

幾何変換を伴う多重解像度分解成分の標本化構造を考慮した 時空間レジストレーションによる空間超解像法の検討

Image Super-resolution by Spatio-temporal Registration of Wavelet Multi-scale Components Considering Color Sampling Pattern with Affine Transformation

松尾 康孝†
Yasutaka Matsuo

境田 慎一†
Shinichi Sakaida

1. まえがき

日本のテレビ放送では、東京オリンピック・パラリンピックが開催される2020年に向けて、高い臨場感や実物感を持つ8K[1]放送が開始される。ここで4Kデジタルシネマ[2]や4K放送の映像を8K放送に用いる場合は、4K映像から8K映像への高解像度化を行う必要がある。さらに、この高解像度化を高画質に行うためには、空間超解像により4K映像が持つナイキスト周波数を超える超解像高周波成分を付加する必要がある。そこで本稿では、この空間超解像法について検討する。

2. 空間超解像法

空間超解像法は様々な方法が提案されているが、大きく分けると学習型超解像と再構成型超解像に分類される[3], [4]。さらに再構成型超解像は、大きく分けるとレジストレーション型超解像[3]と線形/非線形フィルタ型超解像[5]-[8]に分類される。各々の超解像法には、その利点と欠点に応じた相応の適用箇所が存在する。そこで以下では、各々の超解像法の特徴および利点と欠点を説明した後、本目的に適した超解像法を選択して、提案法を設計する。

学習型超解像では、予め学習によりデータベースを生成し、これを参照しながら超解像高周波成分を付加する。このデータベースは、学習用映像に対する低解像度映像を複数所有する。これらの低解像度映像は、学習用映像から撮像系の光学レンズで生じる低域通過フィルタ特性の劣化などを模擬して生成する。その後、被超解像映像のエッジなどの空間高周波成分を領域毎に解析し、データベース内の低解像度映像から合致するものを探索する。この探索結果を用いて、当該領域の空間高周波成分を高解像度な学習用映像の成分で置き換えることで、超解像高周波成分を付加する。学習型超解像では、高精度なデータベースが生成できる場合は高画質な超解像映像が得られるという利点がある。しかし本目的のように、放送などの様々な絵柄を含む一般画像において高画質な超解像映像を得るためには、膨大なデータベースが必要であるという欠点がある。

一方、再構成型超解像におけるレジストレーション型超解像では、被超解像映像内の被超解像フレーム画像とその前後フレーム画像間の位置合わせによって画素密度を向上させる。このため、位置合わせ精度と確度が高い領域では、高精度に超解像高周波成分を付加できる利点がある。しかし静止領域では、位置合わせを行っても画素密度が増えないため、原理的に超解像高周波成分を付加できないという欠点がある。また線形/非線形フィルタ型超解像では、被超解像映像の空間高周波成分を解析して、この空間高周波

成分とナイキスト周波数を超える超解像高周波成分の類似性を仮定した線形/非線形フィルタ処理によって超解像高周波成分を付加する。このため、静止領域においても、一定の精度で超解像高周波成分を生成できるという利点がある。しかし、必ずしも高精度に超解像高周波成分を生成できないという欠点がある。

以上のように各々の超解像法には様々な利点と欠点があるが、一般に4K映像などの高精細映像は、被超解像フレーム画像内にも多くの相似オブジェクトを持つ。これは、例えば花畑の大小様々な花や、ビル群の大小様々なビルなどがあげられる。このため、被超解像フレーム画像とその前後フレーム画像間で位置合わせを行うだけではなく、被超解像フレーム画像内の自己合同性の高さを利用した再構成型[9]のような超解像法が有効である。さらに、被超解像フレーム画像をデータベースとして、これを参照しながら超解像高周波成分を付加するような学習型超解像の良い部分を組み合わせれば、別途膨大なデータベースを用意せずに高画質化を図ることが可能になる。そこで、新しく学習型超解像の良い部分を導入した再構成型超解像法を提案する。

3. 提案法

提案法では、はじめに被超解像映像内の被超解像フレーム画像とその前後フレーム画像からデータベースを生成する。このデータベースは、これらのフレーム画像を多重解像度分解した複数の低解像度画像と、その空間高周波成分の組を所有する。そして被超解像フレーム画像と複数の低解像度画像間の位置合わせを行ったうえで、位置合わせ箇所では被超解像フレーム画像と低解像度画像が似ているため、被超解像フレーム画像のナイキスト周波数を超える未知の超解像高周波成分は、低解像度画像と組になっている空間高周波成分と似ていると仮定してこれを付加する。さらに位置合わせの精度を高めるために、被超解像映像の標本化構造を考慮した位置合わせ精度向上法、および超解像の正解が不明な場合でも高確度で処理できる超解像パラメータ選択法を導入する。以下、3.1では提案法の全体構成を説明し、3.2では提案法を構成する各要素技術を説明する。

