

増分符号を用いた隣接ブロックの適合度評価による 学習型超解像

小野利幸 田口安則 三田雄志 井田孝
(株) 東芝 研究開発センター

あらまし 学習型超解像は、入力画像をブロック分割し、各ブロックを高解像度化するための画像パッチを、事前に学習した辞書から適切に選択することで、高画質な高解像度画像を得る技術である。パッチの選択には、隣接ブロック間でのパッチの整合性を考慮することが重要である。この整合性を、パッチの輝度の増減の変化パターンの類似度で評価する手法を提案する。提案手法を用いることにより、従来よりも鮮鋭で主観的に高画質な高解像度画像を生成できることを実験により示す。

Example-Based Super-Resolution by Evaluating Compatibility of Neighboring Blocks with Increment Sign

Toshiyuki Ono Yasunori Taguchi Takeshi Mita Takashi Ida
Corporate Research and Development Center, Toshiba Corporation

Abstract Example-base super-resolution estimates a clear high-resolution image from an input low-resolution image by using previously learned patches that enhance the image quality. An input image is split into blocks and super-resolved by using patches blockwise. It is important to consider the compatibility of patches used for neighboring blocks when selecting the patches from the dictionary. In this paper, we propose a novel compatibility function that evaluates the patches based on the wave shape of its intensity. Our method creates sharp and visually preferable high-resolution images. It is verified by subjective assessment experiments.

1 はじめに

超解像は、取得した低解像度画像について、単純な線形補間などでは復元できない高周波成分を復元しつつ、解像度を向上させる技術である [1]。近年のハイビジョンテレビの普及に伴い、従来品質のSD(Standard definition) 解像度での放送や、過去に撮影したビデオ映像などをテレビの解像度に合わせて高解像度化して見たいという要求がある。また、フルハイビジョンを超える解像度を持つテレビが実用化されることも予想され、映像資産の高解像度化技術への要求は高まると考えられる。

超解像手法の一つとして、学習データを用いて1枚の入力低解像度画像から高画質な高解像度画像を推定する学習型超解像がある。学習型超解像では、まず学習段階において、学習用高解像度画像とそれを仮定するモデルに従って擬似的に劣化させた低解像度画像の対から、その相関関係を学習して辞書を作成する。そして、超解像段階では、辞書を用いて、入力された低解像度画像を高解像度化する [2, 3]。特に Freeman ら [3] は、画像をブロック分割して、各ブロックを高解像度化するために補う高周波成分を示す画像パッチを特徴量とし、パッチの入力画像に対する適合度を表す生起エネルギーと、隣接するブロックに割り当てられたパッチとの適合度を表す隣接エネルギーを計算し、2つのエネルギーの重み付き和を最小化するパッチを辞書から選択して画像を高解像度化する手法を示した。

しかし、従来型の隣接エネルギーを用いた超解像では、輝度の変動が小さいパッチが選択されやすい傾向にあり、結果として鮮鋭な画像を生成することができなかった。これは、隣接エネルギーを、隣接するブロックに割り当てるパッチの重なり合う領域の輝度のSSD(Sum of Squared Differences)に基づいて算出していることによるもので、パッチの輝度の増減の変化パターン(波形)が似ていない場合にも、輝度の変動が小さいパッチ同士では、そのSSDが小さくなることが原因であった。

この問題を解決するため、パッチの輝度の増減の変化パターンの類似度に基づいた隣接エネルギーの計算手法を提案する。まず、隣接するパッチそれぞれについて、重なり合う領域の画素から増分符号ベクトルを生成する。そして、増分符号ベクトルの類似度に基づいて隣接エネルギーを計算する。このように計算した隣接エネルギーを用いてエネルギー最適化を行うことで、輝度の変動が小さいパッチが選ばれやすいという問題を解決し、高画質な画像を得ることができる。

本稿は、次のように構成されている。まず、2章で超解像問題を定式化し、Freeman らにより提案された従来手法の枠組みを説明する。次に、3章で、増分符号に基づいた新しい隣接エネルギーの計算手法を提案する。4章では、実験結果を示し、5章でまとめる。

