

自己縮小画像と混合ガウス分布モデルを用いた超解像

小川祐樹[†] 堀貴博[†] 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,horitaka}@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp

あらまし 近年、超解像技術はコンピュータビジョンの分野において活発に研究されている。本稿では、混合正規分布 (GMM) を用いた変換関数による超解像を提案する。低解像度画像を高解像度画像に変換する変換関数を、入力画像と入力画像の自己縮小画像を用いた混合正規分布から作成する。入力画像をその変換関数に適用することによって、高解像度画像を得ることができる。さらに、混合正規分布だけでなく、PLS (Partial Least Squares) も用いた変換関数による超解像も提案する。また、入力画像だけを用いているので、従来手法のように大量の学習画像を必要としない。従来手法との比較を行った結果、提案手法 (GMMのみ, GMM+PLS) 共に、従来手法より評価値が優れ、より鮮明な画像を作成することができ、提案手法の有効性を確認した。

キーワード 超解像, 混合ガウス分布モデル, PLS 回帰分析

1. はじめに

近年、デジタルカメラや携帯電話に搭載されたカメラの解像度は、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、高画質化を図る技術が求められている。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、デジタルカメラに比べると解像度が低いため、写真として保存する際には、拡大処理を施して高解像度に変換する必要がある。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うため、低解像度画像になってしまう。また、低解像度画像から高解像度画像を求めることは、解が無数にあり不良設定問題である。そこで本研究では、画像を拡大する際に、適切な高周波成分を付加することによって、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復元する技術であると言える。超解像の手法には、動画を対象にする手法も存在するが本稿では、1枚画像を超解像する手法に焦点をあてる。

2. 関連研究と提案内容

超解像技術は、大きく次の二つに分類することができる。

- (1) Example-based 手法
- (2) 線形回帰による手法

Example-based 手法 [2] は、画像の低解像度パッチと高解像度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。

線形回帰による手法は、 $y = Lx$ (y :低解像度画像 x :高解像度画像 L :劣化関数) によりモデル化し、高解像度画像 x を求める手法である。この手法は L の逆行列を求めることが困難なため、様々な方法を用いてこのモデルを解く手法である。Sparse-coding を用いた手法 [3], Total-variation 正則化を用いた手法 [11], [12], BPLP [4] を用いた手法, Lucy-Richardson [5] を用いた手法などある。また、超解像ではないが、一般的で簡単な画像の拡大手法としてバイキュービック法がある。これらの手法を、どれだけこまやかに表現できるのかという表現力を縦軸に、どれだけ強くモデル化しているのかという軸を横軸にとると、図1のように表せる。

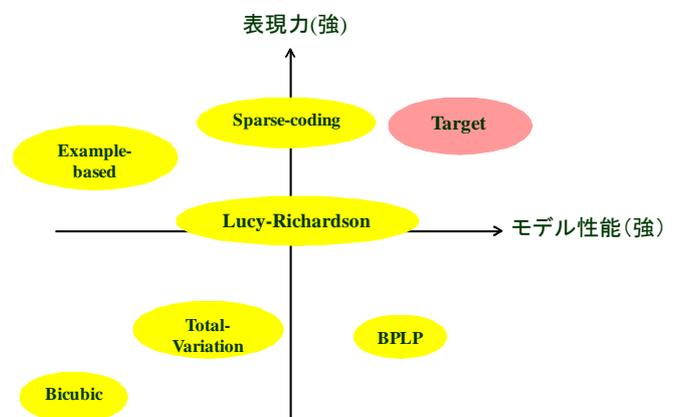


図1 関連研究の分布図

Example-based 手法 [2] は、画像の低解像度パッチと高解像度パッチをペアにしてデータベースを作成しておく、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り、大量の画像データベースがあれば、効果を発揮する。従って、表現力が強く、モデル性が低いと考えられる。

Lucy-Richardson を用いた手法 [5] は、画像のぼけを様々な方法で推定してから、事後確率を最大化する手法なので、中心に配置する。

BPLP を用いた手法 [4] は、低解像度画像から高解像度画像を作成するモデル式を用いるが、固有空間を用いる際に情報損失が起こるので、モデル性が高く表現力が低い位置に配置できる。

Sparse-coding を用いた手法 [3] は、Sparse-coding を用いることにより、情報の分解、再構成を効率よく行うことができるため、表現力が強い位置に配置できる。

Total-variation 正則化を用いた手法 [11], [12] は、ハイパーパラメータに結果が依存しがちであり、他の手法と組み合わせる必要があるため、モデル性と表現力が低い位置に配置している。

