducing Noise in High-resolution Video Sequences by Using Wavelet Shrinkage in the Temporal-spatial Domain," *Proceedings of IEEE PCS*, 2009.
- [14] 松本信幸, 井田孝, "画像のエッジ部の自己合同性を利用した再構成型超解像," 電子情報通信学会論文誌 D, vol. 93-D, pp. 118-126, no. 2, 2010.
- [15] R. R. Schulz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Transactions on Image Processing*, vol. 5, no. 6, pp.996-1011, 1996.
- [16] C. Tomasi and R. Manduchi, "Bilateral Filtering for Gray and Color Images," *Proceedings of IEEE ICCV*, pp.839-846, 1998.
- [17] K. T. Mullen, "The contrast sensitivity of human vision to red- green and blue yellow chromatic gratings," *Journal of Physiology*, no.359, pp. 381-409, 1985.