

単一画像を用いた再構成型超解像法の一考察 A Study on Single Frame Reconstruction-based Super Resolution

佐渡山 史矢[†]
Fumiya Sadoyama

荻野 正[†]
Tadashi Ogino

山田 親稔[†]
Chikatoshi Yamada

1. まえがき

携帯電話やスマートフォンへの高解像度カメラの搭載や、ネットワークの高速化、大量データの高速処理技術の進歩などにより、高解像度の画像が大量に流通している。企業においても製品画像の提供などにとどまらず、製品設計から保守のフェーズにいたるまで高解像度画像を用いることで、顧客とのコミュニケーションをさらに密にし、ビジネスの差別化を図ることが可能となる。流通する画像データは、旧型、あるいは低価格デバイスからの低解像度画像と最新デバイスからの高解像度画像が共存し、これらを統一的に扱うために、画像の高解像度化技術が注目されている。

高解像度化技術として、これまでは補間法が用いられてきたが、補間法を用いた場合、画像が本来持つ高周波成分が生成されずにボケ、エッジ部にジャギーやリング等のアーティファクトが目立った結果となってしまう。そこで近年、補間法に代わる手法として超解像法が注目されている。超解像法の代表的なものとして、再構成型 [1][2][3]、学習型 [4] の超解像法 [5][6] が挙げられる。再構成型超解像とは、補間法により初期設定した高解像度画像から低解像度画像を推定し、推定低解像度画像と入力低解像度画像との差が少なくなるように高解像度を更新する手法である。しかし、この手法には拡大倍率を大きく取れないという問題点が存在する。また、学習型超解像とは学習用の高解像度画像と、高解像度画像を縮小・拡大することで劣化させた低解像度画像の対応を学習させたデータベースから、入力画像の高解像度画像を推定する手法である。この手法は推定の精度が高く、拡大倍率を大きく取ることができる。しかしながら、復元したい情報がデータベースに存在しない場合、画像が本来持たない構造を復元してしまう問題点があり、正確な復元には膨大な学習データが必要となる。

本研究では、唯一枚の入力画像から高拡大倍率を許容し、アーティファクトを抑えた超解像法を実現し、再構成型、学習型の2つの超解像法の持つ問題点を解消することを目的としている。

2. 提案手法

本章では提案手法の原理について述べる。本手法は、唯一枚の画像から高拡大倍率で拡大できる超解像の実現のために、従来の再構成型超解像 [1] と画像の自己合同性に着目した手法 [2][3] を組み合わせ、改良したものである。提案手法のフローチャートを図1に示す。本手法でははじめに入力画像に対し、Lanczos法で所望の倍率で補間処理を行う。次に、補間処理された画像の局所分散を求め、高分散領域と低分散領域とに処理領域の分割を行う。高分散領域には、類似構造の探索と類似データ

の組み合わせを行い、エッジを保存するために局所分散を考慮した Bilateral Filter 処理を行う。最後に、局所分散を考慮した Bilateral Back Projection を行うことで入力画像から高解像度画像を推定する。



図1: Flowchart of proposed method.

2.1 処理領域の分割

エッジ保存を行うために用いる平滑化フィルタなどでは、細かなエッジが潰れてしまう。そのため、本手法では局所分散を求めて処理領域の分割を行う。局所分散とは、注目画素を中心とした任意の大きさの小領域（局所領域）における画素値の分散のことであり、式(1)で表される。

$$v(j, i) = \frac{\sum_{n=-k}^k \sum_{m=-k}^k (f(j+n, i+m) - avg(j, i))^2}{\text{masksize}} \quad (1)$$

ここで、 $f(j+k, i+l)$ は参照画素、 $avg(j, i)$ は注目画素を中心とした領域の平均値、 k はマスク半径、 masksize はマスク内の画素数である。 $v(j, i)$ がしきい値以上の場合を高分散領域とみなし、それ未満の場合を低分散領域とみなす。局所分散がしきい値以上の領域を高分散領域である白、しきい値未満の領域を低分散領域である黒とした画像を図2(b)に示す。図2(b)から、高分散領域は画素値のバラつきが大きいエッジ部、低分散領域は画素値のバラつきが小さい平坦部となっていることがわかる。処理領域の分割に基づき、次節以降3つの処理は高分散領域にのみ行う。

2.2 類似構造の探索

画像内には図3に示すような類似構造が存在することに着目し、類似構造の探索を行う。探索方法としては、Lanczos法で拡大された画像内でブロックマッチングを行い、類似度を求めることで、類似構造を探索する。類似度は式(2)で表される。

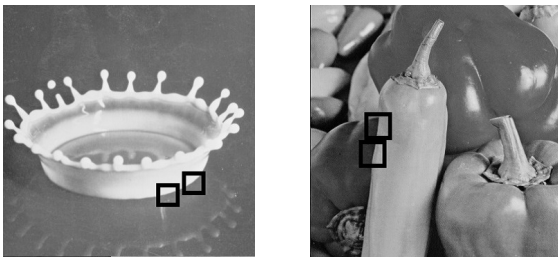
[†]沖縄工業高等専門学校



(a) Original image.

