

高周波強調処理と入力画像の利用による学習型超解像

小川祐樹[†] 滝口哲也^{††} 有木康雄^{††}

[†] 神戸大学大学院システム情報学研究科

657-8501 兵庫県神戸市灘区六甲台町 1-1

^{††} 神戸大学自然科学系先端融合研究環

657-8501 兵庫県神戸市灘区六甲台町 1-1

E-mail: [†]ogawa@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp

あらまし 近年、コンピュータビジョンの分野において、超解像とよばれる手法が活発に研究されている。一枚画像を超解像する手法として代表的なものに、Example-based 手法がある。そのなかでも、ウェーブレット変換と組み合わせた手法も存在するが、高周波成分を改善する余地がある。本稿では、高周波成分を改善するために新しく2つの改善点を提案する。一つ目は画像データベースを作成する際に、従来と違い学習画像のみならず、入力画像も用いることである。二つ目は複数の性質の似たパッチを組み合わせることで、高周波成分をより復元させることである。この二つの改善による効果を実験により確認した。

キーワード 超解像 ウェーブレット変換 1 枚画像

Learning Based Super-resolution Combined with Input Image and Emphasized High Frequency

Yuki OGAWA[†], Tetsuya TAKIGUCHI^{††}, and Yasuo ARIKI^{††}

[†] Graduate School of System informatics, Kobe University

1-1,Rokkodai,Nada,Kobe,Hyogo,657-8501 Japan

^{††} Organization of Advanced Science and Technology,Kobe University

1-1,Rokkodai,Nada,Kobe,Hyogo,657-8501 Japan

E-mail: [†]ogawa@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp

Abstract In recent years, super-resolution techniques in the field of computer vision has been studied actively. The conventional method, as in a typical method for single image super-resolution technique is Example-based. Although this method is combined with wavelet transform, there is some room to be improved for high frequency quality. In this paper, we propose two methods to solve the high frequency components. The first one is to utilize the input image as well as the training images in re-constructing the high frequency components. The second is the combination of multiple patches (small region) with the similar low frequency components in the high frequency component re-construction. We confirmed the effectiveness of the proposed method through the experiments.

Key words super-resolution wavelet-transform single-image

1. はじめに

近年、デジタルカメラや携帯電話搭載のカメラの解像度は飛躍的に向上した。その一方で、価格競争の激化により、撮像素子や光学系のコストアップが深刻な問題となっており、より安価な撮像系を用いて、デジタル画像処理により高画質化を図る技術が注目を集めている。さらに、携帯電話に代表される小型情報機器などのカメラはデジタルカメラと比べると低解像度であり、写真などで保存する際には拡大処理を施して、高解像度に変換すること

が必要になる。一般的な拡大処理手法の代表例である共3次内挿法や線形補間法を用いて画像を拡大すると、画像の輪郭付近の情報を失ってしまい、全体的にぼやけた画像になってしまう。そこで画像を拡大する際に適切な高周波成分を付加してやることによって、より高解像度の画像を作成する方法を考える。近年、この技術はコンピュータビジョンの分野において超解像とよばれる手法として活発に研究されている。この超解像は何らかの原因で失われてしまった観測データから元データを復元する手法である。つまり、超解像とは失われてしまった高

周波成分を何らかの方法で復元する技術であると言える。本稿では、1枚画像を超解像する手法に焦点をあてる。一枚の画像を超解像する手法の中で、代表的なものとして Example-based 手法[1]がある。この手法は、画像の低解像度パッチと高解像度のパッチをペアにしてデータベースを作成しておき、入力画像が与えられると、それを分解した低解像度パッチとデータベースを比較し、高解像度パッチに置き換えることで高解像度の画像を得る手法である。多くの手法は、この手法の一部を改良している[2]～[7]。そのなかで、ウェーブレット変換と組み合わせた手法[2]が存在するが、高周波成分の復元に関して改善する余地がある。本研究では、そのために新しく2つの改善点を提案する。一つ目は画像データベースを作成する際に、従来と違い学習画像のみならず、入力画像も用いることである。二つ目はパッチ(部分領域)を重ね合わせる際に、複数の性質の似たパッチを組み合わせることで高周波成分を強調させることである。

2. ウェーブレット変換

提案手法では、ウェーブレット変換[8]を用いるので、本章ではウェーブレット変換について述べる。

2.1 2次元離散ウェーブレット変換

ここでは、本研究で用いている離散ウェーブレットを使用した2次元画像のウェーブレット変換について述べる。画像データは2次元の離散データとして与えられるものとし、 $f(m, n)$ で表す。2次元離散フーリエ変換と同様に、まず、横方向に離散ウェーブレット変換を行い、その係数に対して縦軸方向に離散ウェーブレット変換を行う。入力画像データ $f(m, n)$ をレベル0のスケール係数 $s_{LL, m, n}^{(0)}$ とみなす。数列 p_k, q_k は予め与えられている。 $\bar{p}_k, \bar{q}_k$ は p_k, q_k の複素共役である。例えば、ドベシのスケール関数を表す数列 p_k は表1で与えられる。提案手法では $N = 1$ の数列を用いている。