2 従来手法

低解像度画像 Y は、高解像度画像 X が撮像システムによる光学的劣化や低解像度化を含む劣化 W を経て取得されるものとしてモデル化される [1]。

$$Y = WX + N \quad (1)$$

ただし、 N はノイズである。本稿では、 W を既知、 N を無視できるものとし、観測可能な Y から X を推定する超解像問題を扱う。

Freeman らは、この問題について、画像をブロック分割して各ブロックを高解像度化するための特徴量を辞書から選択する問題として学習型超解像問題を設定し、隣接するブロック間の共起関係を考慮したエネルギー最適化問題を解くことにより、自然な画像が得られることを示した [3]。以下では、学習型超解像の枠組みを図1を参照しながら説明する。

まず、学習段階では、学習用高解像度画像 (HR) とそれを擬似的に低解像度化させた低解像度画像 (LR) とから、低周波成分画像 (LF) と高周波成分画像 (HF) を生成する。LF は、LR を HR と同じ解像度へ単純な内挿法等を用いて拡大した仮拡大画像 (TH) から高周波成分を抽出して生成する。例えば、TH から LR の最近傍拡大画像を減算して生成する。また、HF は HR から TH を減算して生成する。ここで、HF は HR 中の高周波成分を多く含むという意味で高周波成分画像、LF は HF よりも低い周波数成分を多く含むという意味で低周波成分画像と呼んでいる。こうして得られた、低周波成分画像からブロック L を切り出し、ノルム $\|L\|$ で正規化して低周波成分パッチ l とする。次に、高周波成分画像について L と同じ位置からブロック H を切り出し、 $\|L\|$ で正規化して高周波成分パッチ h とする。そして、 l と h を組にして辞書に登録する。

超解像段階では、入力画像 (IN) から学習段階と同様にして低周波成分画像を生成し、入力低周波成分パッチ l_{IN} を得る。次に、 l_{IN} に対応する高周波成分パッチを複数選択する。例えば、 l_{IN} との距離 $\|l_{IN} - l_n\|^2$ が小さい l_n を複数探索し、その l_n と組になっている h_n を出力する。ここで選択された高周波成分パッチ h は、 $\|L_{IN}\|$ による逆正規化 (式 (2)) を行い H_n とする。

$$H_n = \|L_{IN}\| h_n \quad (2)$$

ここまでの処理で、仮高解像度画像の各ブロックには高画質化のための H_n が複数割り当てられた状態になっており、それらの候補の中から最適な H を選択して、仮高解像度画像へ加算することで、高画質な高解像度画像 (SR) が得られる。このとき、 H_n を対応付けるブロックは、仮高解像度画像上で、重なりがあるよ