(b) Divided image.

図2: Division of processing region based on the local variance.



(a) Milk-drop

(b) Peppers

図3: Examples of image similar structures.

$$d(j, i) = \|P_y(j_0, i_0) - P_y(j, i)\|_2 \quad (2)$$

ここで、 $P_y(j_0, i_0)$ は注目画素を中心とした 7×7 領域のパッチ、 $P_y(j, i)$ は参照画素を中心とした 7×7 領域のパッチである。2つのパッチの各成分の差のユークリッドノルムをとることで、パッチ間のブロックマッチングを行う。類似度 $d(j, i)$ が、しきい値以下であるときに2つのパッチが類似構造であるとみなす。この処理を注目画素を中心とする 100×100 領域内で、参照画素を3ピクセルずつずらして行う。類似構造が見つかった場合、注目画素に対応する参照画素・類似度格納用の構造体であるセット $S(j_0, i_0)$ に、参照画素と類似度 $d(j, i)$ のペアを格納する。類似構造が複数見つかった場合には、類似度が小さい順に、参照画素値と類似度のペアをすべて格納する。セット S 内に格納された参照画素値と類似度は、次節の類似データの組み合わせで用いる。

2.3 類似データの組み合わせ

類似データの組み合わせではセット S に格納された参照画素と類似度 d を用いて加重平均をとり、注目画素の値を更新する。更新式は式 (3) で表される。

$$\tilde{I}_H(j_0, i_0) = \frac{\hat{y}(j_0, i_0) + w_{max} \times y(j_0, i_0)}{w_{sum} + w_{max}} \quad (3)$$

ここで、 w_{sum} はセット $S(j_0, i_0)$ 中の重みの総和、 w_{max} は $S(j_0, i_0)$ 内における重みの最大値を表す。また、

$\hat{y}(j_0, i_0)$ は式 (4) で表される。

$$\hat{y}(j_0, i_0) = \sum_{y(j, i) \in S(j_0, i_0)} w(j, i) y(j, i) \quad (4)$$

重み $w(j, i)$ は式 (5) で表される。

$$w(j, i) = \exp\left(-\frac{d(j, i)}{t}\right) \quad (5)$$

ここで、 t は重みを調節するためのパラメータである。類似している画素を用いて加重平均をとることで、Lanczos 補間によりボケてしまったエッジ部分を本来取りうるべき画素値に近づけることができる。

2.4 局所分散を考慮した Bilateral Filter

類似構造の探索とデータの組み合わせのみではエッジの保存は不十分である。そこで、本手法では局所分散を考慮した Bilateral Filter[7] 処理を行う。この処理は式 (6) で表される。

$$I_H(j, i) = \frac{\sum_{n=-k}^k \sum_{m=-k}^k (f(j+n, i+m) \times w(j, i; n, m))^2}{\sum_{n=-k}^k \sum_{m=-k}^k w(j, i; n, m)} \quad (6)$$

ここで、 $f(j+k, i+l)$ は参照画素、 $w(j, i; n, m)$ は重み、 k はマスク半径である。重み $w(j, i; n, m)$ は式 (7) で表される。

$$w(j, i; n, m) = w_d(j, i; n, m) \times w_r(j, i; n, m) \quad (7)$$

$w_d(j, i; n, m)$ 、 $w_r(j, i; n, m)$ はそれぞれ空間方向、画素値方向の重みであり、式 (8)、(9) で表される。

$$w_d(j, i; n, m) = \exp\left(-\frac{n^2 + m^2}{2\sigma_d(j, i)^2}\right) \quad (8)$$

$$w_r(j, i; n, m) = \exp\left(-\frac{(f(j, i) - f(j+n, i+m))^2}{2\sigma_r^2}\right) \quad (9)$$

ここで、 σ_r は画素値方向のガウス関数の標準偏差である。 $\sigma_d(j, i)$ は式 (10) で表される。

$$w_r(j, i; n, m) = \sigma_d \left(\frac{v(j, i)}{v_{max}} \right)^{\frac{1}{7}} + \epsilon \quad (10)$$