表1 ドベシの数列

N=1	N=2	N=3
0. 707106781	0. 482962913	0. 332670553
0. 707106781	0. 836516304	0. 806891509
	0. 224143868	0. 459877502
	-0. 129409523	-0. 135011020
		-0. 085441274
		0. 035226292

まず、横方向に離散ウェーブレット変換を行う

$$s_{L, m, n}^{(j+1, x)} = \sum_k \bar{p}_{k-2m} s_{k, n}^{(j)} \quad (1)$$

$$w_{H, m, n}^{(j+1, x)} = \sum_k \bar{q}_{k-2m} s_{k, n}^{(j)} \quad (2)$$

ここで $s_{L, m, n}^{(j+1, x)}$ および $w_{H, m, n}^{(j+1, x)}$ はそれぞれ、横軸方

向のスケール係数、及びウェーブレット展開係数を表す。 $j = 0$ の場合を図1に示す。

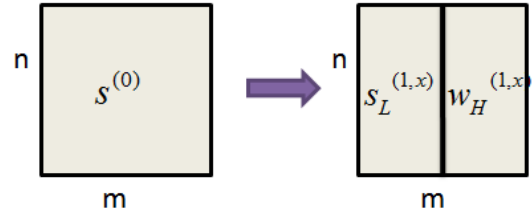


図1 $s_{m, n}^{(0)}$ の分解

次にそれぞれの係数に対し、縦軸方向に離散ウェーブレット変換を行う。

$$s_{LL, m, n}^{(j+1)} = \sum_l \bar{p}_{l-2n} s_{L, m, l}^{(j+1, x)} \quad (3)$$

$$w_{LH, m, n}^{(j+1, h)} = \sum_l \bar{q}_{l-2n} s_{L, m, l}^{(j+1, x)} \quad (4)$$

$$w_{HL, m, n}^{(j+1, v)} = \sum_l \bar{p}_{l-2n} w_{H, m, l}^{(j+1, x)} \quad (5)$$

$$w_{HH, m, n}^{(j+1, d)} = \sum_l \bar{q}_{l-2n} w_{H, m, l}^{(j+1, x)} \quad (6)$$

ここで、 $w_{LH, m, n}^{(j+1, h)}$ は横軸方向にスケール関数、縦軸方向にウェーブレットを作用させた係数、 $w_{HL, m, n}^{(j+1, v)}$ は横軸方向にウェーブレット、縦軸方向にスケール関数を作用させた係数、 $w_{HH, m, n}^{(j+1, d)}$ は縦横方向ともにウェーブレットを作用させた係数を示す。 $j = 1$ の場合を図2に示す。

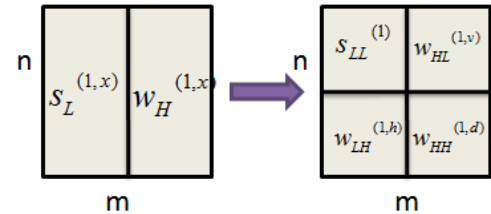


図2 $s_{L, m, n}^{(1, x)}$ と $w_{H, m, n}^{(1, x)}$ の分解

これらの式をまとめると以下ようになる。

$$s_{LL, m, n}^{(j+1)} = \sum_l \sum_k \bar{p}_{k-2m} \bar{p}_{l-2n} s_{k, l}^{(j)} \quad (7)$$

$$w_{LH, m, n}^{(j+1, h)} = \sum_l \sum_k \bar{p}_{k-2m} \bar{q}_{l-2n} s_{k, l}^{(j)} \quad (8)$$

$$w_{HL, m, n}^{(j+1, v)} = \sum_l \sum_k \bar{q}_{k-2m} \bar{p}_{l-2n} s_{k, l}^{(j)} \quad (9)$$

$$w_{HH, m, n}^{(j+1, d)} = \sum_l \sum_k \bar{q}_{k-2m} \bar{q}_{l-2n} s_{k, l}^{(j)} \quad (10)$$

上式の中で $s_{LL, m, n}^{(j+1)}$ の部分つまり、縦横の長さが1/2の大きさの入力画像を入力画像とおき、さらに4つの

成分に分解することにより、2次元の多重解像度解析を行う。

この再構成方法は、以下のように行う。

$$s_{m,n}^{(j)} = \sum_k \sum_l (p_{m-2k} p_{n-2l} s_{LL,k,l}^{(j+1)} + p_{m-2k} q_{n-2l} w_{LH,k,l}^{(j+1,h)} + q_{m-2k} p_{n-2l} w_{HL,k,l}^{(j+1,v)} + q_{m-2k} q_{n-2l} w_{HH,k,l}^{(j+1,d)}) \quad (11)$$

