

オクルージョンや明るさ変化にロバストな超解像処理

矢口陽一 田中正行 奥富正敏

東京工業大学 大学院理工学研究科 機械制御システム専攻

{yag, mtanaka}@ok.ctrl.titech.ac.jp, mxo@ctrl.titech.ac.jp

超解像処理は、複数の低解像度画像間の精密な位置合わせ処理を必要とする。しかし、注目物体にオクルージョンが生じた場合やモーションモデルが不十分な場合などに、位置合わせが不正確になることがある。提案手法では、小領域ごとに類似度と位置合わせ誤差の再評価を行い、超解像処理に利用する画素を選択することによって位置合わせ誤差にロバストな超解像処理を実現する。また、明るさの補正も行い、より高画質な超解像処理結果を得る。提案手法の効果を実験によって確認した。

Robust Super-Resolution under Occlusion and Illumination Change

Youichi YAGUCHI, Masayuki TANAKA, Masatoshi OKUTOMI

Graduate School of Science and Engineering, Tokyo Institute of Technology

A super-resolution requires accurate registration between low-resolution images. However, occlusions or an insufficient motion model degrade the accuracy of the registration. The inaccurate registration yields visible artifacts in super-resolution images. We propose the robust super-resolution against the registration error. The proposed method verifies each pixel based on a similarity and a registration error. Then, the super-resolution image is reconstructed from only validated pixels. We also propose an illumination correction method to improve the super-resolution image. Experiments demonstrate effectiveness the proposed method.

1 はじめに

複数枚の低解像度画像から高解像度画像を再構成する超解像処理では、画像間の高精度な位置合わせが必要である。しかし、撮影された低解像度画像には、図1のようにオクルージョンが生じていることが多い。さらに、オクルージョンが生じている領域を位置合わせに利用した場合や、位置合わせの際に仮定したモーションモデルが不十分な場合などに、位置合わせが不正確になる。オクルージョンが生じている領域や位置合わせが不正確な領域の画素値を利用して超解像処理を行うと、得られる高解像度画像にはノイズ状の誤差が発生する。

また、オクルージョンに伴う影の発生や照明の変化によって、撮影対象の明るさが低解像度画像間で変化することも多い。明るさの変化した領域の画素値をそのまま利用して超解像処理を行うことも、得られる高解像度画像にノイズ状の誤差が発生する原因となる。

このような誤差を低減するため、画素や領域の類似度を利用して超解像処理に利用する画素を選択し、超解像処理を行う研究は従来から行われている。Ivanovski

らによる手法 [1] では、注目フレームと基準フレームの間の対応画素の画素値の差に基づく評価を行い、画素選択を行っている。Zhao らによる手法 [2] では、注目フレームと基準フレームの間の対応画素の周辺領域の相互相関に基づき、画素選択を行っている。戸田らによる手法 [3] では、注目フレームと基準フレームの間の対応画素の周辺領域の輝度値の差、移動量、フレーム間距離に基づいて超解像処理時の重みを決定している。

これらの従来手法では、類似度だけを用いて超解像処理に利用可能な画素の判定を行っているが、類似度による評価だけでは限界がある。

提案手法では、オクルージョンが生じている領域を判定するだけでなく、注目画素の周辺領域を用いてローカルモーションの再推定を行う。その結果、再推定された位置合わせ情報と元の位置合わせ情報との差分が小さければ、その画素を超解像処理に利用し、再推定されたローカルモーションも考慮し、超解像処理を行う。また、画素の明るさを補正し、より高画質な高解像度画像を得ることを目的とする。

なお、提案手法では既に求めた位置合わせ情報を

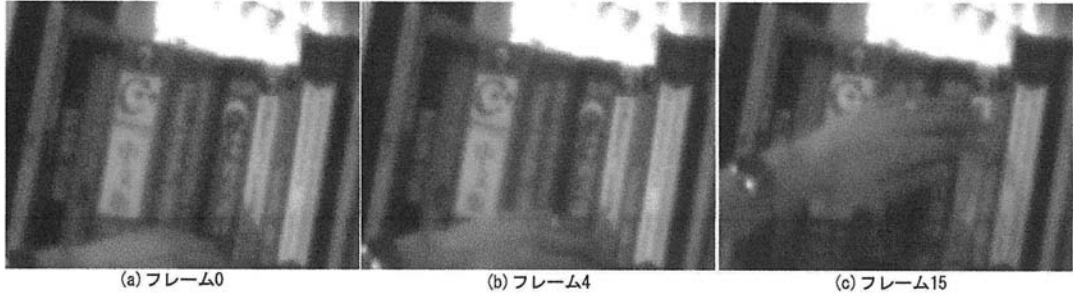


図 1: 入力時系列画像

用いるが、オクルージョンが生じている画像を位置合わせするためにはオクルージョンを考慮した位置合わせ手法を使うことが必要である。たとえば、張らによる手法 [4] では、位置合わせに適した画素を判定、選択し、選択された画素だけを利用して勾配法による位置合わせを行っている。