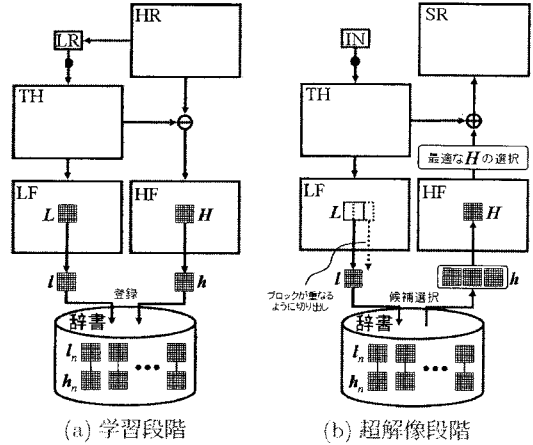


図 1: 学習段階と超解像段階の手順

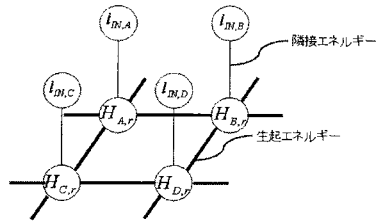


図 2: マルコフネットワークの構成

うに配置する。以降、ブロック A に割り当てられた R 個の高周波成分画像パッチを $H_{A,r} (r = 1, 2, \dots, R)$ と表す。選択する H の最適化は、以下のように行う。まず、図2のようなマルコフネットワークを考える。各ノードは、それぞれ1つの画像ブロックに対応しており、平面状に張られた枝はブロックの隣接関係を、垂直方向に張られた枝は入力画像と高周波成分画像パッチとの関係を表している。そして、マルコフネットワークの枝に対してエネルギー関数を割り当て、ネットワークの総エネルギーを最適化することにより、最適な H を選択する。まず、 l と H を結ぶ枝 (垂直方向に張られた枝) に対して、入力された低周波成分画像パッチ $l_{IN,A}$ に対する高周波成分画像パッチ $H_{A,r}$ の適合度 $D_{A,r}$ を次式で定義する。なお、以降、この適合度を生起エネルギーと呼ぶこととする。

$$D_{A,r} := \|l_{dic,A,r} - l_{IN,A}\|^2 \quad (3)$$

ただし、 $l_{dic,A,r}$ は、辞書中で $h_{A,r}$ と組になっていた低周波成分画像パッチである。また、 H 同士を結ぶ枝 (平面状に張られた枝) に対して、隣接する高周波成分画像パッチの適合度 $V_{A,rA,B,rB}^{SSD}$ (例えばエッジの連続性) を式 (4) で定義する。式 (4) は、パッチの重なり合う領域の輝度の2乗和 (SSD) である。この

適合度を、以降、隣接エネルギーと呼ぶこととする。

$$V_{A,r_A,B,r_B}^{\text{SSD}} := \sum_{k=1}^{K_{ol}} (H_{A,r_A}(p_k) - H_{B,r_B}(q_k))^2 \quad (4)$$

ただし、 $H_{A,r_A}(p_k)$, $H_{B,r_B}(q_k)$ は、 H_{A,r_A} と H_{B,r_B} が重なったときに、重なり領域内で k 番目の同一画素位置に配置される画素であり、 K_{ol} は重なる領域の画素数である。生起エネルギー、隣接エネルギー共に、値が小さいほど高周波成分画像パッチの適合度が高くなる。以上のように各枝で定義されるエネルギーを用いて、マルコフネットワーク全体のエネルギーを、式 (5) で定義する。

$$E := \lambda_D \sum_{A \in \mathcal{G}} D_{A,r} + \lambda_V \sum_{\{A,B\} \in \mathcal{I}} V_{A,r_A,B,r_B}^{\text{SSD}} \quad (5)$$

ただし、 $\lambda_D, \lambda_V \geq 0$, $\mathcal{G}$ は全ノードの集合、 $\mathcal{I}$ は隣接するノードのペアの集合である。式 (5) を最小化する高周波成分画像パッチの組合せ $\{H_{A,\hat{r}_A}\}_{A \in \mathcal{G}}$ を求めることで、周囲のパッチとの整合性をとりつつブロック毎に画像を高画質化する高周波成分画像パッチを選択することができる。なお、 $\{H_{A,\hat{r}_A}\}_{A \in \mathcal{G}}$ は、Belief Propagation を用いて、反復的に近似解を算出可能である。

最後に、 $H_{A,\hat{r}_A}$ を仮高解像度画像 (TH) へ加算することで、高画質な高解像度画像 (SR) を生成する。このとき、加算はブロックの重なりに応じて平均化する。

なお、従来手法で用いられていた辞書の構造では、高い画質を得るためには大量のサンプルを辞書に保存しておく必要がある。しかし、メモリ消費量や、超解像段階での低周波成分画像パッチの探索にかかる計算時間が増大するため、実用的な観点から辞書容量の縮小が求められる。この問題に対して、定められた辞書容量で学習誤差を最小化する辞書の作成方法が [4] で示されている。