図3に、表1でN=1でのウェーブレットを使用した場合の画像の多重解像度解析を示す。

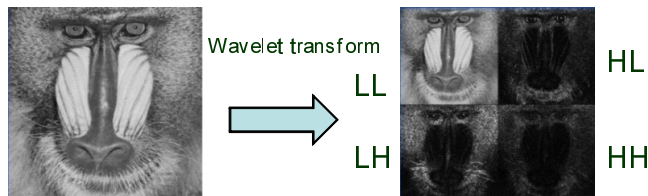


図3 ウェーブレット変換

3. 学習画像データベース作成方法

図4に、超解像処理で用いる学習画像データベースの作成方法を示す。

図に示されている通り、学習画像をウェーブレット変換し、4つの成分(LL, LH, HL, HH)をパックにしてパッチ(小領域)に分解する。パッチは、対応している4つの成分の3x3の大きさの小領域を組み合わせたものである。これを用いて学習画像データベースを作成する。本研究ではさらに、画像データベースに超解像したい画像を加える。これは、ある画像とその縮小画像をそれぞれパッチに分解したときに、ある画像のパッチとその縮小画像のパッチの間で似ているものが存在するというパッチの再起性[6]に基づいている。このことから、学習画像データベースだけを用いて超解像するよりも、さらに入力画像を学習画像データベースに加えて超解像した方が効果が高いと考えられる。

超解像したい入力画像をデータベースに加える際には、単純にそれ1枚をパッチに分解して学習画像データベースに加えるだけではない。まず、入力画像をウェーブレット変換し、パッチに分解する。さらにウェーブレット変換するときが発生する1/2の大きさに縮小された入力画像も同様にして、パッチに分解する。これを1/2画像、1/4画像、1/8画像…と繰り返し、パッチに分解出来なくなるほど小さくなるまで繰り返す。さらにこの処理を入力画像自体を0.99倍に縮小していきながら繰り返す。

この処理を行うことで単純に入力画像をパッチに分解してデータベースに加えるよりも、入力画像に含まれている情報をできるだけ有効に利用することができる。

4. 超解像システム

提案手法の流れを図5に示す。従来の手法は、画像の低解像度パッチと高解像度のパッチをペアにしてデータベースを作成しておき、入力画像が与えられると、それを分解した低解像度パッチとデータベースを比較し、高解像度パッチに置き換えることで高解像度の画像を得る手法である。これに対して、本研究では、学習画像データベースを作成するときに画像をウェーブレット変換し、それぞれの4つの成分(LL, LH, HL, HH)をパックにしてパッチに分解し、さらに入力画像もウェーブレット変換を用いてパッチに分解して学習画像データベースに加える。例えば、4つのパッチの次元数がそれぞれ9次元だとすると、用いるパッチの総次元数は36次元となる。超解像したい入力画像が与えられると、その入力画像をパッチに分解し、入力画像のパッチとデータベースのパッチのLL成分のみを比べ、最も似ているパッチの(LH, HL, HH)成分を置き換えていき、入力画像のパッチの失われた高周波成分を復元していく。その後、パッチに対して、高周波強調処理を行った後、パッチが重なっている部分を平均して重ね合わせ、逆ウェーブレット変換を行って一つの画像を得る。

5. パッチの選択方法

画像の形を決めているのは、ウェーブレット変換したときに発生するLL成分が大部分を占めているので、学習画像データベースにあるパッチのLL成分の画素と、入力画像をパッチに分解した画素部分とを比べて、最もよく似たパッチを検索する。パッチを評価する式は以下の通りである。

$$P(x_i, y_j) = \sum_k \frac{\|x_{ik} - y_{jk}\|^2}{2\sigma^2} \quad (12)$$

x_i は、学習データベースにあるi番目のパッチのLL成分であり、 x_{ik} はその中のk番目の画素の画素値である。 y_j は、入力画像のj番目のパッチであり、 y_{jk} はその中のk番目の画素の画素値である。 σ^2 は x_i と y_j の分散である。

この式の値が低いほど比べているパッチが似ているとして、評価が良いとする。学習画像データベースから評価が高いパッチを順番にN個検索する。次にそのパッチの高周波部分(LH, HL, HH)を入力画像の高周波成分として復元する。パッチをN個用いる理由は、パッチの高周波成分を強調するためである。

6. 高周波強調処理

画像の周波数が局在していることから、高周波成分を3乗するだけで画質が改善できる[9]という報告を基に、パッチを選んだ際に、最も評価が良いパッチだけでなく、それに準じて評価が良いパッチも高周波成分を復元する