ここで、 σ_d は空間方向のガウス関数の標準偏差、 $v(j, i)$ は式 (1) で表される局所分散、 v_{max} は局所分散の最大値、 ϵ は式 (8) で零割を避けるための微小な正の定数である。Bilateral Filter は空間方向の変化だけでなく、画素値方向の変化を考慮した、エッジ保存平滑化能力をもつ非線形フィルタである [8][9]。しかし、Bilateral Filter はエッジ部、平坦部といった局所領域の情報を考慮していないために、平坦部において過度な平滑化が見られる

という問題点がある。そこで、平坦部における過度な平滑化を防ぐため、 σ_d を局所領域に適した空間方向の重みパラメータに $\tilde{\sigma}_d$ に変換する。そのために局所分散 $v(j, i)$ を式 (10) を用いて非線形変換する。図 4 に示すように、 γ の値により関数の形状が変化し、 $\gamma < 1$ のときは下に凸、 $\gamma > 1$ のときは上に凸となり、平滑化の度合いを調節することができる。空間方向の重みパラメータを可変とすることで、局所分散が大きい強エッジ部では、ガウス関数の広がりが大きくなり、アーティファクト除去、先鋭効果が期待できる。また、局所分散が比較的小さい平坦部等ではガウス関数の広がり小さくなり、過度な平滑化を防ぐことができる。

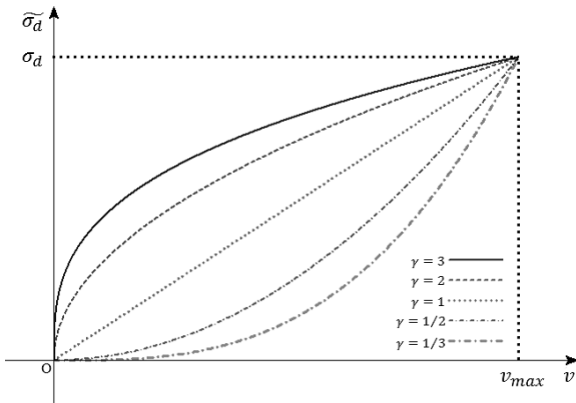


図 4: Change in shape of the tone curve by γ .

2.5 局所分散を考慮した Bilateral Back Projection

最後に局所分散を考慮した Bilateral Back Projection (BBPLV) [7][10] 処理を行い、高解像度画像を更新・推定する。この処理は式 (11) で表され、再構成型超解像に、問題点であるアーティファクトを低減するために、局所分散を考慮した Bilateral Filter を加えたものである。

$$I_H^{t+1} = (I_H^t + \tau (I_H^0 - ((I_H^t * G) \downarrow_s) \uparrow_s) * P) * B \quad (11)$$

ここで τ はステップ、 t は繰り返し回数、 G はガウシアンフィルタ、 P は Back Projection カーネル、 B は局所分散を考慮した Bilateral Filter、 $*$ は畳み込み演算子、 \uparrow_s はアップサンプリング演算子、 \downarrow_s はダウンサンプリング演算子である。Back Projection カーネル P では、平均化フィルタと先鋭化フィルタを順に適用する。平均化フィルタのフィルタ係数は $[0,1,0;1,2,1;0,1,0]/6$ 、先鋭化フィルタのフィルタ係数は $[0,-0.25,0;-0.25,2,-0.25;0,-0.25,0]$ とする。繰り返し回数 $t(t = 0, 1, \dots, T-1)$ による最終的な推定高解像度画像 I_H^T は次の最小化問題に帰着する。

$$I_H^T = \arg \min_{I_H} \sum \sum \| (I_H^0 - ((I_H^t * G) \downarrow_s) \uparrow_s) \|^2 \quad (12)$$

式 (12) に示す最小化問題を解くアルゴリズムが式 (11) となる。

3. 実験

3.1 実験方法

原画像を縮小・高解像度化し、図 1 および 2 章で述べた提案手法を用いて、2 倍・4 倍の超解像処理を施して従来手法との比較を行う。従来手法は Lanczos 法とし、入力画像に SIDBA (Standard Image Data-BAsE) の標準画像を用いる。評価指標は PSNR とする。

3.2 実験結果

実験による超解像処理後の PSNR による評価を表 1、処理画像を図 5 および図 6 に示す。示した画像は 2 倍の処理結果である。これらの結果からボケの除去効果、アーティファクトの低減、PSNR が向上していることが確認できる。

3.3 考察

画質の改善が不十分であった理由として、拡大する際に失われる高周波成分の復元が不十分である点が原因のひとつであると考えられる。そのため、拡大アルゴリズムの改良や高周波成分に特化した復元方法についての検討が必要である。また、適用するフィルタのパラメータが不適切である可能性もあるため、再検討が必要である。

表 1: PSNR.