バイキュービック法は、単純な拡大手法であるため、1番左下に配置している。

本研究では、従来法では研究されていない領域、すなわち、モデル性が強く、表現力が強い手法を提案する。そのために、確率モデルとしての表現力が強い混合正規分布 (GMM) を用いた超解像を提案する。具体的には、学習段階では、入力画像とその縮小画像から GMM を作成し、これをもとに低解像度画像から超解像画像を生成する変換関数を推定する。従来手法のように大量の学習画像を用いることなく、入力画像のみで処理を行う。推定段階では、入力画像を変換関数に適用することで、超解像画像を得る。本稿では、GMM だけではなく PLS (Partial Least Squares) [1] も用いて変換関数を作成し、超解像を得る手法も提案する。

3. 混合正規分布を用いた画像変換法

混合正規分布を用いた画像変換法を説明する前に、混合正規分布 (GMM) について述べる。

3.1 GMM

一般に、パターン認識に用いられる特徴量は多次元ベクトルとなる。多次元ベクトルに対する多次元正規分布は、次元数を n 、データの特徴量ベクトルを $\mathbf{x} \in \mathbf{R}^n$ とすると次式で表される。

$$N(\mathbf{x}; \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)\right\} \quad (1)$$

ここで $\mu \in \mathbf{R}^n$ は平均ベクトル (mean vector), $\Sigma \in$

$\mathbf{R}^{n \times n}$ は分散共分散行列 (variance covariance matrix) と呼ばれる正定値対称行列である。

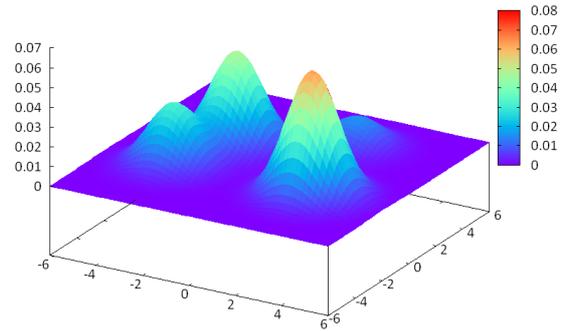


図2 GMM の例

正規分布は単一のピークを持つ単純な確率密度関数 (Probability Density Function, PDF) であり、複雑な分布を表現することはできない。そこで、複雑な分布を、複数の正規分布に対する重み付き和を用いて表現する。正規分布を重み付きで足し合わせた確率密度関数は、混合正規分布と呼ばれる。データの次元数を n 、正規分布の混合数を M 、特徴量ベクトルを $\mathbf{x} \in \mathbf{R}^n$ 、 i 番目の分布に属する平均ベクトルを $\mu_i \in \mathbf{R}^n$ 、分散共分散行列を $\Sigma_i \in \mathbf{R}^{n \times n}$ 、 i 番目の分布の重み係数を λ_i とする。各正規分布を $N(\mathbf{x}; \mu_i, \Sigma_i)$ とすると、GMM の出力確率 $b(x)$ は、

$$b(x) = \sum_{i=1}^M \lambda_i N(\mathbf{x}; \mu_i, \Sigma_i) \quad (2)$$

と表すことができる。ここで、重み係数 λ_i は、

$$\sum_{i=1}^M \lambda_i = 1 \quad (3)$$

である。

図2に、次元数2、混合数4のGMMの例を示す。このように、GMMは複雑な確率密度関数を表現ことができ、パラメータの推定もEMアルゴリズム (Expectation Maximization) を用いることで、比較的容易に行えるため、広い分野で用いられている。

3.2 変換関数

次に、GMMを用いて画像を変換する方法について述べる。元となる画像と目標となる画像の特徴量をそれぞれ、 $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$, $\mathbf{y} = [y_1, y_2, \dots, y_n]^T$ とする。元画像 x から目標画像 y へ変換する関数は、文献 [6] に述べられている。

$$y = F(x) = E[y|x]$$

$$= \sum_{i=1}^M w_i(x) [\mu_i^y + \Sigma_i^{yx} (\Sigma_i^{xx})^{-1} (x - \mu_i^x)] \quad (4)$$

$$w_i(x) = \frac{\alpha_i N(x; \mu_i^x, \Sigma_i^{xx})}{\sum_{j=1}^M \alpha_j N(x; \mu_j^x, \Sigma_j^{xx})} \quad (5)$$

μ_i^x と μ_i^y はそれぞれ、元画像と目標画像の分布 i における平均ベクトルである。 Σ_i^{xx} は、元画像分布 i における共分散行列であり、 Σ_i^{yx} は元画像と目標画像の分布 i における相互共分散行列である。式 (4) を用いることで、元画像を目標画像へと変換することができる。変換関数のパラメータ $(\alpha_i, \mu_i^x, \mu_i^y, \Sigma_i^{xx}, \Sigma_i^{yx})$ は、結合ベクトル $\mathbf{z} = [\mathbf{x}^T \mathbf{y}^T]^T$ の確率分布を、次式で示すような GMM で表すことで推定する。

$$P(\mathbf{z}) = \sum_{i=1}^M \alpha_i N(\mathbf{z}; \mu_i^z, \Sigma_i^z) \quad (6)$$