3 増分符号を用いた隣接エネルギー

3.1 従来型隣接エネルギーの問題点

高周波成分画像パッチの重なり領域の輝度の SSD によって定義される隣接エネルギー (式 (4)) を用いた場合、エネルギー最適化により輝度の変動が小さいパッチが選択される傾向が強いことが実験的にわかった (4.2 節及び図 8 参照)。この原因の 1 つとして考えられる要因を図 3 を用いて説明する。図 3 は、重なり合う 2 つの高周波成分画像パッチの輝度の例を、簡単のために 1 次元で示した図である。図 3(a) は、2 つのパッチの輝度の増減の変化パターンは類似しており、また、輝度の変動量も大きい。一方、図 3(b) は、(a) と比較して変化パターンは類似しておらず、また、輝度の変動量は小さい。(a) と (b) の 2 つ

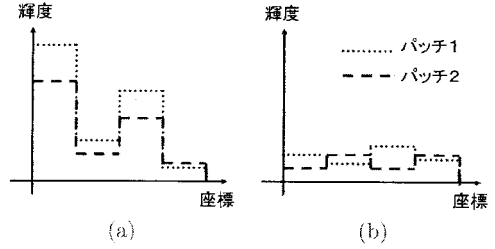


図 3: 輝度の増減の変化パターンの類似度と SSD 値の非類似性

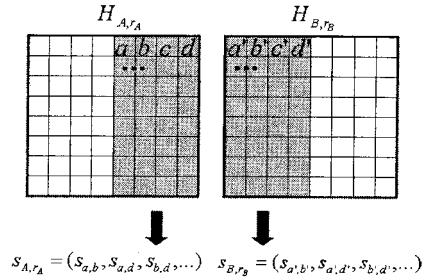


図 4: 増分符号の作成方法

の例を比較した場合、(a) はパターンが類似しているが輝度の開きが大きいので SSD の値が大きくなる。一方、(b) はパターンは類似していないにもかかわらず輝度の開きが小さいため SSD の値が小さくなる。つまり、(b) のような、輝度の変動量が小さいパッチ同士の組合せでは、パッチの輝度の増減の変化パターンが類似していない場合にも、隣接エネルギーが小さく評価される。その結果、エネルギー最適化により選択されるパッチは、輝度の変化が小さいパッチが多い傾向にあったと考えられる。

3.2 増分符号を用いた隣接エネルギーの計算

3.1 節で述べた問題点を解決する手法として、隣接エネルギーをパッチの輝度の増減の変化パターンの一致度に基づいて算出する手法を提案する。図 4 は、サイズ 8×8 画素のパッチ H_{A,r_A} と H_{B,r_B} が、左右方向に 4 画素ずれて重なる場合について、隣接エネルギーを算出する例を示している。まず、 H_{A,r_A} , H_{B,r_B} のそれぞれについて、重なる領域 (灰色で示した領域) の画素から増分符号ベクトル s_{A,r_A} , s_{B,r_B} を生成する。増分符号ベクトルは、重なる領域の画素の全ての組合せについて増分符号を求め、それを成分とするように生成する。つまり、重なる領域の画素数が K_{ol} (図 4 では 32) のとき、増分符号ベクトルの次元数 (Dim とする) は $K_{ol} C_2$ となる。ただし、 s_{A,r_A} と s_{B,r_B} の各成分は、重なる領域で同一位置の画素の組合せから作成する。2 つの画素 a と b の増分符号

$s_{a,b}$ は、

$$s_{a,b} = \begin{cases} 2 & \text{if } I_a - I_b \geq \sigma \\ 1 & \text{if } -\sigma < I_a - I_b < \sigma \\ 0 & \text{if } I_a - I_b \leq -\sigma \end{cases} \quad (6)$$

により計算する。ただし、 I_a, I_b はそれぞれ画素 a, b の輝度、 $\sigma \geq 0$ は閾値である。

そして、新しい隣接エネルギー $V_{A,rA,B,rB}^{prop}$ を式 (7) で定義する。

$$V_{A,rA,B,rB}^{prop} := V_{A,rA,B,rB}^{IS} + \mu V_{A,rA,B,rB}^{SSD} \quad (7)$$