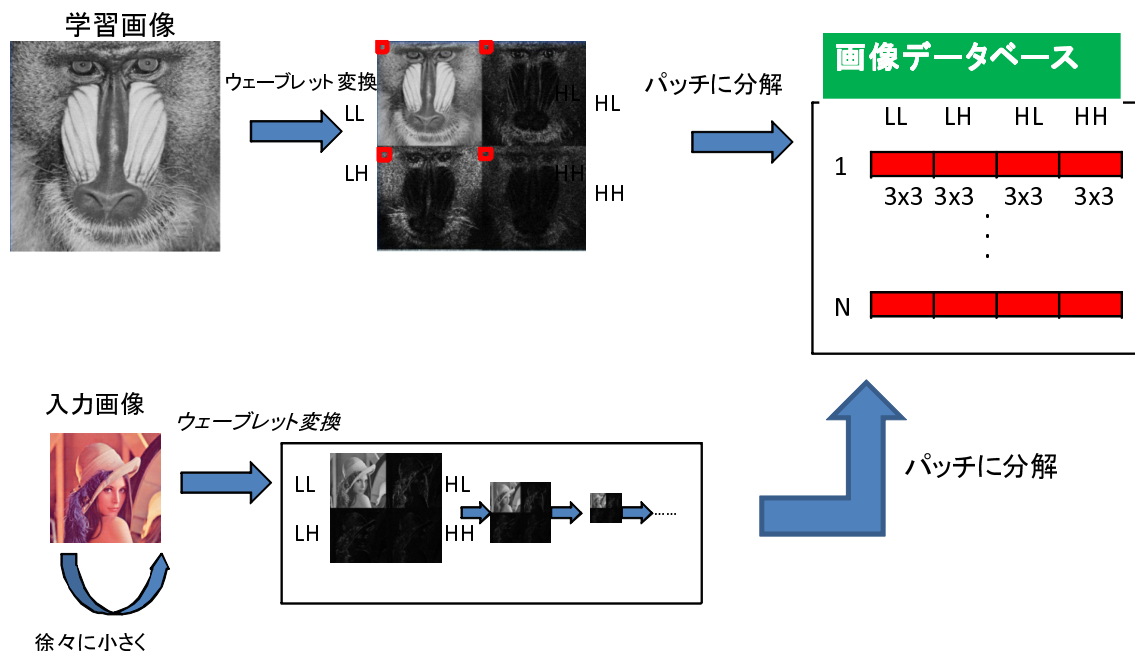


図 4 学習画像データベース作成方法

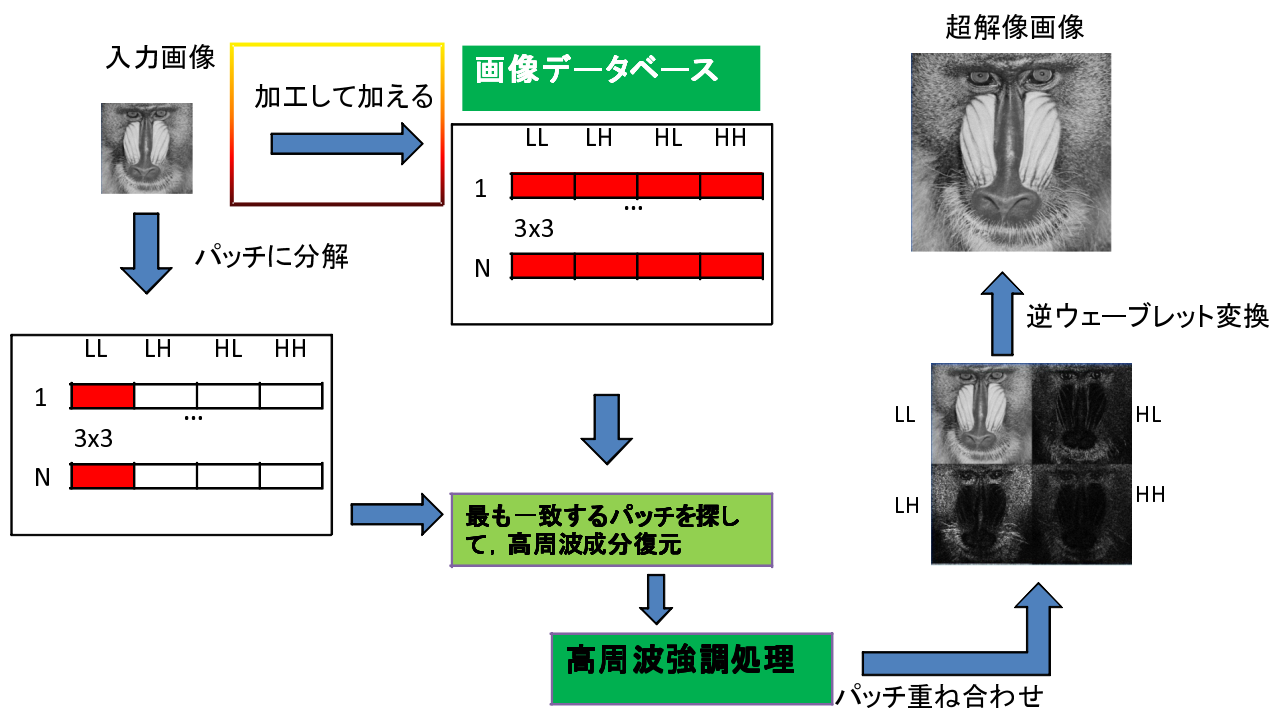


図 5 超解像システム

のに利用できると思われる。

そこで、評価が良い順に選ばれたパッチを $x_1, x_2, \dots, x_N$ とし、高周波成分 (LH, HL, HH) のパッチを $h_1, h_2, \dots, h_N$ として取り出す。高周波を強調したパッチを y とすると、 y は以下の式で表せる。

$$y = h_1 + \sum_{i=2}^N \frac{h_i}{(i-1)N} \quad (13)$$

N の値は任意であるが、実験により決定している。

7. 実験準備

7.1 入力画像

学習画像 (768x576) として 5 枚～14 枚用いる。入力画像の例 3 つを図 6 に示す。

7.2 評価方法

実験の評価方法として SSIM を用いる [10]。ここで SSIM について述べる。構造的類似性 (Structural SIMi-

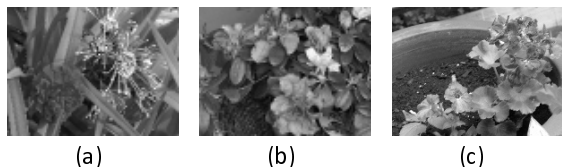


図6 入力画像

larity) 指数は、二つのイメージの類似性を計測する新しい指標である。SSIM は片方の画像が完璧な画質と考えた場合、比較対象となる画像の品質を測る指標と言う事ができる。SSIM の値が1 の値に近ければ近いほど、元の画像と似ている。PSNR と、比べると SSIM の方が人間の印象と一致しやすいという特徴がある。SSIM を求める式は以下の通り

$$SSIM(x, y) = \frac{(2\mu_x\mu_y)(2\sigma_{xy})}{(\mu_x^2 + \mu_y^2)(\sigma_x^2 + \sigma_y^2)} \quad (14)$$

