

スパースコーディングを用いたマルチフレーム超解像

加藤 利幸^{†1,a)} 日野 英逸^{†2,b)} 村田 昇^{†1,c)}

概要: 複数の低解像度の観測画像から高解像度画像の推定を行う、マルチフレーム超解像の一手法を提案する。提案手法は高解像度画像を推定するために、信号のスパース表現を用いる。スパース表現を用いる超解像手法の従来手法の多くはシングルフレーム超解像の手法である。また、スパース表現を用いるマルチフレーム超解像の従来手法は、画像間のサブピクセル精度の位置関係を考慮しないものであったため、複数枚の画像を効率的に利用できなかった。提案手法は低解像度画像に対応する辞書を、画像の劣化過程に従って作成し、その際にサブピクセル精度の位置ずれを取り入れることでこの問題を解決した。観測された低解像度画像の位置関係は、サブピクセル精度のブロックマッチングを用いて求める。超解像に用いる低解像度パッチのサブセットはマッチングの結果として得られる類似度に応じて選択する。提案手法の優れている点は、入力画像の中の一部の画像しか超解像に利用できない場合は、その条件に応じて出来る限り良い推定画像を作成することができる点である。特に一つのみの低解像度パッチが選択される場合、提案手法はシングルフレームの超解像手法としても実行される。実画像を使用した実験により、提案手法は従来のシングルフレームやマルチフレームの超解像手法と同等もしくはより良い結果を示すことを確認した。

1. はじめに

近年、動画像を高画質化する技術として超解像が大きな注目を集めている。超解像は一般的な画像補間法とは異なり、画像に関する仮定や事前知識を利用することで鮮鋭な画像を出力する。超解像が注目されるようになった背景としては映像機器の進化が挙げられる。高画質なコンテンツを再生することができる出力機器が身近になったことで、従来の低画質なコンテンツを高画質化して出力したいというニーズが高まっている。また、高い分解能を持つ撮影機器はコストが大きいと、従来の低分解能の撮影機器と組み合わせて超解像を用いることも考えられる。

代表的な超解像の方法としては再構成型超解像 [1], [2], [3] と、事例ベース超解像 [4], [5], [6] が知られている。再構成型超解像は入力画像に対して標本（画素）を追加するアイデアにより推定画像を作成する方法である。再構成型超解像は複数の入力画像を利用する「マルチフレーム超解像」で用いられることが多い。再構成型超解像では確率場を導入して画像を再構成することで良好な結果が得られることが確認されている。しかし、複雑なモデルは大量の計算

を伴う。一方、事例ベース超解像は学習用画像から作成した事例に基づき高解像度画像を推定する方法である。事例ベース超解像は 1 枚の入力画像から推定画像を作成する「シングルフレーム超解像」で用いられることが多い。事例ベースの超解像手法の例としては Freeman の方法 [4] や Neighbor Embedding (NE) [5] がある。Yang らは NE の超解像を改良して、超解像にスパースコーディング [7] の考え方を導入した「スパースコーディング超解像」を提案した [8], [9]。スパースコーディングでは信号が小数の基底の線形和によって表現できることを仮定し、それらの基底を画像の劣化過程を表す事例のように利用する。

シングルフレームに対するスパースコーディング超解像は複数提案されており、一般的な画像補間法と比較してより鮮明な高解像画像を生成することが可能である。単一の画像からその解像度を向上させるシングルフレーム超解像はインパクトの大きい手法であるが、例えば動画像の連続したフレームなど、同一の対象（フレーム）に対して複数の低解像度画像が得られる状況もあり、複数の低解像度を有効に利用することでより鮮明な画像の生成が期待できる。

これらの背景をふまえ、本稿ではスパースコーディングに基づくマルチフレーム超解像を提案する。提案手法ではブロックマッチングにより小領域ごとに位置ずれ推定を行う。ブロックマッチングを利用することで、大域的な位置ずれのみではなく、オブジェクト毎の位置ずれや歪み・回転を含む多くの状況に対応することができる。

^{†1} 現在、早稲田大学
Presently with Waseda University

^{†2} 現在、筑波大学
Presently with Tsukuba University

a) kato-toshiyuki@asagi.waseda.jp

b) hinohide@cs.tsukuba.ac.jp

c) noboru.murata@eb.waseda.ac.jp

2. 画像の観測モデル

本稿で仮定する画像の観測モデルについて述べる。まず、観測された離散的な画像 X は、真の撮影シーンを表す輝度値の連続的な分布 $\tilde{X}$ がワープ W 、ブラー H 、離散化 $\downarrow$ 、ノイズ ε の効果を経て得られているものであると仮定すると次の式が得られる。

$$X[m, n] = \left[\mathcal{H}(x, y) * W \left(\tilde{X}(x, y) \right) \right] \downarrow + \varepsilon[m, n] \quad (1)$$

(x, y) は連続的な空間での座標を表し、 $[m, n]$ は離散化された空間での座標を表す。* は畳み込みを表すオペレータである。

本稿では一般的な超解像の問題と同様に、離散的なピクセル領域で近似を行い、得られている低解像画像 $\mathbf{Y} \in \mathbb{R}^{p_l}$ が推定したい真の高解像画像 $\mathbf{X} \in \mathbb{R}^{p_h}$ が劣化して観測されたものであることを、ベクトルと行列を用いて表現する。

$$\mathbf{Y} = SHW\mathbf{X} + \varepsilon \quad (2)$$

$\mathbf{Y}$, $\mathbf{X}$ は画素が縦に積まれた列ベクトルである。また、 W はワープ (ゆがみ, 位置ずれ), H はブラー, S はダウンサンプリングを表すオペレータである。 ε は観測ノイズを表しており、正規乱数を仮定する。これらを式 (2) のように画像全体に適用する場合、各オペレータ及びノイズの次元はそれぞれ $W \in \mathbb{R}^{p_h \times p_h}$, $H \in \mathbb{R}^{p_h \times p_h}$, $S \in \mathbb{R}^{p_l \times p_h}$, $\varepsilon \in \mathbb{R}^{p_l}$ となる。

提案手法を含むスパースコーディング超解像では画像を小さなパッチに分割して処理を行う。高解像画像 $\mathbf{X}$ から切り出したパッチを $\mathbf{x} \in \mathbb{R}^{q_h}$ 、それに対応する低解像画像 $\mathbf{Y}$ から切り出したパッチを $\mathbf{y} \in \mathbb{R}^{q_l}$ と表記する。推定したい画像パッチ $\mathbf{x}$ とそれに対応する観測画像のパッチ $\mathbf{y}$ には以下の関係が成り立つ。

$$\mathbf{y} = SHW\mathbf{x} + \varepsilon \quad (3)$$

本稿では画像パッチの劣化過程は画像全体の劣化過程と共通の記号を用いて表す。画像パッチの劣化過程を記述するには、各オペレータ及びノイズの次元はそれぞれ $W \in \mathbb{R}^{q_h \times q_h}$, $H \in \mathbb{R}^{q_h \times q_h}$, $S \in \mathbb{R}^{q_l \times q_h}$, $\varepsilon \in \mathbb{R}^{q_l}$ となる。