ただし、 $V_{A,rA,B,rB}^{IS}$ は、2つの増分符号ベクトルの非類似度であり、次式 (8) で定義する。

$$V_{A,rA,B,rB}^{IS} := \sum_{k=1}^{Dim} |s_{A,rA}(k) - s_{B,rB}(k)| \quad (8)$$

式 (7) は、第1項でパッチの輝度のパターンの類似度を評価し、第2項でパッチの輝度の差を評価している。このように定義した隣接エネルギーを用いて、最適化するエネルギー関数を式 (9) で再定義し、エネルギー最適化を行うことにより、鮮鋭な画像を生成することが可能となる。

$$E := \lambda_D \sum_{A \in \mathcal{G}} D_{A,r} + \lambda_V \sum_{\{A,B\} \in \mathcal{I}} V_{A,rA,B,rB}^{prop} \quad (9)$$

4 実験

4.1 方法

提案手法の効果を確認するため、次のような条件で実験を行った。まず、図5に示した4枚のテスト画像 (サイズは 900×600 及び 760×436) について、面積平均により縦横それぞれ $1/4$ 倍に縮小して擬似的に入力低解像度画像を作成した。その入力画像を縦横それぞれ4倍に拡大する実験を行った。低周波成分画像パッチと高周波成分画像パッチのサイズは、いずれも 8×8 画素とした。ただし、低周波成分画像パッチについては主成分分析を用いて9次元に次元圧縮した。

辞書は、まず、大量のサンプルを保持した辞書として、従来手法の構造で3000万組の対を記憶した辞書を準備した。この辞書を辞書Iとする。また、低容量化を図った辞書IIを、次のように作成した [4]。まず、学習用画像から1億組の低周波成分画像パッチと高周波成分画像パッチの対を作成した後、低周波成分画像パッチ10万個を生成し、それぞれの低周波成分パッチに対して16個の高周波成分画像パッチを対応付けて作成した。

超解像段階では、仮高解像度画像に対して 8×8 画素のブロックを、縦横4画素ずつ重なるように配置

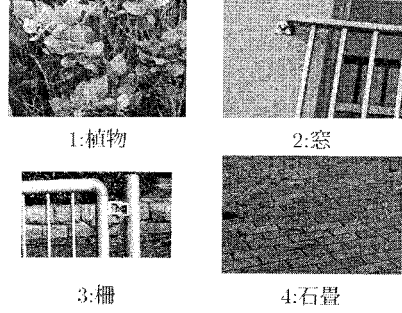


図5: テスト画像

した。辞書Iを用いる場合は、1つのブロックに対して、辞書から16個の低周波成分画像パッチを探索し、それらに対応する高周波成分画像パッチ16個を候補として割り当てた。辞書IIを用いる場合は、1つのブロックに対して、辞書から1個の低周波成分画像パッチを探索し、それに対応する16個の高周波成分画像パッチを高画質化するパッチの候補として割り当てた。

なお、辞書学習段階と超解像段階で用いる、仮高解像度画像の作成には、再構成制約を満足する高解像度化フィルタ [5] を用いた。このフィルタは、三次畳込み内挿法をベースとして、主にテクスチャ領域を鮮鋭化する効果があるフィルタであり、超解像結果画像の画質が底上げされる。また、超解像画像を縮小すると入力画像に一致するという再構成制約 (式 (1)) が保持される。

提案型の隣接エネルギーの計算に用いる増分符号ベクトルの次元数は、ブロックの重なりがそれぞれ32画素であるから、 ${}_{32}C_2 = 496$ 次元となる。また、係数パラメタは $\mu = 0.1$ とした。従来型の隣接エネルギーの計算は、式 (4) により行った。

隣接エネルギーの計算方法の差による結果をより明らかにするため、ネットワーク全体のエネルギーを定義する際の係数パラメタは、 $\lambda_D = 0, \lambda_V = 1$ とし、隣接エネルギーのみの最適化を行った。生起エネルギーは最適化されないが、 H の候補をブロックに割り当てる段階で、入力画像への適合度が評価されている。なお、Belief Propagationの反復回数は画質に大きな変化が無くなる10回とした。