μ_i^z と Σ_i^z は、次のように表現できる。

$$\Sigma_i^z = \begin{bmatrix} \Sigma_i^{xx} & \Sigma_i^{xy} \\ \Sigma_i^{yx} & \Sigma_i^{yy} \end{bmatrix}, \mu_i^z = \begin{bmatrix} \mu_i^x \\ \mu_i^y \end{bmatrix} \quad (7)$$

パラメータの推定は EM アルゴリズムによって行う。

4. 提案する超解像システム

図 3 に提案する超解像システムの流れを示す。

学習段階では、低解像度画像から超解像度画像を作成する変換関数を、入力画像とその縮小画像を用いた GMM により作成する。従来手法のように大量の学習画像を用いることなく、入力画像のみで処理を行う。

推定段階では、入力画像を変換関数に適用することで、超解像画像を得る。

4.1 変換関数の学習

次の手順により、低解像度画像から目標画像への変換関数を求める。

(1) 低解像度画像と高解像度画像のペアを求める。

図 3 左に示すように、入力画像 I を一度縮小して自己縮小画像 I_R を得る。自己縮小画像 I_R をバイキュービック法で拡大することで、入力画像の低解像度画像 I_L を得る。これにより、入力画像の低解像度画像 I_L と高解像度画像 I のペアを得ることが出来る。

(2) 低解像度画像の画像特徴量を求める。

低解像度画像 I_L に複数のハイパスフィルタ ($H1, H2, H3, H4$) を適用することで、高周波成分 ($I_{LH1}, I_{LH2}, I_{LH3}, I_{LH4}$) を抽出する。このハイパスフィルタは、次のような水平、垂直方向の第 1 次微分成分を抽出するフィルタと、水平、垂直方向の第 2 次微分成分を抽出するフィルタである。

$$H1 = [0, 0, 1, 0, 0, -1]$$

$$H2 = [0, 0, 1, 0, 0, -1]^T$$

$$H3 = [1/2, 0, 0, -1, 0, 0, 1/2]$$

$$H4 = [1/2, 0, 0, -1, 0, 0, 1/2]^T$$

(3) 復元する目標画像の特徴量を求める。

自己縮小画像を拡大した低解像度画像 I_L と、高解像度画像である入力画像 I との差分をとることで、差分画像 I_F を得る。これは、高解像度画像が持つ高周波成分であり、低解像度画像では失われてしまった成分である。この特徴量が、復元すべき目標画像の特徴量となる。

(4) 低解像度画像から目標画像への変換関数を求める。

関連している低解像度の画像特徴量 $I_{LH1}, I_{LH2}, I_{LH3}, I_{LH4}$ と目標画像の特徴量 I_F を小さなパッチ (画像の小領域) に分解する。分解された画像特徴量は次式のように表せる。

$$\mathbf{x}_m = [I_{LH1m}, I_{LH2m}, I_{LH3m}, I_{LH4m}]^T \quad (8)$$

$$\mathbf{y}_m = [I_{Fm}]^T \quad (9)$$

m はパッチ番号である。結合ベクトル \mathbf{z} は、 \mathbf{x} と \mathbf{y} を結合することで得られる。結合ベクトルから GMM を作成し、式 (4) に適用することで、 \mathbf{x} から \mathbf{y} を生成するための変換関数 $y = f(x)$ を得る。

4.2 超解像画像の推定

次の手順により、低解像度画像から高解像度画像を生成する。

(1) 入力画像の拡大画像を求める。

図 3 右に示すように、入力画像 I をバイキュービック法で拡大することで拡大画像 I'_L を得る。拡大画像 I'_L に対して、複数のハイパスフィルタを適用して、高周波成分 $I_{LH1}, I_{LH2}, I_{LH3}, I_{LH4}$ を得る

(2) 変換関数により目的画像を生成する。

高周波成分をパッチに分解することで、分解された画像特徴量 \mathbf{x}_m を得る。目標画像 I^F (高解像度画像復元に必要な失われた高周波成分) を得るために、拡大画像 I'_L のパッチ \mathbf{x}_m を、式 (4) に適用して \mathbf{y}_m を得る。

(3) 目標画像より超解像画像を求める。

拡大画像 I'_L に、復元した高周波成分 I'_F を足し合わせることで、超解像画像 I_S を得る。

5. GMM と PLS を用いた超解像

GMM による超解像変換を更に改良した手法として、PLS (Partial Least Squares) [1] と GMM を組み合わせた手法を提案する。この手法は、本来、音声の研究分野において声質変換として研究されてきた手法を超解像に利用したものである。GMM のみで変換関数を作成すると、過学習が起こる可能性があり、これを防ぐために PLS を