提案手法はマルチフレーム超解像なので複数枚の観測画像を利用する。マルチフレーム超解像では、はじめに N 枚の観測画像 $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N$ から基準画像を決定する。最終的な出力は基準画像を拡大したものとなる。本稿では1枚目の観測画像が基準画像であるものとして提案手法を述べる。複数枚の観測画像の劣化過程は次のように記述される。

$$\mathbf{Y}_j = SHW_j\mathbf{X} + \varepsilon \quad j = 1, \dots, N \quad (4)$$

観測された画像はそれぞれ異なる位置ずれを持つため、 W

は画像の枚数分考えなければならない。ただし、基準画像に対応する W_1 は単位行列となる。ブラー H とダウンサンプリング S については、提案手法では全ての観測画像に共通のオペレータを用いることができるとする。それぞれの画像に対応する観測パッチ $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N$ は後述するクリッピングオペレータ $C_j \in \mathbb{R}^{q_l \times p_l}$ ($j = 1, \dots, N$) を用いて切り出す。

提案手法ではパッチ間のワープ W は平行移動に限ると仮定する。これはシンプルな仮定であるが、パッチ毎に位置関係を推定することで、歪みや回転などを含む現実的な問題にも対応できる。ブラー H は一般的にカメラのレンズぼけのようなものとして導入される。光学系のぼけ関数としては Point Spread Function (点広がり関数) が知られている。本稿ではブラー H を二次元のガウシアンフィルタ畳み込みとして記述する。ダウンサンプリング S は離散的にサンプリングを行い、画像の縮小を行うオペレータとする。

3. 位置ずれ推定

3.1 ブロックマッチング

マルチフレーム超解像を行うためには観測された複数枚の画像間の位置関係を推定することが必要となる。提案手法では画像を小さな矩形領域に分割して位置合わせを行うブロックマッチングを用いる。スパースコーディング超解像は同じく画像を小さな矩形領域に分割して処理を行うため、ブロックマッチングの結果をそのまま利用することができる。ブロック間の位置関係には平行移動を仮定する。ブロックのサイズを十分小さくすることで、平行移動のみで歪みや回転などを含む現実的な問題に対応することができる。

ブロック間の正確な位置関係はサブピクセル (実数) 精度で記述される。本稿では高速にサブピクセル精度の位置関係を推定することができる類似度補間法 [10], [11] を採用する。類似度補間法はピクセル (整数) 精度のマッチングで得られる類似度を補間して、連続的な領域で類似度が最大となる場所を推定する。具体的な方法として様々な手法が提案されているが、本稿ではその中で、縦方向の位置と横方向の位置を同時に推定することで高精度な推定結果が得られる「2次元同時サブピクセル推定法」[12] を利用する。

位置関係の推定では基準画像のいずれかのパッチと、それ以外の画像の間でマッチングを行う。本稿ではマッチングを行った際の類似度と閾値 δ を比較して、観測されたパッチを超解像に用いるかどうかを決定する。その際に用いる類似度は正規化した平均二乗誤差 $MSE \in [0, 1]$ 、またはそれを補間して得られた値とする。

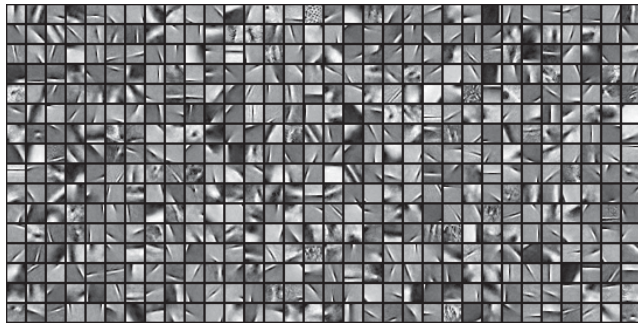


図 1 Lee の辞書学習法を用いて作成した辞書
Fig. 1 Learned atoms by Lee's method.

4. スパースコーディング

スパースコーディング [7] は、多数の基底から少数の基底を選択し、その線形結合で信号を近似する方法論である。代表的な方法として単純に少ない数の基底で信号を再構成する ℓ^0 ノルムスパースコーディングがある。

$$\text{minimize } \|\alpha\|_0 \quad \text{subject to } \|\mathbf{D}\alpha - \mathbf{s}\|_2^2 \leq \epsilon \quad (5)$$

ここで、 $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_K] \in \mathbb{R}^{p \times K}$ は K 個の基底からなる辞書、 $\alpha \in \mathbb{R}^K$ は係数ベクトル、 $\mathbf{s} \in \mathbb{R}^p$ は再構成する信号を表している。 $\epsilon \in \mathbb{R}$ は再構成を行う際に許容する残差の大きさを表す。この最適化は NP 困難であり [13]、多くの場合、Matching Pursuit タイプのアルゴリズム [14], [15] で近似的に解が求められる。

他のアプローチとして係数ベクトルの ℓ^1 ノルムをスパースさの基準とする、 ℓ^1 ノルムスパースコーディングがある。

$$\text{minimize } \|\alpha\|_1 \quad \text{subject to } \|\mathbf{D}\alpha - \mathbf{s}\|_2^2 \leq \epsilon \quad (6)$$

この問題はラグランジュ乗数 $\eta \geq 0$ を用いて次のように書き直すことができる。

$$\text{minimize } \frac{1}{2} \|\mathbf{D}\alpha - \mathbf{y}\|_2^2 + \eta \|\alpha\|_1 \quad (7)$$

η は係数ベクトルのスパースさの制約の強さを表している。制約の強さを適切に調整することで係数 α は十分スパースになる。また、 ℓ^1 ノルムスパースコーディングは安定的な解を与えることが知られている [16]。式 (7) は回帰で用いられる Lasso [17] と同じ形式となっている。この解を効率的に求める方法としては LARS アルゴリズム [18] や Feature-sign Search アルゴリズム [19] などが知られている。本稿では、 ℓ^1 ノルムスパースコーディングを利用する。

スパースコーディングはデータから学習した基底の辞書 $\mathbf{D}$ を用いて信号を再構成することで、高精度な信号の近似が可能となる。辞書の学習は十分な量の学習用データ $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n] \in \mathbb{R}^{p \times n}$ からあらかじめ行っておく。辞書を学習するためのアルゴリズムとしては Lee の方法 [19] や K-SVD アルゴリズム [20] などが知られている。本稿で