以上の条件で、辞書IIを用いて従来型の隣接エネルギーの最適化を行う手法をDicII-SSD、提案型の隣接エネルギーの最適化を行う手法をDicII-Propとする。また、辞書Iの容量は、パッチの輝度をdouble型で保持した場合に17GBを超えるため、実用には現実的でないが、参考のため、Freemanの枠組みを再現する手法として、辞書Iを用いて従来型の隣接エネルギーの最適化を行う実験も行い、これをDicI-SSDとする。また、一般的な画像拡大手法として、三次

畳込み内挿法 (CC) を準備した。以上 4 手法について、主観評価実験と、PSNR による客観評価とにより結果画像の比較を行った。

主観評価実験は、一対比較法を用いて、「どちらがきれいですか？」という質問で、15 名の被験者に対して行い、サーストン法 [6] による主観評価値で結果を評価した。

4.2 結果

図 7 は、「1:植物」の結果画像の一部を切り出した比較である。葉や茎の輪郭を比較すると、CC や DicII-SSD の結果はぼやけており、また DicI-SSD はエッジ部にジャギーがあるのに対して、DicII-Prop では鮮鋭でジャギーの少ない画像が得られることがわかる。図 8 は、DicII-SSD と DicII-Prop により推定された高周波成分画像と、原画像から仮拡大画像を減算して求めた Ground truth との比較である。差をわかりやすくするため、振幅を 3 倍に拡大して示している。3.1 節でも触れたように、従来手法では輝度の変動が小さい高周波成分画像パッチが選択されるために、推定される高周波成分の強度が弱くなってしまふ。一方、提案手法では、高周波成分の強度が弱くなる問題を改善している。

図 6 に主観評価実験の結果を示す。値が大きいほど、評価が良い。DicII-Prop と DicII-SSD との比較では、DicII-Prop が全ての画像で優位であることから、提案した隣接エネルギーの有効性を確認することができる。既存手法との比較としては、DicI-SSD と DicII-SSD の比較からわかるように、従来型の隣接エネルギーを用いた場合、辞書容量を小さくしたことで画質が低下してしまう。しかし、提案型の隣接エネルギーを用いることで、画質を維持・向上できることが、DicI-SSD と DicII-Prop の比較からわかる。紙面の都合上割愛したが、辞書 I を用いて、提案型の隣接エネルギーを適用した場合にも、同じく画像の鮮鋭化を確認することができた。つまり、いかなる辞書を用いた場合でも、提案型の隣接エネルギーを用いることで、画像の鮮鋭化の効果が期待できると言える。

表 1 に各テスト画像に各手法を適応した結果画像の PSNR を示す。PSNR の比較では、テスト画像に依存して各手法間の優劣にばらつきがあり、また差も小さく有意な差は認められなかった。これは、隣接エネルギーに原画像への近さを評価する尺度が含まれていないためであると考えられる。

5 おわりに

隣接するブロックに割り当てられたパッチの適合度を示す隣接エネルギーについて、増分符号を用いてパッチの輝度の増減の変化パターンに基づいて適合度を評価する手法を提案した。提案した隣接エネ

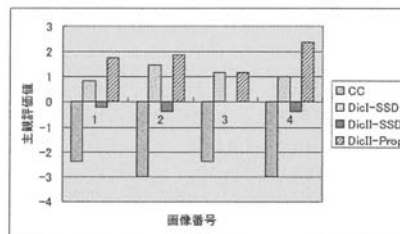


図 6: サーストン法による主観評価値

表 1: PSNR[dB] の比較

画像番号	CC	DicI-SSD	DicII-SSD	DicII-Prop
1	29.04	29.97	29.74	29.83
2	32.52	35.18	34.21	34.79
3	25.43	25.88	25.93	25.80
4	26.90	27.45	27.27	27.59

ルギーを用いた超解像により、従来型の隣接エネルギーを用いた場合と比較して、鮮鋭でジャギーの少ない画像が得られることを確認した。また、主観画質評価実験を行い、提案手法が視覚的に好ましい画像を生成することを確認した。