x, y は比較する画像であり、 μ_x, μ_y は画像 x, y それぞれの全画素値の平均である。 σ_x^2, σ_y^2 は画像 x, y それぞれの全画素値の分散であり、 σ_{xy} は画像 x, y それぞれの全画素値の共分散である。

画像を評価する有名な方法として PSNR という指標がある。ここでは、なぜ PSNR ではなく SSIM を評価方法として用いる理由について述べる。比較のために、例を図7を示す。表7.2に図7の PSNR と SSIM を示す。

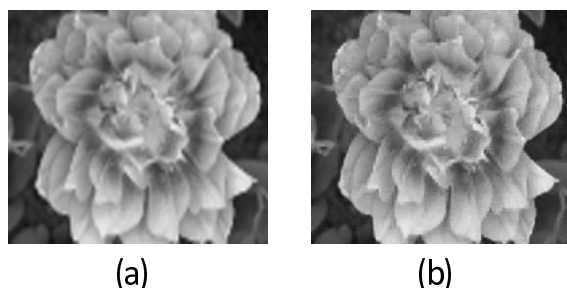


図7 画像比較

表2 図7に対する PSNR と SSIM の比較

	(a)	(b)
PSNR	39.990	39.917
SSIM	0.984	0.985

2 枚の画像を比較してみると、(b) の画像の方が (a) の画像よりも明らかに鮮やかなことがわかる。SSIM の精度を比較すると、(b) の方が良くなっているが、PSNR では、(a) の精度の方が良い。このことから、PSNR は人間の印象と必ずしも一致していないことがわかる。よって、SSIM の方が PSNR に比べ、人間の主観的な印象に一致していると考え、画像の指標として SSIM を使用することにした。

8. 評価実験

8.1 比較実験1

まず、学習画像データベースに入力画像を用いる場合と用いない場合を比較する。入力画像 (96x72) を縦横2倍に拡大する。(1) 学習画像のみの画像データベースを用いて拡大する手法と、(2) 学習画像に入力画像を組み合わせた画像データベースを用いて拡大する手法(3章)を比較した。グレースケールで行い、評価方法として SSIM を用いた。学習画像として (768x576) の画像を1枚~14枚用いた。拡大した画像は図6であり、SSIM における評価値を図8、図9、図10に示す。