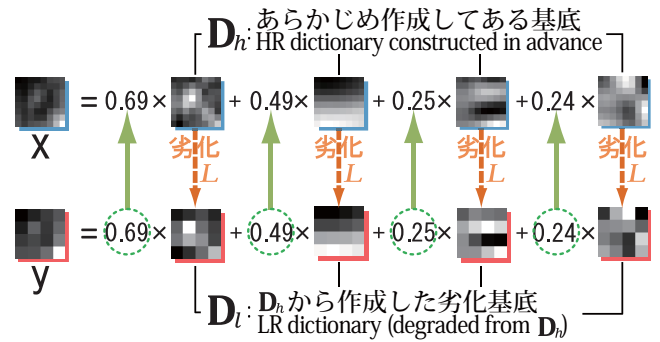


図 2 スパースコーディング超解像の枠組み
Fig. 2 A framework of super resolution based on sparse coding.

は ℓ^1 ノルムスパースコーディングで用いる辞書の学習を高速に行う方法として一般的な Lee の方法を採用する。Lee の方法で学習した辞書を図 1 に示す。

5. スパースコーディング超解像

5.1 スパースコーディング超解像の考え方

スパースコーディングを用いて、超解像で推定画像を作成する方法を述べる。また、再構成画像をより良いものにするために行われるいくつかの工夫についてまとめる。

5.1.1 推定画像の作成方法

スパースコーディング超解像では画像を小さなパッチに分割して処理を行う。まず、スパースコーディングを用いて超解像を行うには高解像パッチ $\mathbf{x}$ に対応する辞書 $\mathbf{D}_h \in \mathbb{R}^{q_h \times K}$ と、低解像パッチ $\mathbf{y}$ に対応する辞書 $\mathbf{D}_l \in \mathbb{R}^{q_l \times K}$ をそれぞれ用意する必要がある。画像の劣化過程 SHW をまとめて $L \in \mathbb{R}^{q_l \times q_h}$ を用いて表すと、以下の関係が成り立つ。

$$\mathbf{y} = L\mathbf{x} = L\mathbf{D}_h\alpha \quad (8)$$

式 (8) において高解像パッチ $\mathbf{x}$ は辞書 $\mathbf{D}_h$ と係数 $\alpha \in \mathbb{R}^K$ によって再構成され则认为している。ここで、 $L\mathbf{D}_h$ を低解像画像に対応する辞書 $\mathbf{D}_l$ とみなすと観測された低解像画像は辞書 $\mathbf{D}_l$ と係数 α によって再構成され则认为される。この時、 $L\mathbf{D}_h$ は高解像画像に対応する辞書 $\mathbf{D}_h$ を画像と同じ劣化過程によって劣化させたものを表す。そして、観測された低解像画像 $\mathbf{y}$ と辞書 $\mathbf{D}_l = L\mathbf{D}_h$ から係数 α を推定すれば、 $\mathbf{x} = \mathbf{D}_h\alpha$ として高解像画像を再構成することが出来る。このスパースコーディング超解像の枠組みを図 2 に示す。

5.1.2 パッチ重複部の処理と工夫

パッチ毎に処理を行う際、単純に画像を格子状に分割して処理するだけでは、再構成画像においてパッチとパッチの境界がなめらかに繋がらず、ブロック状のアーティファクトがあらわれてしまう。このような現象が発生しないように、スパースコーディング超解像ではパッチを重複させて処理を行う。重複させる領域を増やすほど精度が向上す

るが、重複させる領域を増やすほど計算量が増大する。

各パッチに関する超解像の結果をうまく統合するために様々な工夫が提案されているが、平均または重み付き平均を用いる方法がシンプルで効果的である。提案手法では2次元のハニング窓の重み付き平均で各パッチに関する結果を統合し、再構成画像を作成する。

5.1.3 バックプロジェクション法による解の補正

スパースコーディング超解像では画像を小さなパッチに分割して処理を行うため、劣化基底に推定係数を掛け合わせたものが原信号を劣化させたものと大域的に一致することを仮定していない。これが原因となり、最も基本的な制約である $\mathbf{Y} = \mathbf{SHX}$ はパッチ毎の係数推定のみによる再構成画像によっては満たされない。そこで、バックプロジェクション法 [21] を用いて解の補正を行う。超解像による推定画像を $\mathbf{X}_0 \in \mathbb{R}^{p \times h}$ として、次の最適化問題を解く。

$$\mathbf{X}^* = \arg \min_{\mathbf{X}} \|\mathbf{SHX} - \mathbf{Y}\|_2^2 + c\|\mathbf{X} - \mathbf{X}_0\|_2^2 \quad (9)$$

この最適化は勾配法によって繰り返し計算で行うことができる。 ν をステップ幅、 c を重み係数とすると、最適化を行うための更新式は以下のように与えられる。

$$\mathbf{X}_{t+1} = \mathbf{X}_t - \nu[H^T S^T(\mathbf{SHX}_t - \mathbf{Y}) + c(\mathbf{X} - \mathbf{X}_0)] \quad (10)$$

5.2 スパースコーディングを用いたマルチフレーム超解像

スパースコーディングを用いた超解像をマルチフレーム超解像へと応用する方法について述べる。マルチフレーム超解像は劣化過程、特に位置ずれの異なる複数の観測画像を利用して高解像の画像を再構成する問題である。シングルフレームの時と同様に、再構成すべき高解像画像を $\mathbf{X}$ 、そこから抜き出したパッチを $\mathbf{x}$ とし、観測された N 枚の劣化画像を $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N$ 、そこから抜き出したパッチを $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N$ とする。

マルチフレームのスパースコーディング超解像はシングルフレームのスパースコーディング超解像と同様の手順で実行される。異なる点は係数推定の際に、複数枚の観測パッチを利用する点である。そのために、後述する結合辞書 $\tilde{\mathbf{D}}_l \in \mathbb{R}^{q \times K}$ と結合観測ベクトル $\tilde{\mathbf{y}}_l \in \mathbb{R}^q$ から係数推定を行う。

$$\min_{\alpha} \|\tilde{\mathbf{D}}_l \alpha - \tilde{\mathbf{y}}_l\|_2^2 + \eta \|\alpha\|_1 \quad (11)$$

$$\text{with } \tilde{\mathbf{D}}_l = \begin{bmatrix} \mathbf{D}_{l1} \\ \vdots \\ \mathbf{D}_{lN} \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{y}}_l = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} \quad (12)$$

この定式化では観測パッチと各基底に対応するベクトルそれぞれを縦につないで一つのベクトルにしている。辞書 $\mathbf{D}_l$ には各基底に対応するベクトルが列方向に配置されているので、辞書を縦につなげることは各基底を縦につなげ