3.1 全体構成

図1に提案法の全体図を示す。「幾何変換」ユニットでは、高精度なデータベースを生成するために、被超解像映像内の被超解像フレーム画像とその前後フレーム画像を幾何変換する。続いて「多重解像度分解」ユニットでは、データベースとして複数の低解像度画像とその空間高周波成分を生成するために、幾何変換画像を多重解像度分解する。「位置合わせ」ユニットでは、被超解像フレーム画像と複

†日本放送協会, Japan Broadcasting Corporation

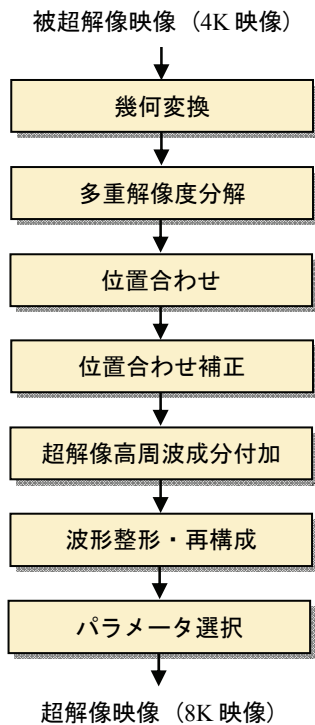


図 1 提案法全体図

数の低解像度映像間で位置合わせを行う。「位置合わせ補正」ユニットでは、位置合わせの精度を高めるために、各色信号の標本化構造を考慮した位置合わせ補正を行う。「超解像高周波成分付加」ユニットでは、位置合わせ補正情報を用いた超解像高周波成分の付加を行う。「波形整形・再構成」ユニットでは、超解像による過強調や強調不足を補正するために、付加した超解像高周波成分を波形整形した後、再構成して超解像画像を得る。最後に「パラメータ選択」ユニットにおいて被超解像フレーム画像を基準に DI PSNR (Double Interpolation PSNR) [10] 値が最大となる波形整形パラメータを選択することで、最適化された超解像画像が出力される。

3.2 詳細

以下では、図 1 の各ユニットにおける処理の詳細を説明する。「幾何変換」ユニットでは、被超解像フレーム画像とその前後フレーム画像をアフィン変換する。いま被超解像映像を F 、その水平、垂直、時間位置を x, y, t 、アフィン変換された映像を ${}^A F$ 、その水平、垂直、時間位置を u, v, t 、回転角度を θ 、水平、垂直拡大率を s_x, s_y 、スキュー角を δ とすると、アフィン変換後の ${}^A F$ の水平、垂直位置 (u, v) は式 1 から求められる。

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} s_x & 0 \\ 0 & s_y \end{pmatrix} \begin{pmatrix} 1 & \delta \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (1)$$

「多重解像度分解」ユニットでは、 ${}^A F$ に線形位相性を持つ CDF 9/7 ウェーブレット[11]を用いた N 階ウェーブレット分解を適用して、空間低周波帯域成分 ${}^A F_{LL}^N$ および水平、垂直、斜め高周波帯域成分 ${}^A F_{LH}^N, {}^A F_{HL}^N, {}^A F_{HH}^N$ を得る。

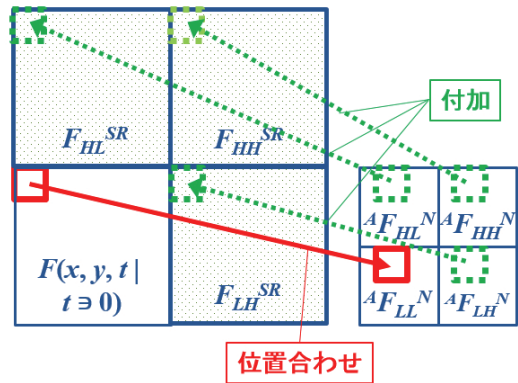
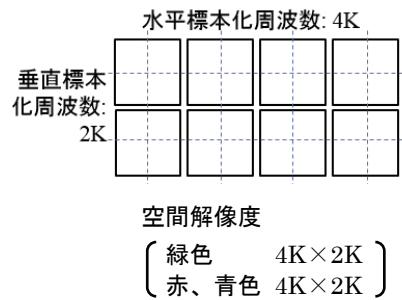
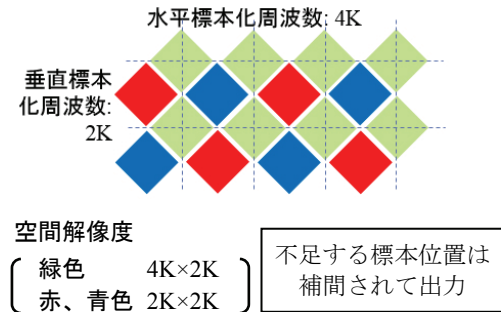