参考文献

- [1] S. C. Park, M. K. Park, and M. G. Kang, "Super-resolution image reconstruction - a technical overview-," *IEEE Trans. on Signal Processing*, vol. 20, no. 3, pp. 21–36, May 2003.
- [2] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1167–1183, Sept. 2002.
- [3] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, "Learning low-level vision," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 25–47, 2000.
- [4] 田口安則, 小野利幸, 三田雄志, 井田孝, "歪みを最小化する高周波成分の学習に基づく超解像度化方法," 情報処理学会研究報告, 2007-CVIM-160, vol. 2007, no. 87, pp. 257–264, 2007.
- [5] 田口安則, 井田孝, "再構成制約を満足する高解像度化フィルタ," PRMU 研究会, 3月, 2008 にて発表予定.
- [6] 菅民郎, "すべてがわかる アンケートデータの分析," 現代数学社, ISBN4-7687-0253-8, 1998.

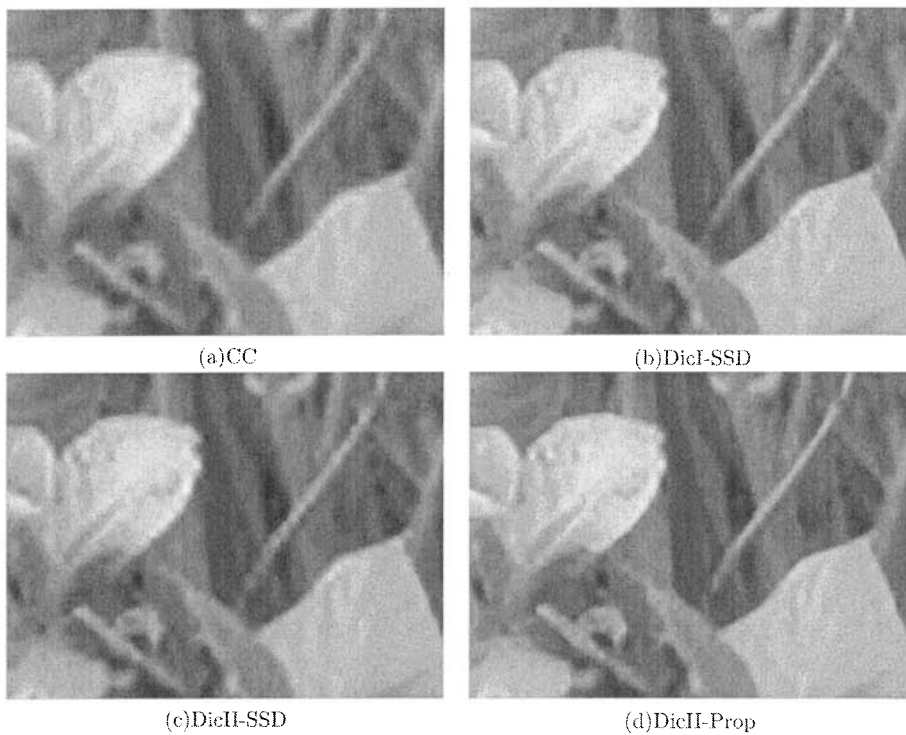


図 7: 結果画像の比較

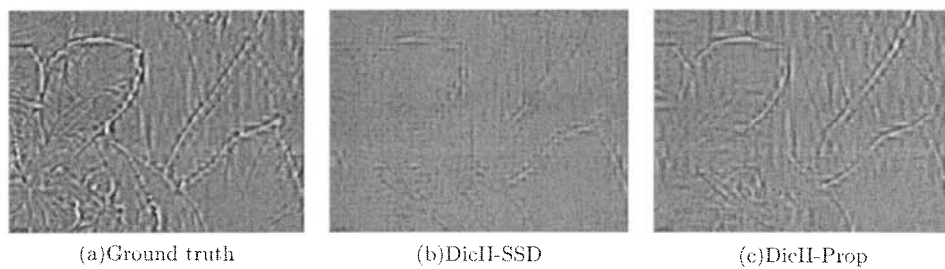


図 8: 高周波成分画像の比較