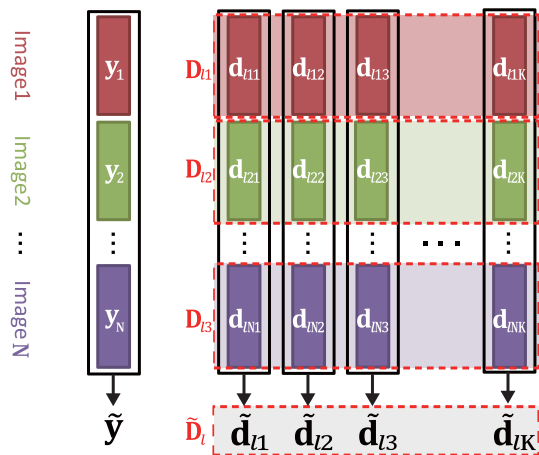


図 3 マルチフレーム超解像のためのベクトルの結合

Fig. 3 Stack vectors for multi-frame SR.

ることを意味する。ベクトルの結合の様子を図 3 に示す。

具体的に観測された複数枚の画像から $\tilde{\mathbf{y}}$ と $\tilde{\mathbf{D}}_l$ を作成するにはいろいろな方法が考えられる。基準画像を 1 枚目とすると $\mathbf{y}_1$ と $\mathbf{D}_{l1}$ はシングルフレーム超解像の時と同様にして作成することができるので、1 枚目以外の画像に対応する $\mathbf{y}_2, \dots, \mathbf{y}_N$ と $\mathbf{D}_{l2}, \dots, \mathbf{D}_{lN}$ が重要となる。

提案手法と類似したモチベーションで提案されている Wang の方法 [22] では、観測された劣化画像を再構成したい画像のサイズに拡大した後、整数ピクセル精度でブロックマッチングを行い $\mathbf{y}_1$ に対応するパッチを $\mathbf{Y}_2 \dots \mathbf{Y}_N$ の中から決定し、 $\mathbf{y}_2 \dots \mathbf{y}_N$ としている。また、辞書は各画像について共通のものを使用し、 $(\mathbf{D}_{l1}, \dots, \mathbf{D}_{lN}) = (\mathbf{D}_l, \dots, \mathbf{D}_l)$ としている。Wang の方法では、明示的なサブピクセル精度の位置ずれ推定は行わない。よって、ノイズが比較的小さく、サブピクセル精度の位置ずれ推定を高精度に行える場合は、それを利用することでより高精度な推定画像が期待できる。

5.3 提案手法

5.3.1 サブピクセル精度の位置ずれの考慮

提案手法では複数枚の画像間のサブピクセル精度の位置ずれを考慮して、結合辞書 $\tilde{\mathbf{D}}_l \in \mathbb{R}^{q \times K}$ と結合観測ベクトル $\tilde{\mathbf{y}}_l \in \mathbb{R}^q$ を作成する。簡単のために一次元の例を用いた概念図を図 4 に示す。図の横軸は画像上の位置、縦軸は輝度値を表している。まず、観測された N 枚の低解像画像 (g)~(i) のそれぞれに対応する高解像画像 (d)~(f) と元となる輝度値の連続分布 (a)~(c) を考える。ここで、式 (1) は (a) と (d) または (g)、式 (2) は (d) と (g) の関係を表していると解釈することができる。輝度値の連続分布 (a)~(c) は、平行移動の位置ずれ情報を用いて、ある画像の分布から他の画像の分布を求めることができる。一方離散的な位置で値を持つ高解像画像 (d)~(f) は、いずれか一つから他の画像に対応するものを正確に計算することはできな

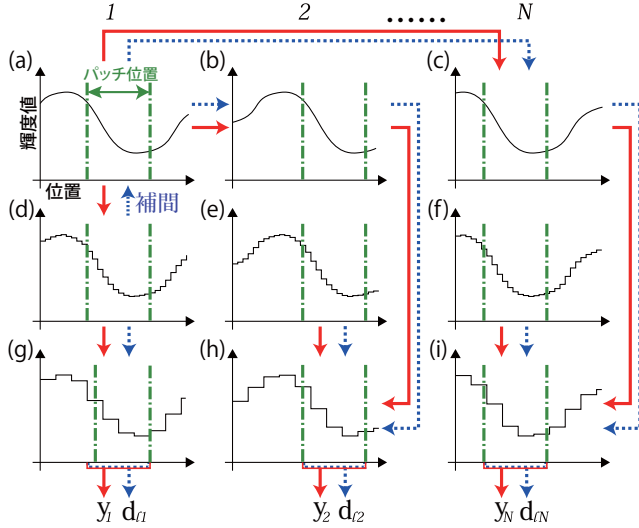


図 4 サブピクセル精度の位置ずれを考慮したサンプリング
Fig. 4 Sampling with considering sub-pixel shift.

い。図 4 では、実線の矢印が観測画像の劣化過程、点線の矢印が基底ベクトルの作成手順、破線がサブピクセル精度の位置関係を表している。観測された複数枚の画像はサブピクセル精度の位置ずれを含む。観測ベクトル $\mathbf{y}$ は、ブロックマッチングで求めた位置関係に応じて得られている画素を直接クリッピングしたものとする。画素のクリッピングではパッチの位置関係を領域で考え、その領域の内側に入っている画素が持っている輝度値を抜き出してベクトルとする。基底ベクトル $\mathbf{d}$ は、基底学習アルゴリズムにより (d) に相当する基準画像の高解像基底があらかじめ得られているので、補間を行い (a) に相当する連続分布を作成する。その後連続分布の平行移動により (b) や (c) を求め、画像の劣化過程に従い (h) や (i) を求める。実装では、次節で述べるように一般的な画像補間法を用いた操作で (d) から直接 (e) や (f) を計算し、ブラーとダウンサンプリングを加えて (h) や (i) を作成する。

5.3.2 実装

提案手法の結合観測ベクトル $\tilde{\mathbf{y}}$ の作成方法を述べる。基準画像の再構成したいパッチの位置に合わせて、各低解像画像からその領域内のピクセルをクリッピングし、一つのベクトルにつなぎ合わせる。その様子を図 5 に示す。観測画像からクリッピングを行うオペレータを $C_j \in \mathbb{R}^{q_l \times p_l}$ ($j = 1, \dots, N$) とすると、 $\tilde{\mathbf{y}}$ は以下のように表される。

$$\tilde{\mathbf{y}} = \begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} C_1 \mathbf{Y}_1 \\ \vdots \\ C_N \mathbf{Y}_N \end{bmatrix} \quad (13)$$