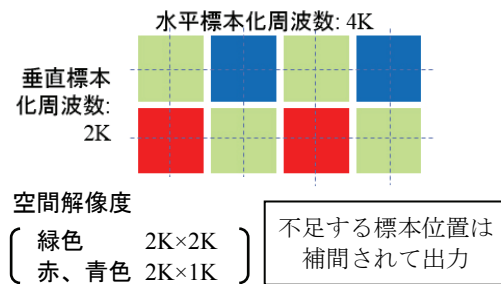
図 2 位置合わせと超解像高周波成分付加



(a) 正方形子型



(b) SONY F65 型 (SONY F65 カメラ)



(c) BAYER 型 (RED EPIC カメラなど)

図 3 4K カメラの標本化構造

「位置合わせ」ユニットでは、図 2 に示すように F 内の $t=0$ における被超解像フレーム画像 $F(x, y, t | t \ni 0)$ を基準フレーム、多重解像度分解ユニットで生成した ${}^A F_{LL}^N$ を参照フレームとして、似ている箇所を探索する。この位置合わせでは、 $F(x, y, t | t \ni 0)$ を 4×4 の小ブロックに分割し、小ブロック毎に参照フレームに対して SSD (Sum of Squared Difference) を評価関数とする全探索ブロックマッチングを行い、SSD 値が 4×4 (画素) $\times 8$ (レベル) の閾値以下であれば、その位置合わせ結果の確度は高いとして採用する。

「位置合わせ補正」ユニットでは、位置合わせ結果に対して、4K 映像を撮像したカメラが持つ各色信号の標準化構造を考慮した位置合わせ補正を行う。ここで図 3 に示すように、現在広く用いられている 4K カメラのうち、SONY F65 や RED EPIC カメラの標準化構造は、緑色信号の標準化周波数が赤、青色よりも高いため、映像内容が単色オブジェクトでない場合は緑色信号の位置合わせ精度は赤、青色よりも高い。このため位置合わせを行うブロック領域が単色ではなく、かつ各色信号の位置合わせ結果が 2 画素未満であれば、赤、青色信号の位置合わせ結果を緑色信号の結果で置き換える。

「超解像高周波成分付加」ユニットでは、図 2 に示すように $F(x, y, t | t \ni 0)$ のナイキスト周波数を超える超解像水平、垂直、斜め高周波帯域成分 F_{LH}^{SR} , F_{HL}^{SR} , F_{HH}^{SR} を、位置合わせまたは位置合わせ補正結果を用いて付加する。ここでは、 $F(x, y, t | t \ni 0)$ 内の任意の小ブロック領域が ${}^A F_{LL}^N$ 内で位置合わせされた位相位置の成分と似ているならば、 F_{LH}^{SR} , F_{HL}^{SR} , F_{HH}^{SR} 内の当該小ブロック成分は ${}^A F_{LL}^N$ 内で位置合わせされた位相位置の ${}^A F_{LH}^N$, ${}^A F_{HL}^N$, ${}^A F_{HH}^N$ に似ているとして、これらを付加する。

「波形整形・再構成」ユニットでは、超解像高周波成分を付加した後、その過強調や強調不足を補うために F_{LH}^{SR} , F_{HL}^{SR} , F_{HH}^{SR} に分散値 σ , ゲイン γ のガウシアンフィルタを適用する。その後、 $F(x, y, t | t \ni 0)$ を空間低周波帯域、 σ と γ を変化させてガウシアンフィルタを適用した F_{LH}^{SR} , F_{HL}^{SR} , F_{HH}^{SR} を空間高周波帯域成分として 1 階ウェーブレット再構成することで、 σ と γ による複数の超解像画像 $F_{\sigma, \gamma}^{SR}$ を生成する。

「パラメータ選択」ユニットでは、 $F(x, y, t | t \ni 0)$ を基準に、DI PSNR 値が最大となる σ と γ を 2×2 画素領域毎に選択することで、最終的な超解像画像 F^{SR} が得られる。