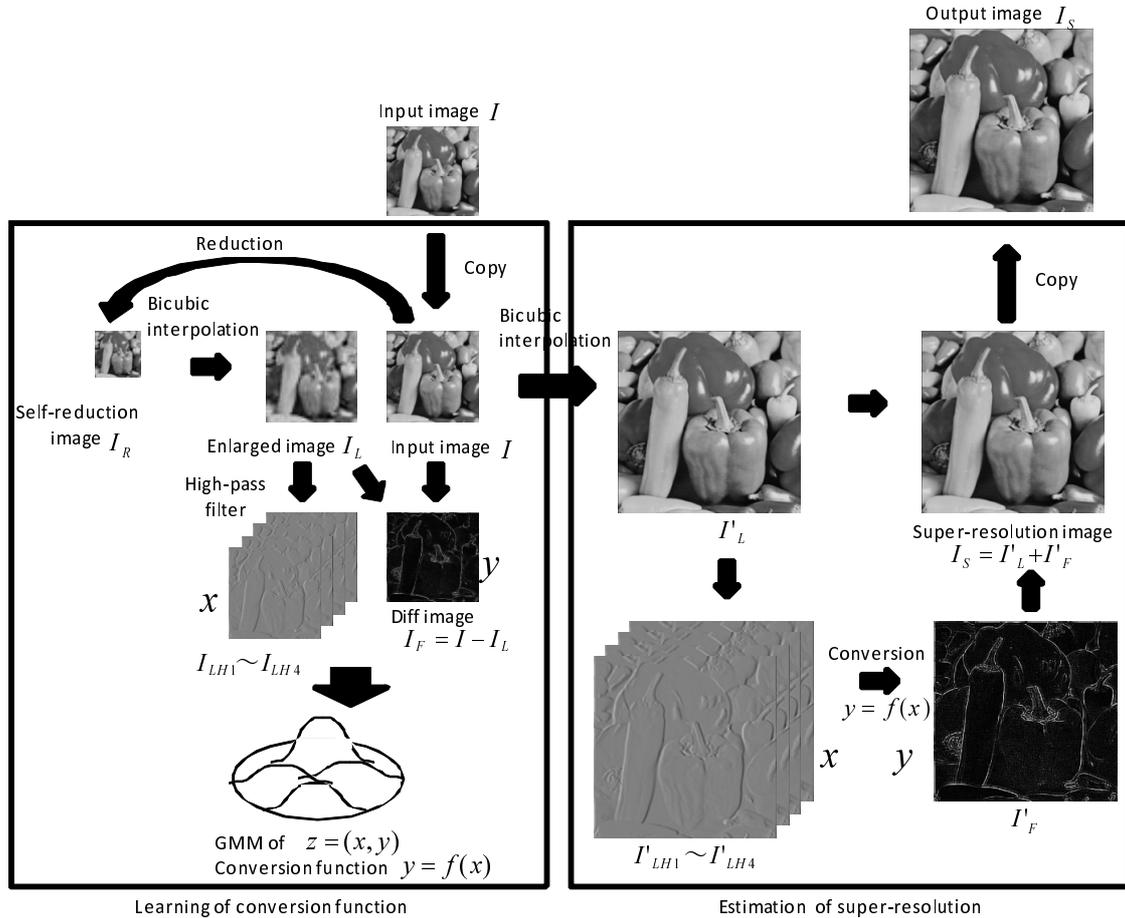


図3 超解像画像処理方法

利用している. まず, PLS について述べた後に, PLS と GMM を用いて超解像を行う方法について述べる.

5.1 PLS(Partial Least Squares)

PLS 法は, 計量化学の分野で Wold(1975) によって開発され, その分野でよく用いられている回帰分析手法である. PLS 回帰はデータをそのまま使わずに潜在変数を計算し, その潜在変数への回帰を行う点が通常の重回帰と異なる. 回帰の係数を求める際には, 潜在変数と従属変数の共分散が最大になるようにし, 潜在変数が互いに無相関となるよう求めていく手法である. PLS 回帰のアルゴリズムはいくつか提案されており, 本研究では SIMPLS アルゴリズム [8] を用いる.

行列 $X = [x_1, \dots, x_m, \dots, x_t]$, 目標行列 $Y = [y_1, \dots, y_m, \dots, y_t]$ を考える. x_m, y_m は行列 X, Y の要素ベクトルである. この x_m, y_m に対して PLS 回帰を行うと次式のようなになる.

$$y_m = \beta x_m + e \quad (10)$$

β は回帰係数であり, e は残差である.