次に、低解像画像に対応する辞書 $\tilde{\mathbf{D}}_l$ の作成方法について述べる。基底の作成では高解像画像に基底 $\mathbf{d}_h$ を埋め込み、各画像の劣化過程に合わせて画像を劣化させる。平行移動は、共一次内挿法 [23] を用いて行う。共一次内挿法は輝度値を推定したい位置の近傍の 4 画素を線形に補間して、値

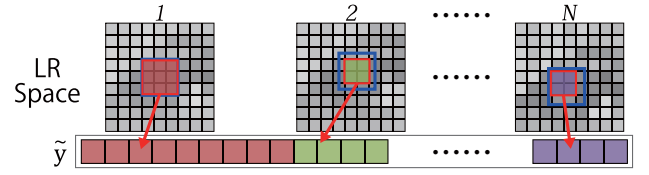


図 5 観測ベクトル $\tilde{\mathbf{y}}$ の作り方
Fig. 5 The procedure of generating stacked LR observation $\tilde{\mathbf{y}}$.

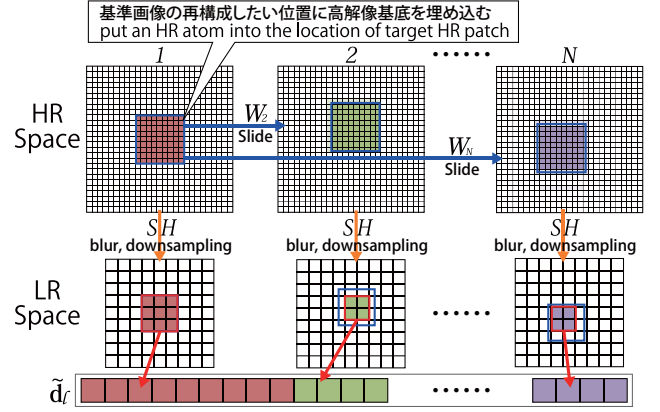


図 6 劣化基底ベクトル $\tilde{\mathbf{d}}_l$ の作り方
Fig. 6 The procedure of generating stacked LR dictionary $\tilde{\mathbf{D}}_l$.

の推定を行う方法である。推定したい位置を (x, y) として、 u_1, u_2 を x, y の値を越えない最大の整数、 $v_1 = x - u_1$ 、 $v_2 = x - u_2$ とすると次式により輝度値が得られる。

$$Y(x, y) = \begin{bmatrix} 1-v_1 & v_1 \end{bmatrix} \begin{bmatrix} Y[u_1, u_2] & Y[u_1+1, u_2] \\ Y[u_1, u_2+1] & Y[u_1+1, u_2+1] \end{bmatrix} \begin{bmatrix} 1-v_2 \\ v_2 \end{bmatrix} \quad (14)$$

行列 W を作成する際には、推定したい位置の 4 近傍の画素とのピクセル精度の対応関係を表す 4 つの行列を作成し、その重みつき和を計算すればよい。

$$W = (1-v_1)(1-v_2)W_{[w_1, w_2]} + (1-v_1)v_2W_{[w_1+1, w_2]} + v_1(1-v_2)W_{[w_1, w_2+1]} + v_1v_2W_{[w_1+1, w_2+1]} \quad (15)$$

w_1, w_2 は横方向と縦方向の位置ずれの大きさを整数で表し、 $W_{[w_1, w_2]}$ は横方向に w_1 ピクセル、縦方向に w_2 ピクセルの平行移動を行うワープオペレータである。劣化させた画像からは観測信号と同様にしてクリッピングを行い、ベクトルを結合して $\tilde{\mathbf{d}}_l$ を得る。その様子を図 6 に示す。それぞれの基底について図の操作を行い、得られた K 個のベクトル $\tilde{\mathbf{d}}_l$ をまとめて結合辞書 $\tilde{\mathbf{D}}_l$ とする。高解像画像に基底の埋め込みを行うオペレータを $R \in \mathbb{R}^{p_h \times q_h}$ とすると、 $\tilde{\mathbf{D}}_l$ は以下のように表すことができる。

$$\tilde{\mathbf{D}}_l = \begin{bmatrix} \mathbf{D}_{l1} \\ \vdots \\ \mathbf{D}_{lN} \end{bmatrix} = \begin{bmatrix} C_1 SHW_1 R \mathbf{d}_h \\ \vdots \\ C_N SHW_N R \mathbf{d}_h \end{bmatrix} \quad (16)$$

Algorithm 1 提案手法

Input: LR Images $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N$ and HR Dictionary $\mathbf{D}_h$
{ スパースコーディングを用いた超解像 }
for each patch $\mathbf{y}_k$ **do**
 [Step1] BlockMatching {cf. 3 節}
 [Step2] Make $\tilde{\mathbf{y}}_l, \tilde{\mathbf{D}}_l$ {cf. 図 5, 図 6}
 [Step3] $m \leftarrow \text{mean}(\tilde{\mathbf{y}}_l)$
 $\tilde{\mathbf{y}}_l \leftarrow \tilde{\mathbf{y}}_l - m$
 $\alpha \leftarrow \text{SparseCoding}(\tilde{\mathbf{D}}_l, \tilde{\mathbf{y}}_l)$
 [Step4] $\mathbf{x} \leftarrow \mathbf{D}_h \alpha + m$
end for
Make $\mathbf{X}_0$ from all $\mathbf{x}$ {cf. 5.1.3 節}
{ 再構成結果の補正 }
repeat
 $\mathbf{X}_{t+1} = \mathbf{X}_t - \nu[H^T S^T(SH\mathbf{X}_t - \mathbf{Y}) + c(\mathbf{X} - \mathbf{X}_0)]$
until convergence
return $\mathbf{X}$

5.3.3 提案手法のまとめ

最終的に提案する手法の処理の手順を以下に記述する。提案手法は、大きく分けると2つの処理からなる。1つはスパースコーディングを用いた超解像、2つは再構成結果の補正である。

超解像では画像を小さく分割したパッチ、それぞれについて処理を行い、最後に5.1.2節で述べたように結果を重み付き平均で統合して仮の結果 $\mathbf{X}_0$ を作成する。パッチについて行う処理は4つのステップからなる。ステップ1では3節で述べたブロックマッチングを用いてサブピクセル精度の位置ずれ推定を行う。ステップ2では5.3節の方法で結合観測ベクトル $\tilde{\mathbf{y}}_l$ と結合辞書 $\tilde{\mathbf{D}}_l$ を作成する。ステップ1でブロックマッチングを行った際、類似度が閾値より高いブロックが見つからなかった画像は結合観測ベクトルと結合辞書の作成に用いない。ステップ3では4節で述べた ℓ^1 ノルムスパースコーディングの方法で $\tilde{\mathbf{y}}_l$ と $\tilde{\mathbf{D}}_l$ から係数 α を推定する。この時、 $\tilde{\mathbf{y}}_l$ からバイアス m を取り除いてから処理を行う。ステップ4では元の基底 $\mathbf{D}_h$ と α から推定パッチ $\mathbf{x}$ を作成する。また、ステップ3で取り除いたバイアスを再構成結果に加える。