4. 実験

図 4 に実験に用いる評価用映像を示す。このうち「Kimono」は NHK 8K カメラ (8K, 60 フレーム/秒, 正方格子型標準化構造), 「Vegetable」は SONY F65 カメラ (4K, 24 フレーム/秒, SONY F65 型標準化構造), 「Cherry blossom」は RED EPIC カメラ (4K, 60 フレーム/秒, BAYER 型標準化構造) で撮像された映像である。実験では、PSNR 値を測定するために、これらの映像に式 2 に示す 11 タップの低域通過型フィルタを適用して、「Kimono」は 4K, 「Vegetable」と「Cherry blossom」は 2K に縮小した映像を被超解像映像とする。

$$[21 \ 0 \ -52 \ 0 \ 159 \ 256 \ 159 \ 0 \ -52 \ 0 \ 21] ./ 512 \quad (2)$$

(./は分子の各係数を分母値で除算)



(a) Kimono
(8K, 60P, 正方格子型標準化構造)



(b) Vegetable
(4K, 24P, SONY F65 型標準化構造)



(c) Cherry blossom
(4K, 60P, BAYER 型標準化構造)

図 4 評価用映像

表 1 超解像法

Bi-cubic (BC)
Multi-frame registration (MFR) [5]
Wavelet zero padding (WZP)
WZP and directional cycle spanning (DCS) [7]
Discrete and stationary wavelet decomposition (DSW) [8]
提案法 (被超解像フレーム画像のみを使用) (SPM)
提案法 (PM)

表 2 PSNR 値 (赤(R), 緑(G), 青(B)色信号) ([dB])

		BC	MFR	WZP	DCS	DSW	SPM	PM
Kimono	R	39.21	39.93	40.55	40.68	40.55	41.45	41.49
	G	37.46	38.01	38.36	38.44	38.36	39.27	39.36
	B	24.20	24.52	24.67	24.67	24.66	25.70	25.90
Vegetable	R	41.31	41.90	42.84	43.51	42.84	44.46	44.57
	G	38.17	38.71	39.44	39.66	39.44	41.06	41.18
	B	30.96	31.54	32.00	32.04	32.00	33.52	33.72
Cherry.	R	33.49	35.34	39.96	40.40	40.02	40.55	40.67
	G	33.28	35.58	40.57	41.10	40.63	41.53	41.58
	B	30.71	33.72	36.48	36.88	36.53	37.69	37.75

表1に実験に用いる超解像法を示す. このうち Wavelet zero padding (WZP), WZP and directional cycle spanning (DCS), Discrete and stationary wavelet decomposition (DSW), 提案法 (被超解像フレーム画像のみを使用) (SPM), 提案法 (PM)では, CDF 9/7 ウェーブレットを用いる. そして Bi-cubic (BC), WZP, DCS, DSW, SPM は被超解像映像 F 内の被超解像フレーム画像 $F(x, y, t | t \ni 0)$ のみを用い, Multi-frame registration (MFR)と PM では被超解像フレーム画像とその前後2フレーム画像 ($F(x, y, t | t \ni -2, -1, 0, 1, 2)$)を用いる. 予備実験により, 式1における θ, s_x, s_y, δ は, $\theta \ni -\pi/4, 0, \pi/4, s_x, s_y \ni 1.5, 2.0, \delta \ni 0$ とした. さらに波形整形のパラメータ選択における σ と γ は, $\sigma \ni 0.1, 1.0, 2.0, \gamma \ni 0, 0.05, 0.2, 1.0$ とした.

表2に, PSNR 測定結果を示す. PMはBC, MFR, WZP, DCS, DSW に対して 0.27~1.68[dB]高い PSNR 値が得られた. さらに PM は SPM に対して被超解像フレーム画像の前後フレーム画像からの位置合わせを導入することで, PSNR 値が 0.04~0.20[dB]向上した. これらの結果のうち「幾何変換」ユニットの処理を行うことで, PSNR 値が約 0.05[dB]向上した. また「Sakura」, 「Vegetable」では, 「位置合わせ補正」ユニットの処理を行うことで, R, B の PSNR 値が約 0.03[dB]向上した.

図5に「Kimono」, 「Vegetable」, 「Cherry blossom」の BC, DCS, PM の切り出し画像を示す. 同図より, PM は BC に対してぼやけが大きく復元されていることを確認した. さらに DCS やその他の方法に対しても斜め方向のエッジ成分などの過強調などが抑えられており, 良好な超解像画質が得られた.