次に, β を求めるアルゴリズムについて述べる.

- (1) R, V, Q, T を初期化して空の行列を用意する.
- (2) $C = XY^T$ を計算し, x と y の相互共分散を求

める.

- (3) $C^T C$ の最大固有値に対応した固有ベクトル q を求める.
- (4) $r = Cq, t = X^t r$ とする.
- (5) t の平均を抽出する.
- (6) $r = r/\|t\|, t = t/\|t\|$ により, r, t を正規化する.
- (7) $p = Xt, q = Yt, u = Y^T q$ とする.
- (8) $v = p$ とする.
- (9) 繰り返し回数 $i > 1$ ならば, $v = v - VV^T p, u = u - TT^T u$ を行う.
- (10) $v = v/\|v\|$ を行い, v の正規化を行う.
- (11) $C = C - vv^T C$ を行う.
- (12) r, q, v を i 番目の列ベクトルとして, R, Q, V に割り当てる.

これら (2)-(12) の処理を PLS の要素 $i = 1, 2, \dots$ の数だけ繰り返す. 要素数はユーザーが決めるハイパーパラメータである. これらの処理によって

$$\beta = RQ^T \quad (11)$$

を計算することで, β を求めることが出来る.

5.2 PLS を用いた超解像

超解像度画像を求める手順は, 4. で述べた手順および

図3に示した手順と同じである。異なっているのは、変換関数が式(4)から式(10)に変わることである。

学習段階では、入力画像とその縮小画像を用いてGMMとPLSを学習し、変換関数を生成する。従来手法のように大量の学習画像を用いず、入力画像のみで処理を行う。学習データは低解像度画像のパッチ x_m からなる行列 $X = [x_1, \dots, x_m, \dots, x_t]$ と、目標画像のパッチ y_m からなる行列 $Y = [y_1, \dots, y_m, \dots, y_t]$ である。

推定段階では、次のように入力画像を変換関数に適用することで、超解像画像を得る。まず、学習データのパッチ x_m から、式(5)により、重み $w_{i,m}$ を求めることで、推定したい目標画像のパッチ \hat{y}_m を、 $y_{i,m}$ の重み付け和で次式のように求める。

$$\hat{y}_m = \sum_{i=1}^M w_{i,m} \hat{y}_{i,m} \quad (12)$$

ここで、PLS回帰を用いることにより、式(12)は次式のように表せる。

$$\hat{y}_m = \sum_{i=1}^M w_{i,m} \beta_i x_m + e_m \quad (13)$$

SIMPLSは0平均の分散のデータを用いる。まず、低解像度画像のパッチ集合 $\{x_m\}$ 、目標画像のパッチ集合 $\{y_m\}$ の平均を求める。

$$\mu_x^i = \frac{\sum_{m=1}^T w_{i,m} x_m}{T} \quad (14)$$

$$\mu_y = \frac{\sum_{m=1}^T y_m}{T} \quad (15)$$

求めた平均を用いて、パッチ集合 $\{x_m\}$ 、 $\{y_m\}$ 共に0平均化を行う。

$$\tilde{y}_m = y_m - \mu_y \quad (16)$$

$$\tilde{x}_m = \begin{bmatrix} w_{1,m} x_m - \mu_1^x \\ w_{2,m} x_m - \mu_2^x \\ \vdots \\ w_{M,m} x_m - \mu_M^x \end{bmatrix} \quad (17)$$

パッチ集合を0平均化することにより、式(10)を次式のように書き直すことができる。

$$\tilde{y}_m = \beta \tilde{x}_m + e_m \quad (18)$$

$$\beta = [\beta_1, \beta_2, \dots, \beta_M] \quad (19)$$

式(18)に対してPLS回帰を行うことにより、 β を求める。 β の求め方は前節で述べた通りである。従って、目標画像のパッチ \hat{y}_m を求める変換関数は次式のように表せる。

$$\hat{y}_m = \beta \tilde{x}_m + \sum_{n=1}^M w_{i,m} \mu_n^y \quad (20)$$

$$\mu_n^y = \frac{\sum_{m=1}^T w_{i,m} y_m}{T} \quad (21)$$

6. 評価実験

実験で用いる入力画像を図4に示す。画像のサイズは576x576画素である。提案手法が様々な種類の画像に対応できるのか調べるために、入力画像として、テクスチャが複雑な画像、自然風景の画像、被写体がある画像、人工物の画像を用いて実験を行う。



図4 入力画像

6.1 評価方法

画像を客観的に評価する手法として様々な手法が提案されている。MSE(Mean Square Error)やPSNR(Peak Signal to Noise Ratio)は、従来から良く用いられている手法であるが、信号レベルの劣化度合いを計算するのみであり、その劣化が人の視覚特性にどのような影響を与えるかを考慮していない。そこで、人の視覚特性を考慮した客観的評価手法として、SSIM[9]やVSNR[10]が提案されている。