再構成結果の補正を行うステージでは5.1.3節で述べたようにバックプロジェクション法で補正を行う。提案手法はシングルフレームの超解像と同様にして、仮の超解像結果 $\mathbf{X}_0$ と基準画像 $\mathbf{Y}_1$ のみを用いて補正を行う。

6. 実験

静止画から人工的に作成したデータと、実動画から抜き出したフレームを用いて提案手法の有効性を示す。画像への適用では5枚の劣化画像から、その中の1枚の画像を拡大した画像を作成する。動画への適用では基準とするフレームとその前後2フレームの合計5フレームの画像から、基準とするフレームを拡大した画像を作成する。

6.1 画像への適用

はじめに数値的な評価を行うために人工的に作成したデータを用いて実験を行う。評価基準にはPSNRを用いる。

$$\text{PSNR[dB]} = 10 \log_{10} \frac{255^2}{\text{MSE}} \quad (17)$$

MSEは元画像と再構成した画像の平均二乗誤差を表す。

画像の観測条件は2節で述べたモデルに従って次のように仮定する。複数枚の劣化画像がワープ、ブラー、ダウンサンプリング、ノイズの劣化過程を経て観測されているとする。ワープは上下左右方向への一様分布 $[-5, 5]$ の平行移動、ブラーは 9×9 ピクセルで標準偏差 $\sigma_h = 1$ のガウシアンフィルタの畳み込み、ノイズは標準偏差 $\sigma_n = \sqrt{2}$ の正規分布を加えるもの(SNRはおよそ37~40[dB])、スケールレートは3倍として5枚の劣化画像を作成する。ブラー、ノイズの強さは既知とする。

本稿では提案手法を次の4つの方法と比較して結果の評価を行う。バイキュービック補間、Yangの方法[9]、Zeydeの方法[24]、そして、Wangの方法[22]である。Wangらのマルチフレームの超解像手法は提案手法と同様に、基準画像を含む5枚の画像に対して適用した。一枚のみの画像を用いて高解像度画像を作成するバイキュービック補間、Yangの方法、Zeydeの方法は、基準画像のみにに対して適用した。また、提案手法はブロックマッチングを用いて画像間の位置関係を推定とした場合と、位置関係を既知とした場合の二通りの結果を示す。

提案手法には7つのパラメータがある。1) スパースコーディングにおけるノルムの制約 η 、2) スパースコーディングで用いる基底の数 m 、3) パッチのサイズ q_h 、4) パッチの重複幅、5) ブロックマッチングの結果を超解像に用いるかどうかを決定する類似度の閾値 δ 、6) 勾配法を用いた解の補正の制約 c 、7) 解の補正におけるステップ幅 ν である。

スパースコーディングの制約 η と解の補正の制約 c は、ノイズの大きさなどを考慮して値の設定を行う。本稿では実験的にチューニングを行い、 $\eta = 0.05$ 、 $c = 0.0001$ とした。類似度の閾値 δ は、ブロックマッチングの結果として得られるブロック間の類似度から、基準画像以外の画像から得られるサンプルを超解像に用いるかどうかを決定するパラメータである。 δ の値が大きいほど多くの画像からサンプルを行い、超解像に利用する。画像への適用では $\delta = 1$ として、5枚の画像全てから超解像に用いるための観測ベクトルをサンプルするように設定した。残りのパラメータは計算量と再構成精度のトレードオフを考慮して、表1のように設定した。

実験対象は5種類の画像(Lena, Cameraman, Flower, Girl, Parthenon)とした。実験結果のPSNRを表2に示す。また、Lena, Cameramanの結果画像を図7、図8に示す。提案手法では数値的にも視覚的にも良い結果が得られている。視覚的にはエッジ部分がより鮮鋭になっている。

表 2 超解像結果の PSNR
Table 2 PSNRs of SR methods.

Image	Input	Bicubic	Yang(2010)	Zeyde(2010)	Wang(2011)	提案手法 (位置ずれ未知)	提案手法 (位置ずれ既知)
Lena	26.34 $\pm$ 0.00	28.01 $\pm$ 0.01	28.64 $\pm$ 0.01	29.38 $\pm$ 0.01	29.05 $\pm$ 0.01	29.98 $\pm$ 0.12	30.89 $\pm$ 0.20
Cameraman	24.37 $\pm$ 0.00	26.87 $\pm$ 0.01	27.85 $\pm$ 0.01	29.08 $\pm$ 0.01	28.12 $\pm$ 0.01	29.81 $\pm$ 0.35	32.28 $\pm$ 0.42
Flower	33.57 $\pm$ 0.00	35.50 $\pm$ 0.00	35.95 $\pm$ 0.00	36.67 $\pm$ 0.00	36.41 $\pm$ 0.00	36.57 $\pm$ 0.12	37.07 $\pm$ 0.13
Girl	29.87 $\pm$ 0.00	31.11 $\pm$ 0.00	31.49 $\pm$ 0.00	31.85 $\pm$ 0.00	31.66 $\pm$ 0.00	31.92 $\pm$ 0.06	32.25 $\pm$ 0.08
Parthenon	23.52 $\pm$ 0.00	24.32 $\pm$ 0.00	24.45 $\pm$ 0.00	24.95 $\pm$ 0.00	24.64 $\pm$ 0.00	25.38 $\pm$ 0.11	26.37 $\pm$ 0.18