5. むすび

本稿では, 4K 映像などの高精細映像が被超解像フレーム画像内にも多くの相似オブジェクトを持つことに着目し, 被超解像フレーム画像とその幾何変換を伴うウェーブレット分解空間低周波成分 (被超解像フレーム画像の低解像度画像)間で位置合わせを行うことを特徴とする, 新しいレジストレーション型超解像法を提案した. さらにこの位置合わせでは, 被超解像フレーム画像の前後フレーム画像との位置合わせも同様に行った. これらの位置合わせでは, 4K 映像を撮像したカメラの標準化構造を考慮した位置合わせ補正を行った. そして波形整形とそのパラメータ選択も行い, 超解像画質を高めた. 実験により, 提案法は従来法に対して高い PSNR 値と主観画質が得られることを確認した.

6. 参考文献

- [1] Rec. ITU-R BT.2020-1, "Parameter values for ultra-high definition television systems for production and international programme exchange", 2014.
- [2] Digital Cinema Initiatives LLC, "Digital cinema system specification version 1.2", 2008.
- [3] 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. 21-26, 2003.
- [4] 奥富正敏, 田中正行, 竹島秀則, 松本信幸, "画像超解像処理技術の最新動向", 電子情報通信学会誌, vol. 93, no. 8, pp.693-698, 2010.

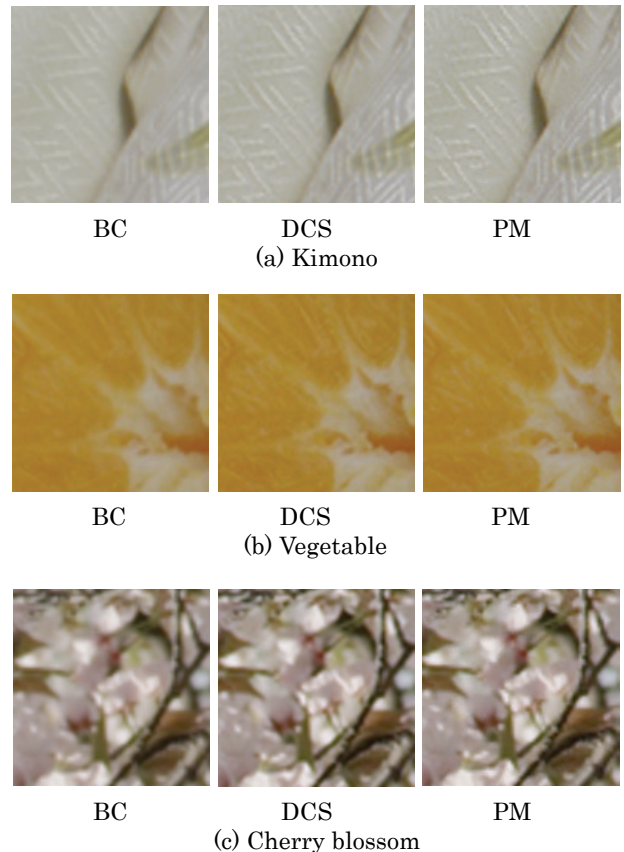


図5 超解像画像 (切り出し画像)

- [5] A. Temizel, and T. Vlachos, "Wavelet domain image resolution enhancement using cycle-spanning", Journal of Electronics Letters, vol. 41, no. 3, pp. 119-121, 2005.
- [6] D. Hasan, and A. Gholamreza, "IMAGE resolution enhancement by using discrete and stationary wavelet decomposition", IEEE Transactions on Image Processing, vol. 20, no. 2, pp. 1458-1460, 2011.
- [7] 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.
- [8] 桜井優, 吉川明博, 鈴木彰太郎, 後藤富朗, 平野智, "Total Variation 正則化手法と事例学習法を組合せた超解像度画像の復元法", 映像情報メディア学会誌, vol. 64, no. 11, pp. 1613-1620, 2011.
- [9] 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", 電子情報通信学会誌 D, vol. J93-D, no. 2, pp. 118-126, 2010.
- [10] T. Saito, K. Ishikawa, and T. Komatsu, "Super-resolution interpolation with a quasi blur-hypothesis", Proceedings of IEEE ICIP, pp. 1169-1173, 2011.
- [11] A. Cohen, I. Daubechies, and J. -C. Feauveau, "Biorthogonal bases of compactly supported wavelets", Communications on Pure and Applied Mathematics, vol. 45, no. 5, pp. 485-560, 1992.