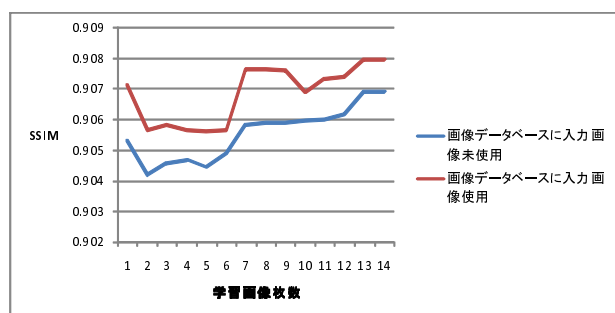


図8 画像(a)の SSIM 値

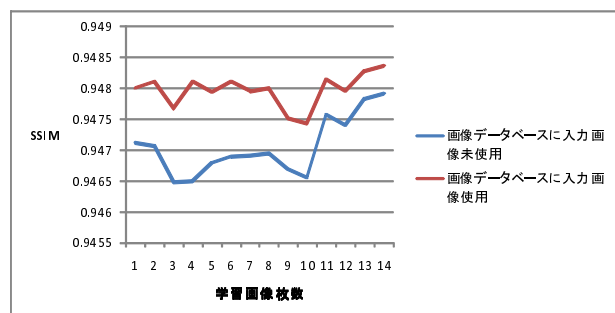


図9 画像(b)の SSIM 値

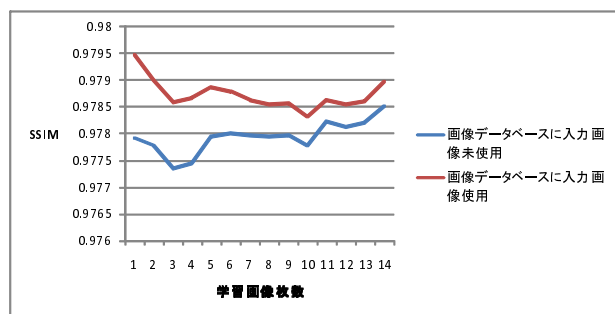


図10 画像(c)の SSIM 値

入力画像を学習画像データベースに組み合わせることにより、SSIM の値の上昇がみられ、学習画像に入力画像を組み合わせる手法の有効性が示された。SSIM の値の上昇値に差があるのは、学習画像のみを用いて超解像す

る際に使うパッチに加えて、入力画像を加工して得られたパッチで補完出来ているか否かの差が、表れているためである。

8.2 パラメータ設定

次に、第6章の高周波強調処理において、最も SSIM が高くなる N の値を実験により求める。図6の入力画像に対して、提案手法により縦横2倍に拡大し、 $N = 1, 2, \dots, 10$ の範囲で最も効果の高い N を求める。

実験の結果を図11に示す。

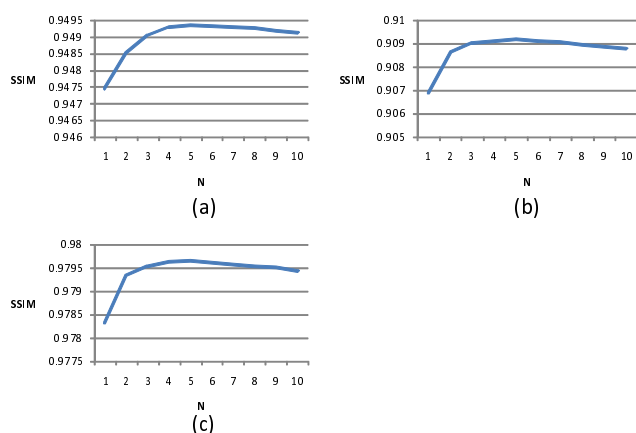


図11 N のパラメータ比較

足し合わせるパッチの数を増やしていくと、SSIM の値が上昇していることが分かる。これは、足し合わせていくことで高周波成分が強調されているためである。(a), (b), (c) は $N=5$ で SSIM の値が最大となった。 N の値が大きくなると、SSIM の値が減少していく。この理由は、選ばれるパッチの一致度が下がっていき、選ばれたパッチの高周波成分が不適切になったためと考えられる。この実験の結果より、今後の高周波強調処理は、 $N=5$ で行う。

8.3 比較実験2

最後に、入力画像を用いることと高周波強調処理により、どれだけ精度が上昇したか調べる。(1)Bicubic 法で拡大した画像と、(2) 学習画像のみの画像データベースを用いて拡大する手法、(3) 提案手法、3つの手法により拡大した画像を比較した。縦横2倍に拡大し、グレースケールで行った。評価方法として SSIM を用いた。学習画像として10枚用いる。