本研究では、より客観的に評価することを目的として、複数の手法を用いて評価を行う。実験で使用する評価手法は、PSNR, SSIM, VSNRである。

6.1.1 PSNR

PSNR(ピーク信号対雑音比)は、信号が取りうる最大のパワーと劣化をもたらすノイズの比率を表す工学用語である。多くの信号はダイナミックレンジが非常に広いので、PSNR比は通常10を底にした常用対数で表される。

PSNRは一般的に、画像圧縮や画像復元などで使用され、元画像と処理画像を用いて値を算出することで画像の品質を評価することができる。値が高ければ高いほど、処理画像が元画像に近い。ただし、信号レベルの劣化度合いを計算しているだけなので、人の視覚特性を考慮しておらず、人の評価とは一致しないことがある。

PSNRの定義として、共に、 $m \times n$ の大きさの元画像 I と、何らかの処理をして得た処理画像 J があるとき、MSEは次式で表せる。

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I_{i,j} - J_{i,j})^2 \quad (22)$$

MSEを用いてPSNRは次式で表せる。

$$PSNR = 10 \log_{10} \frac{MAX^2}{MSE} \quad (23)$$

MAX は、画像が取りうる最大ピクセル値である。ピクセルが1サンプルあたり8ビットで表現されている場合、MAX の値は255である。

6.1.2 SSIM [9]

SSIM について説明する。構造的類似性 (Structural SIMilarity : SSIM) 指数は、二つのイメージの類似性を計測する指標である。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定することができる。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ているという特徴がある。SSIM を求める式は以下の通りである。

$$SSIM(x, y) = \frac{(2\mu_x\mu_y)(2\sigma_{xy})}{(\mu_x^2 + \mu_y^2)(\sigma_x^2 + \sigma_y^2)} \quad (24)$$

x, y は画像、 μ_x, μ_y は x, y それぞれの全画素値の平均、 σ_x^2, σ_y^2 は x, y それぞれの全画素値の分散、 σ_{xy} は x, y それぞれの全画素値の共分散である。2枚の画像の平均と分散が等しく、共分散が画像の分散に等しい場合、 $SSIM = 1$ となる。

6.1.3 VSNR [10]

従来の PSNR は、単なる信号レベルの劣化量に基づいて計算されるが、VSNR はそれが人の目に届いた後に知覚されるであろう劣化量に基づいて計算される。そのため、その計算過程において人の視覚特性がモデル化されている。人の視覚には次の重要な特徴がある。

- (1) 画像のコントラストが大きいほど、その中の小さな雑音を検知しにくい。
- (2) 空間周波数が高い雑音ほど検知しにくい。

VSNR の計算過程では、画像 I のコントラスト $C(I)$ を計算する仕組み、および人の目が知覚する誤差 VD (Visual Distortion) を計算する仕組みを持つ。それらを用いて VSNR は次式のように表せる。

$$VSNR = 10 \log_{10} \left\{ \frac{C(I)}{VD} \right\}^2 \quad (25)$$

$C(I)$ と VD の導出については、文献 [10] で詳しく述べられている。

6.2 パラメータ決定のための実験

評価が最も高くなる画像のパッチサイズと混合数を決定する。実験に用いた画像は図4の画像(c)を代表として選び、この画像を基準にパラメータを決定する。画像の評価手法として PSNR, SSIM, VSNR を用いた。まず、混合数を10に固定し、パッチサイズの大きさを変更した。画像のパッチのサイズは、3x3, 6x6, 12x12, 18x18, 24x24, 30x30, 36x36 で実験を行った。結果を表1に示す。

表1より、画像のパッチサイズの大きさは、3x3 が適切であることがわかった。次にパッチサイズを3x3に固定した上で、混合数を3, 6, 9, 12, 15で実験を行った。

表1 パッチサイズの比較結果

パッチのサイズ	PSNR	SSIM	VSNR
3x3	38.92	0.9308	13.92
6x6	38.82	0.9293	13.92
12x12	38.55	0.9246	13.92
18x18	38.13	0.9152	13.92
24x24	37.53	0.9006	13.92
30x30	36.52	0.8711	13.90

表2 混合数の比較結果

混合数	PSNR	SSIM	VSNR
3	38.89	0.9305	13.92
6	38.91	0.9309	13.92
9	38.92	0.9308	13.92
12	39.00	0.9315	13.93
15	38.71	0.9295	13.92