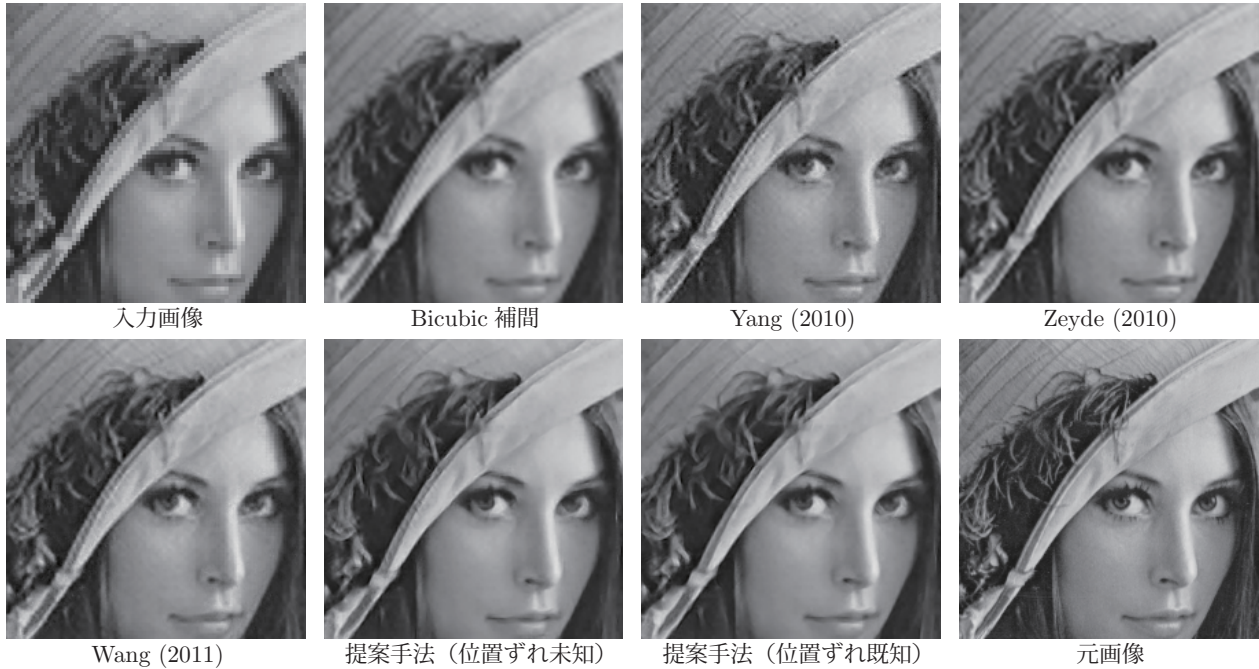


図 7 超解像結果 (Lena)
Fig. 7 Images estimated from LR observations (Lena).

表 1 パラメータの設定 (画像への適用)
Table 1 Parameter settings (Still Images).

パラメータ	値
ノルムの制約 η	0.05
基底の数 m	512
パッチのサイズ q_h	$15 \times 15 = 125$
パッチの重複幅	9
類似度の閾値 δ	1
解の補正の制約 c	0.0001
解の補正のステップ幅 μ	0.001

6.2 動画への適用

提案手法が現実的な問題に対しても適用できることを示すために、実動画を用いた場合の実験結果を示す。基本的には 6.1 節と同様の観測条件、設定条件で実験を行った。

元動画から基準画像と前後 2 フレームの計 5 フレームを抜き出し、ブラー、ダウンサンプリング、ノイズで劣化させて実験用の観測画像とした。

パラメータの設定は、基準画像以外の画像からサンプルした観測ベクトルを超解像に用いるどうかを決定する類

似度の閾値 δ 以外は 6.1 節と同様である。類似度の閾値 δ は $\delta = 0.003$ と $\delta = 0.001$ の二通りの値でそれぞれ実験を行った。動画から抜き出したフレームでは、あるフレームに写っているオブジェクトが別のフレームでは他のオブジェクトの裏側に隠れてしまうことや、そのオブジェクトが画面外に出て行くことなどを考慮しなければならない。 $\delta = 0.003$ は類似度がやや高いパッチのみを用いて超解像を行う場合、 $\delta = 0.001$ は類似度がより高いパッチのみを用いて超解像を行う場合である。

実験対象は 1 種類の動画 (MacArthur) とした。実験に用いた 5 フレームを図 9 に示す。実験結果を図 10 に示す。提案手法では、顔の輪郭などのエッジ部分や手すりなどのライン状の特徴が鮮鋭になっている。一方、服のシワや模様のようなテクスチャ成分は綺麗に再現されていない所も見られる。

7. おわりに

提案手法は複数の画像を利用して、その中の 1 枚の画像を超解像する手法である。実験では、人工的に作成した実



図 8 超解像結果 (Cameraman)

Fig. 8 Images estimated from LR observations (Cameraman).

験用の画像のみではなく、実動画に対しても提案手法が有効であることを示した。提案手法は複数のフレームの情報を利用して、シングルフレームのスパースコーディング超解像より鮮鋭な画像を作成する。提案手法は類似したモチベーションで提案された Wang らの手法よりもよい結果をもたらすことが実験的に確かめられた。Wang らの方法は、画像間の位置ずれを整数単位でとらえるのに対し、提案手法は実数精度の位置ずれ推定結果を超解像に利用する。提案手法の優れている点として入力画像の中の一部の画像しか超解像に利用できない場合は、その条件に応じて出来る限り良い推定画像を作成することができる点がある。特に入力画像に超解像に利用できるフレームが 1 フレームしかない場合は提案手法はシングルフレームの超解像として実行される。

将来的な課題としては再構成精度の向上と処理速度の高速化がある。これらの課題を実現するためには、発展的なスパースコーディングの方法や超解像のアルゴリズムを考案する必要がある。例えばシングルフレームのスパースコーディング超解像では様々な工夫が提案されているので [24], [25], [26], [27], それらの工夫をマルチフレームのスパースコーディング超解像へと取り入れることで性能の向上が期待できる。

参考文献

- [1] Hardie, R., Barnard, K. and Armstrong, E.: Joint MAP registration and high-resolution image estimation using a sequence of undersampled images, *IEEE Transactions on Image Processing*, Vol. 6, No. 12, pp. 1621–1633 (1997).
- [2] Farsiu, S., Robinson, M., Elad, M. and Milanfar, P.: Fast and robust multiframe super resolution, *IEEE Transactions on Image Processing*, Vol. 13, No. 10, pp. 1327–1344 (2004).
- [3] Kanemura, A., Maeda, S. and Ishii, S.: Hyperparameter Estimation in Bayesian Image Superresolution with a Compound Markov Random Field Prior, *IEEE Workshop on Machine Learning for Signal Processing*, pp. 181–186 (2007).
- [4] Freeman, W., Jones, T. and Pasztor, E.: Example-based super-resolution, *Computer Graphics and Applications, IEEE*, Vol. 22, No. 2, pp. 56–65 (2002).
- [5] Yeung, D.-Y., Yeung, D.-Y. and Xiong, Y.: Super-resolution through neighbor embedding, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 1, pp. I-275–I-282 Vol.1 (2004).
- [6] Sun, J., Zheng, N.-N., Tao, H. and Shum, H.-Y.: Image hallucination with primal sketch priors, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 2, pp. II-729–36 vol.2 (2003).
- [7] Olshausen, B. A. and Field, D. J.: Emergence of simple-cell receptive field properties by learning a sparse code for natural images, *Nature*, Vol. 381, pp. 607–609 (1996).
- [8] Yang, J., Wright, J., Huang, T. and Ma, Y.: Image super-resolution as sparse representation of raw image patches, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–8 (2008).
- [9] Yang, J., Wright, J., Huang, T. and Ma, Y.: Image Super-Resolution Via Sparse Representation, *IEEE Transactions on Image Processing*, Vol. 19, No. 11, pp. 2861–2873 (2010).
- [10] Tian, Q. and Huhns, M. N.: Algorithms for subpixel registration, *Comput. Vision Graph. Image Process.*, Vol. 35, No. 2, pp. 220–233 (1986).
- [11] Dvornychenko, V. N.: Bounds on (Deterministic) Correlation Functions with Application to Registration, *IEEE*



図 9 動画から抜き出した実験用画像 (MacArthur)

Fig. 9 Images captured from movies (MacArthur).