実験の結果を図12に示す。(a-1), (a-2), (a-3) は、(a) の入力画像に対し (1), (2), (3) による処理を行った結果である。(b-1), (b-2), (b-3) は、(b) の入力画像に対し (1), (2), (3) による処理を行った結果である。(c-1), (c-2), (c-3) は、(c) の入力画像に対し (1), (2), (3) による処理を行った結果である。

次に評価実験を行った。学習画像を10枚用いて、11.1で述べた手法(1), (2), (3)を行い、SSIM の評価値を得

た。結果を図13に示す。

8.4 実験考察

bicubic 法と提案手法を比べてみると、SSIM の値が上昇し、画像を比較してみると明らかに画像も鮮明になっている。これは画像データベースを用いることにより、高周波を復元することができ、その結果、SSIM 値が上昇したと考えられる。学習画像のみの画像データベースを用いた超解像手法と提案手法を比べてみると、エッジがわずかではあるが鮮明となっており、SSIM の値も少し上昇している。そして、例えば入力画像図6(a)の場合、図8のグラフを見ると、入力画像を用いていない実験で、学習画像が14枚の場合と、入力画像を用いた実験で、学習画像が1枚の場合の SSIM の値がほぼ同じである。これより、入力画像1枚分の情報量が学習画像13枚分と同じであることが分かる。図8, 図9, 図10のグラフより、学習画像枚数が増えるにつれて、入力画像を用いた場合と用いていない場合の差が徐々に小さくなっている。よって、入力画像を加工してパッチとして用いることは、学習画像が少ない場合、効果的な方法である。

9. まとめ

従来手法のウェーブレット変換を用いた超解像手法において、入力画像を利用すること、似た性質のパッチを重ね合わせることで高周波成分を強調させる超解像手法を提案した。拡大したい入力画像を学習画像として用いることで、少ない学習画像データであっても、超解像に有効な情報として活用することができた。そして、高周波成分を強調することで、画像のエッジを鮮明にすることができた。今後の課題は、高周波を強調させる処理の改良と、入力画像の情報をより利用できるように改善することである。

文 献

- [1] Freeman, W. T., Jones, T. R., Pasztor, E. C. : Example-Based Super-Resolution, IEEE Computer Graphics and Applications 22(2), 56-65 (2002)
- [2] Shu-Fan Lui, Jin-Yi Wu, Hsi-Shu Mao, and Jenn-Jier James Lien. :Learning-Based Super-Resolution System Using Single, ACCV 2007, Part II, LNCS 4844, pp. 96-105, 2007. Facial Image and Multi-resolution Wavelet Synthesis.
- [3] Kwang In Kim, Matthias O. Franz, and Bernhard Scholkopf. :Kernel Hebbian Algorithm for Single-Frame Super-Resolution, ECCV 2004 Workshop, Prague, Czech Republic, May 2004, pp. 135-149.
- [4] Dmitry Datsenko, Michael Elad. :Example-based single document image super-resolution: a global MAP approach with outlier rejection, Multidim Syst Sign Process (2007) 18:103-121
- [5] G. Dedeoglu, T. Kanade, J. August. :High-zoom video hallucination by exploiting spatio-temporal regularities. Proceedings of the IEEE Conference on CVPR, vol. 2, June 2004, pp. 151?158.
- [6] Daniel Glancer, Shai Bagon, Michal Irani: Super-