	Milk-drop		Lenna	
	Lanczos	Proposed	Lanczos	Proposed
2 倍	31.06	32.72	29.17	30.93
4 倍	26.48	26.91	25.10	25.40

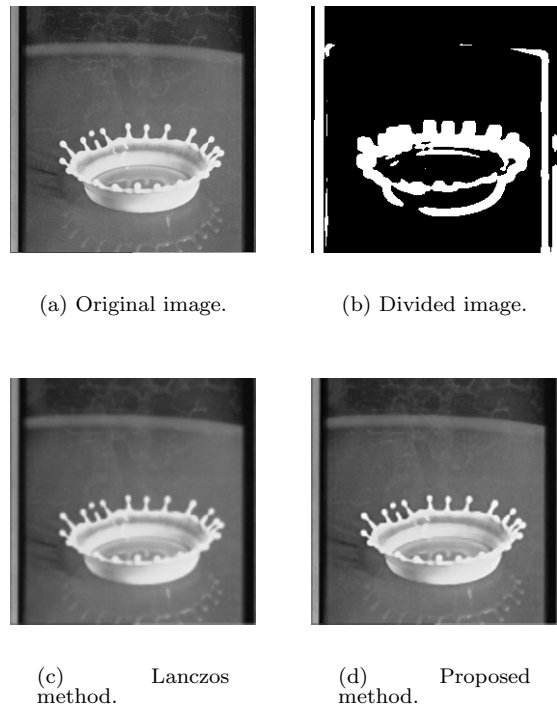


図 5: Milk-drop.



図 6: Lenna.

4. まとめ

再構成型超解像には、精度・許容倍率が低いという問題点があり、学習型超解像には大量の学習画像が必要で、入力画像をに対する汎用性が低いという問題点が存在している。この問題を排除した、入力画像によらず唯一枚の画像から、高精度、高拡大倍率を許容する超解像法の実現を目的として、単一画像を用いた超解像法の提案を行った。提案手法を実装し、従来手法との比較を行い、画像全体のボケ除去効果、エッジ部におけるジャギーやリングング等のアーティファクトの低減、PSNRの値の向上を確認した。しかし、画質改善はまだ不十分である。そのため、適用するフィルタの再検討、拡大アルゴリズムの見直し等が今後の課題である。

参考文献

- [1] M. Irani and S. Peleg : “Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency”, *JV-CIP*, Vol.4, pp.324-335 (1993)
- [2] W. Dong, L. Zhang, G. Shi, and X. Wu : “Non-local back-projection for adaptive image enlargement”, *ICIP*, pp.349-352 (2009)
- [3] 藤原 英晃・近藤 克哉 : 「エッジ保存を考慮した単一画像の再構成型超解像」, 電子情報通信学会技術研究報告. SIS, スマートインフォメディアシステム, Vol.111, No.457 pp.43-48 (2012)
- [4] W. Wu and Z. Liu : “Learning-based super resolution using kernel partial least squares”, *Image*

and Vision Computing, Vol.29, No.6 pp.394-406 (2011)

- [5] Y.W. Tai, S. Liu, M.S. Brown, and S. Lin : “Super Resolution using Edge Prior and Single Image Detail Synthesis”, *CVPR*, pp.2400-2407 (2010)
- [6] Y. Fan, Z. Gan, Y. Qie, and X. Zhu : “Single Image Super Resolution Method Based on Edge Preservation”, *ICIG*, pp.394-399 (2011)
- [7] 橋上 英宜・延原 肇 : 「局所分散を考慮したバイラテラルバックプロジェクションと逆フィルタを用いた一枚超解像手法」, 電子情報通信学会技術研究報告. SIS, スマートインフォメディアシステム, Vol.110, No.445 pp.31-36 (2011)
- [8] C. Tomasi and R. Manduchi : “Bilateral filtering for gray and color images”, *ICCV*, pp.839-846 (1998)
- [9] 浦浜 喜一 : 「画像の本質を引き出すバイラテラルフィルタによるノイズ除去とイラスト風画像の生成」, 映像情報メディア学会誌, Vol.62, No.8 pp.1268-1273 (2008)
- [10] S. Dai, M. Han, Y. Wu, and Y. Gong : “Bilateral backprojection for single image super resolution”, *Multimedia and Expo*, pp.1039-1042 (2007)
- [11] 佐渡山 史矢・山田 親稔 : 「単一画像のエッジ保存型超解像」, 電子情報通信学会ソサイエティ大会講演論文集 (2013)
- [12] F. Sadoyama, C. Yamada, and S. Ichikawa : “Single Frame Super Resolution using Reconstruction Method”, *International Workshop on Networking, Computing, Systems, and Software* (2013)