結果を表2に示す。

混合数を変化させてもあまり変化は見受けられなかったが、混合数12の時に最も高い評価値が得られたため、混合数12で実験を行っている。

以上のことより、画像のパッチサイズは3x3、GMMの混合数は12で実験を行った。

6.3 比較実験

512x512画素の画像を1/4の大きさにして作成した256x256画素の画像を縦横2倍に拡大し、失われた高周波成分がどれだけ復元できたか調べる。比較した手法は次の5種類である。(1)Bicubic法、(2)Example-based手法[2]、(3)スパースコーディング[3]、(4)提案手法(GMMのみ)、(5)提案手法(GMM+PLS)。グレースケールで実験を行い、評価方法としてPSNR, SSIM, VSNRを用いた。PSNR, SSIM, VSNRの評価値は、値が高ければ高いほど、画像の質が高く、鮮やかであることを示している。処理結果の評価値を表3、処理結果の画像の一部を図5~9に示す。

評価値を見てみると、Bicubic法、スパースコーディングと比べて、提案手法(GMM,GMM+PLS)が特に優れていることが分かる。これは、提案手法(GMM)、提案手法(GMM+PLS)が他の手法と比べて、適切に高周波成分を復元できることを示している。また、図5~9を見ると、提案手法(GMM)、提案手法(GMM+PLS)共に、従来手法と比べて鮮やかであることがわかる。

しかし、提案手法(GMM)と提案手法(GMM+PLS)を比較しても明確な差を得ることができなかった。GMMのみで変換関数を作成すると、過学習が発生する可能性があり、それを防ぐためにGMMとPLSを組み合わせた変換関数を作成したが、明確な差が得られなかった。これは、パッチサイズ3x3に比べて画像サイズ512x512が充分大きく、パラメータに対して学習データ量が充分であったためと考えられる。

表 3 PSNR,SSIM,VSNR による画像 (a),(b),(c),(d),(e) の比較

Image	Measure	PSNR	SSIM	VSNR
(a)	Bicubic	33.07	0.8015	16.20
	example-based	32.23	0.754	15.35
	Sparse-coding	34.20	0.8608	17.05
	GMM	35.57	0.9157	17.43
	GMM + PLS	36.74	0.9162	17.32
(b)	Bicubic	38.15	0.8977	17.27
	example-based	37.02	0.8682	16.71
	Sparse-coding	39.68	0.9240	17.72
	GMM	40.98	0.9476	17.89
	GMM + PLS	40.82	0.9474	17.90
(c)	Bicubic	36.43	0.8816	13.63
	example-based	35.42	0.8493	13.32
	Sparse-coding	37.71	0.9084	13.87
	GMM	39.00	0.9315	13.93
	GMM + PLS	39.50	0.9317	13.90
(d)	Bicubic	34.82	0.9097	15.19
	example-based	33.56	0.8832	14.21
	Sparse-coding	36.91	0.9454	15.75
	GMM	39.22	0.9731	15.87
	GMM + PLS	41.08	0.9730	15.76
(e)	Bicubic	32.50	0.7745	13.86
	example-based	32.03	0.7450	13.44
	Sparse-coding	33.36	0.8346	14.35
	GMM	33.97	0.8802	14.52
	GMM + PLS	33.71	0.8844	14.57

7. 結 論

本論文では、1枚の画像を超解像する手法として、自己縮小画像と混合ガウス分布モデルを用いた超解像を提案した。従来手法との比較を行った結果、2つの提案手法(GMMのみ, GMM+PLS)共に、従来手法より評価値が優れ、より鮮明な画像を作成することができ、提案手法の有効性を確認した。今後の課題として、パラメータの自動推定や更に有効な変換関数の作成方法などが挙げられる。

文 献

- [1] Wold, H. "Soft Modeling by Latent Variables, the Nonlinear Iterative Partial Least Squares Approach," in Perspective in Probability and Statistics, Paper in Honour of M. S. Bartlett, Academic Press, 520-540, 1975.
- [2] Freeman, W.T., Jones, T.R., Pasztor, E.C., "Example-Based Super-Resolution," IEEE Computer Graphics and Applications 22(2), 56-65, 2002.
- [3] Mario A. T. Figueiredo Michael Elad and Yi Ma, "On the role of sparse and redundant representations in image processing," Proceedings of the IEEE, vol. 98, no. 6, pp. 972-982, June 2010.
- [4] 天野敏之, 佐藤幸男, "固有空間法を用いたBPLPによる画像補間," 電子情報通信学会論文誌D-II, Vol.J85-D-II, No.3, pp.457-465, 2002.
- [5] Richardson, William Hadley, "Bayesian-Based Itera-