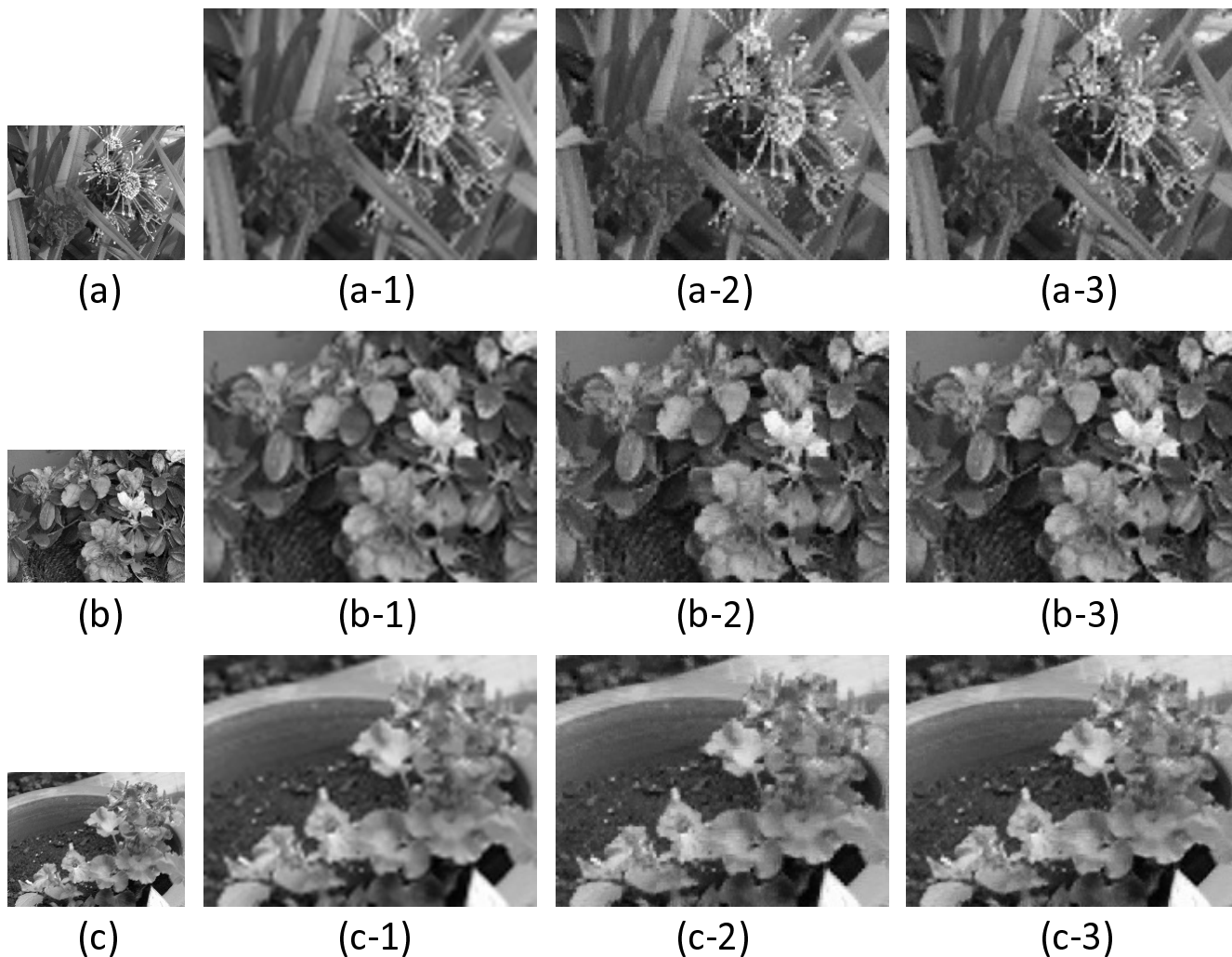


図 12 画像の比較

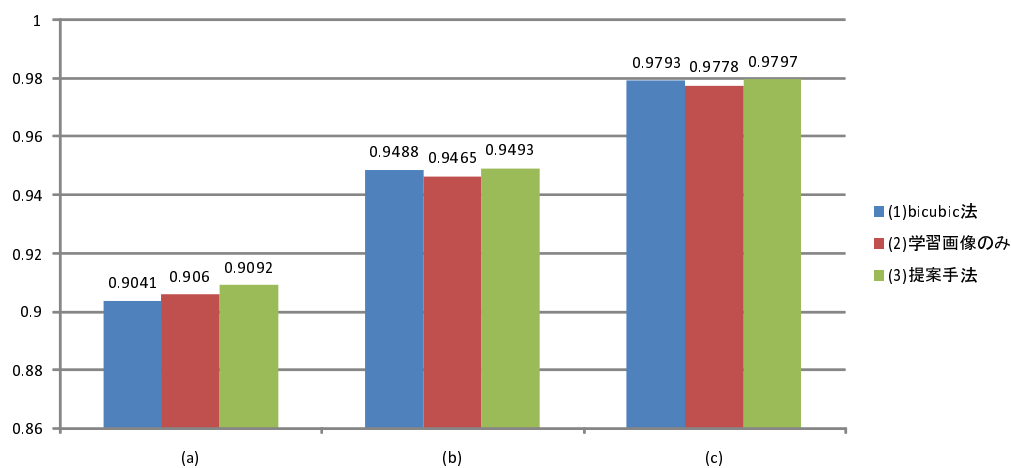


図 13 SSIM 値の比較

- Resolution from a Single Image, In ICCV (2009)
- [7] Freeman, W. T., Pasztor, E. C., Carmichael, O. T. : Learning Low-Level Vision, IJCV 40(1), 25-47 (2000)
- [8] 中野宏毅, 山本鎮男, 吉田靖夫, “ ウェーブレットによる信号処理と画像処理 ” 共立出版株式会社, 1999
- [9] 合志 清一, 寺川 雅嗣, 三上 浩, 今井 繁規:非線形特性

- を応用した画像の超解像度化, FIT2009 (第 8 回情報科学技術フォーラム)
- [10] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli: Image quality assessment From error visibility to structural similarity, IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, (2004)