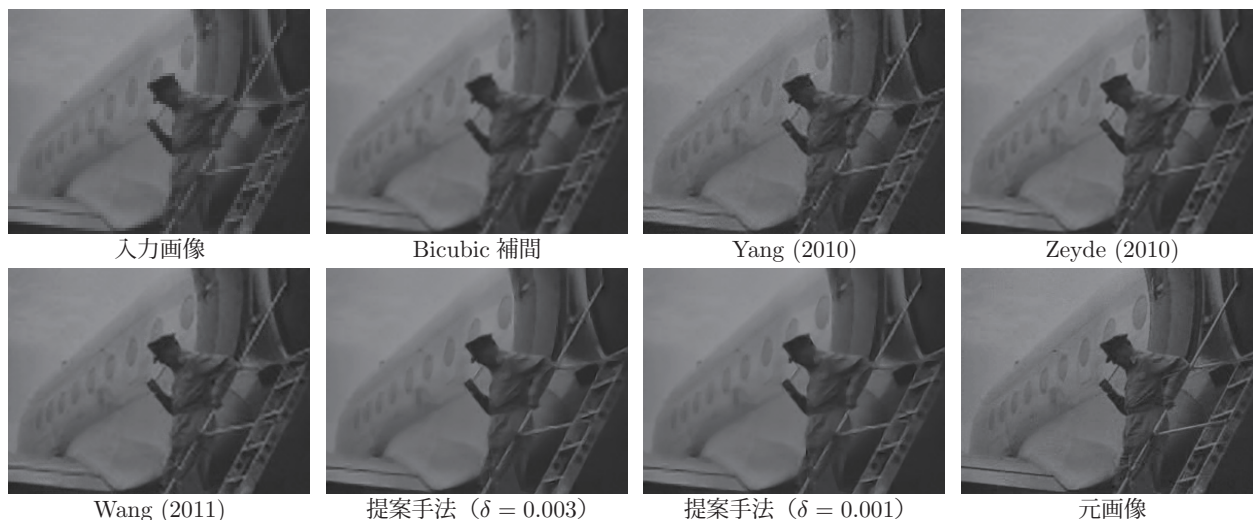


図 10 超解像結果 (MacArthur)

Fig. 10 Images estimated from LR observations (MacArthur).

- Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-5, No. 2, pp. 206–213 (1983).
- [12] 清水雅夫, 奥富正敏: 領域ベースマッチングのための 2 次元同時サブピクセル推定法 (画像処理, 画像パターン認識), 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理, No. 2, pp. 554–564 (オンライン), 入手先 (<http://ci.nii.ac.jp/naid/110003171041/>) (2004).
- [13] Davis, G., Mallat, S. and Avellaneda, M.: Adaptive greedy approximations, *Constructive Approximation*, Vol. 13, No. 1, pp. 57–98 (1997).
- [14] Mallat, S. and Zhang, Z.: Matching pursuits with time-frequency dictionaries, *IEEE Transactions on Signal Processing*, Vol. 41, No. 12, pp. 3397–3415 (1993).
- [15] Pati, Y. C., Rezaiifar, R., Rezaiifar, Y. C. P. R. and Krishnaprasad, P. S.: Orthogonal Matching Pursuit: Recursive Function Approximation with Applications to Wavelet Decomposition, *Proceedings of the 27th Annual Asilomar Conference on Signals, Systems, and Computers*, pp. 40–44 (1993).
- [16] Chen, S. S., Donoho, D. L., Michael and Saunders, A.: Atomic decomposition by basis pursuit, *SIAM Journal on Scientific Computing*, Vol. 20, pp. 33–61 (1998).
- [17] Tibshirani, R.: Regression Shrinkage and Selection Via the Lasso, *Journal of the Royal Statistical Society, Series B*, Vol. 58, pp. 267–288 (1994).
- [18] Trevor, B. E., Hastie, T., Johnstone, L. and Tibshirani, R.: Least Angle Regression, *Annals of Statistics*, Vol. 32, pp. 407–499 (2002).
- [19] Lee, H., Battle, A., Raina, R. and Ng, A. Y.: Efficient sparse coding algorithms, *In NIPS*, pp. 801–808 (2007).
- [20] Aharon, M., Elad, M. and Bruckstein, A.: K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation, *IEEE Transactions on Signal Processing*, Vol. 54, No. 11, pp. 4311–4322 (2006).
- [21] Irani, M. and Peleg, S.: Motion Analysis for Image Enhancement: Resolution, Occlusion, and Transparency, *Journal of Visual Communication and Image Representation*, Vol. 4, pp. 324–335 (1993).
- [22] Wang, P., Hu, X., Xuan, B., Mu, J. and Peng, S.: Super Resolution Reconstruction via Multiple Frames Joint Learning, *International Conference on Multimedia and Signal Processing (CMSP)*, Vol. 1, pp. 357–361 (2011).
- [23] 末松良一, 山田宏尚: 画像処理工学, コロナ社 (2000).
- [24] Zeyde, R., Elad, M. and Protter, M.: On single image scale-up using sparse-representations, *Proceedings of the 7th international conference on Curves and Surfaces*, Berlin, Heidelberg, Springer-Verlag (2012).
- [25] Min, L., Shihua, L., Fu, W., Xiang, L., Hong, J. and LianJun, J.: Super-resolution based on improved sparse coding, *International Conference on Computer Application and System Modeling (ICCSM)*, Vol. 15, pp. V15–398–V15–401 (2010).
- [26] Dong, W., Zhang, D., Shi, G. and Wu, X.: Image Deblurring and Super-Resolution by Adaptive Sparse Domain Selection and Adaptive Regularization, *IEEE Transactions on Image Processing*, Vol. 20, No. 7, pp. 1838–1857 (2011).
- [27] Lu, X., Yuan, H., Yan, P., Yuan, Y. and Li, X.: Geometry constrained sparse coding for single image super-resolution, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1648–1655 (2012).