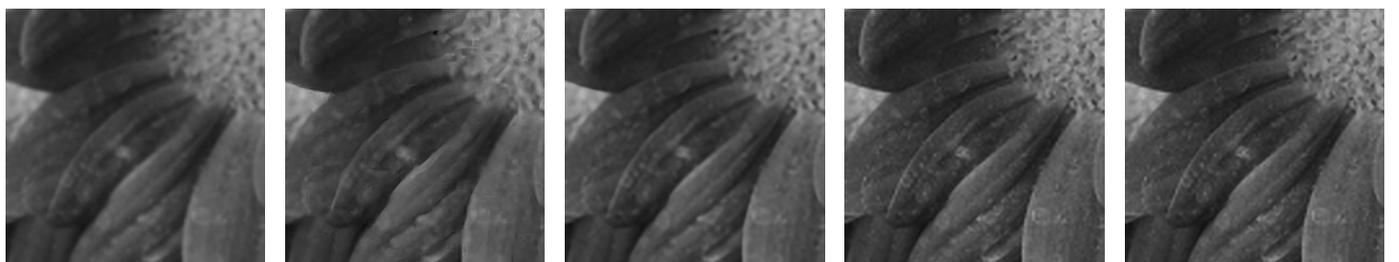
tive Method of Image Restoration". JOSA 62 (1): 55-59, 1972.

- [6] Olivier Cappe Yannis Stylianou and Eric Moulines, "Statistical methods for voice quality transformation," EUROSPEECH, pp. 447-450, September 1995.
- [7] Elina Helander, Tuomas Virtanen, Jani Nurminen and Moncef Gabbouj, "Voice Conversion Using Partial Least Squares Regression," IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 18, NO. 5, pp.912-921, JULY 2010
- [8] S. de Jong, "SIMPLS: An alternative approach to partial least squares regression," Chemometrics Intell. Lab. Syst., vol. 18, no. 3, pp.251-263, Mar 1993.
- [9] Hamid R. Sheikh Zhou Wang and Eero P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE TRANSACTIONS ON IMAGE PROCESSING, vol. 13, no. 4, pp. 600-612, April 2004.
- [10] Damon M. Chandler and Sheila S. Hemami, "Vsnr: A wavelet-based visual signal-to-noise ratio for natural images," IEEE TRANSACTIONS ON IMAGE PROCESSING, vol. 16, no. 9, pp. 2284-2298, September 2007.
- [11] 齊藤 隆弘, "(第2回) サンプリング定理の壁を打ち破る :1枚の画像からの超解像度オーバーサンプリング(次世代デジタルカメラ/デジタルムービーを予測する)(2)," 映像情報メディア学会誌 : 映像情報メディア 62(2), 181-189, 2008-02-01
- [12] 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 : 映像情報メディア = The journal of the Institute of Image Information and Television Engineers 64(11), 1613-1620, 2010-11-01



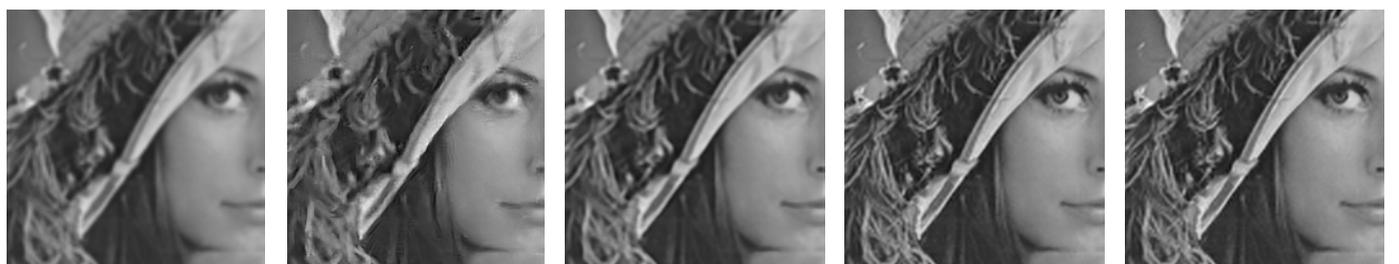
(1) Bicubic interpolation (2) Example-based (3) Sparse-coding (4) Proposed (GMM) (5) Proposed (GMM+PLS)

図 5 画像 (a) の処理結果



(1) Bicubic interpolation (2) Example-based (3) Sparse-coding (4) Proposed (GMM) (5) Proposed (GMM+PLS)

図 6 画像 (b) の処理結果



(1) Bicubic interpolation (2) Example-based (3) Sparse-coding (4) Proposed (GMM) (5) Proposed (GMM+PLS)

図 7 画像 (c) の処理結果



(1) Bicubic interpolation (2) Example-based (3) Sparse-coding (4) Proposed (GMM) (5) Proposed (GMM+PLS)

図 8 画像 (d) の処理結果



(1) Bicubic interpolation (2) Example-based (3) Sparse-coding (4) Proposed (GMM) (5) Proposed (GMM+PLS)

図 9 画像 (e) の処理結果