本論文の構成は次の通りである。2章で超解像処理の概要を説明し、3章で提案手法を説明する。4章で実験結果を示し、5章で結論を述べる。

2 超解像処理概要

超解像処理とは、複数枚の低解像度画像と画像間の位置合わせ情報が与えられたときに、高解像度画像を再構成する処理 [5] である。本章では、概要を説明する。

2.1 観測モデル

図 2 は、シーンがカメラに入力され、連続-離散変換、画素の運動や変形、ぼけやふれ、ダウンサンプリング、ノイズ付加を経て、時系列画像が観測される様子を示す。このモデルに基づき、求めたい高解像度画像と観測画像との関係を記述する。

低解像度時系列画像の k 枚目 ($k = 1, 2, \dots, K$) を $y_k(\mathbf{x}_l)$ 、高解像度画像を $z(\mathbf{x}_h)$ とする。ただし、 $\mathbf{x}_l = [x_l, y_l]^T$ および $\mathbf{x}_h = [x_h, y_h]^T$ は、低解像度画像および高解像度画像における画像座標を示す。それぞれの画素値をラスタスキャン順に並べた縦ベクトルを \mathbf{y}_k 、 \mathbf{z} とすると、観測モデルは以下のように定式化される。

$$\mathbf{y}_k = \mathbf{D}_k \mathbf{B}_k \mathbf{M}_k \mathbf{z} + \mathbf{n}_k \quad (1)$$

ここで、 \mathbf{D}_k はダウンサンプリング、 \mathbf{B}_k は PSF の影響、 \mathbf{M}_k は高解像度画像の各画素のモーションを表す行列、 \mathbf{n}_k は低解像度画像に含まれるノイズを表すベクトルである。

以上の $\mathbf{D}_k \mathbf{B}_k \mathbf{M}_k$ をまとめて \mathbf{A}_k とし、さらに $k =$

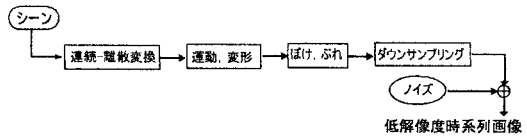


図 2: 観測モデル

$1, 2, \dots, K$ をまとめて記述すると以下ようになる。

$$\mathbf{y} = \mathbf{A}\mathbf{z} + \mathbf{n} \quad (2)$$

ただし、

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_K \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_K \end{bmatrix}, \quad \mathbf{n} = \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \\ \vdots \\ \mathbf{n}_K \end{bmatrix} \quad (3)$$

である。

2.2 高解像度画像の再構成

高解像度画像の再構成を行うためには、以下の評価関数 $E[\mathbf{z}]$ を最小化するような \mathbf{z} を求めればよい。

$$E[\mathbf{z}] = \|\mathbf{y} - \mathbf{A}\mathbf{z}\|^2 + f(\mathbf{z}) \quad (4)$$

$f(\mathbf{z})$ は拘束項と呼ばれ、一般的には高解像度画像に平滑化拘束を与える項や、色差を抑制する項を利用する [6]。

しかし、一般に行列 \mathbf{A} は非常に大きいので、解析的に \mathbf{z} の解を得ることは現実的ではない。そのため実際の超解像処理では、最急降下法などの繰り返し計算による最適化手法を利用する。

3 オクルージョンや明るさ変化への対応

3.1 超解像処理に利用可能な画素

式 (4) の評価関数を最小化する基本的な超解像処理では、各低解像度画像のすべての画素を超解像処理に利

用している。しかし、実際には超解像処理に利用すべきでない画素が存在する。超解像処理に利用すべきでない画素を超解像処理に利用すると、得られる高解像度画像にはノイズ状の誤差が発生する。

超解像処理に利用可能な画素を、以下の条件を満たす画素として定義する。

- オクルージョンが生じていない
- 位置合わせが正確である
- 光源の明るさが基準画像と等しい

オクルージョンが生じているか否かは、対応点の周辺領域との類似度を適当な閾値と比較することによって判定する。位置合わせの誤差は、対応点の周囲との類似度分布から推定されるローカルモーションを用いて修正する。ただし、元の位置合わせ情報の誤差が低解像度画像空間で 0.5 画素より大きければ、ローカルモーション推定結果の信頼度が低いと考え、超解像処理に利用しない。光源の明るさが変化した画素に関しては、明るさを補正して超解像処理に利用する。

3.2 ローカルモーション再推定と画素選択

まず、各画素ごとの位置合わせ誤差を評価するために、画素ごとにローカルモーションの再推定を行う。このローカルモーション再推定では、注目画素の近傍領域を利用し、対応点と 8 近傍の類似度を 2 次曲面にあてはめることにより、平行移動量を推定する。後段の処理で明度補正が行われることを考慮し、類似度には NCC (Normalized Cross Correlation) を利用した。次に、位置合わせ誤差が小さく、対応点との類似度が高い画素のみを選択して超解像処理に利用する。超解像処理に利用する画素を選択する処理を画素選択と呼ぶ。ローカルモーション再推定と画素選択について、具体的な方法を以下で述べる。

まず、ローカルモーション再推定を行うため、 $y_k(\mathbf{x}_i)$ と $y'_b(\mathbf{x}_i + \mathbf{i})$ との類似度の値 $R(\mathbf{x}_i, \mathbf{i})$ を利用する。ただし、 $y_k(\mathbf{x}_i)$ は入力フレーム、 $y'_b(\mathbf{x}_i)$ は基準フレームを変換して入力フレームに位置合わせした画像である。 $\mathbf{x}_i = [x_i, y_i]^T$ は注目画素の座標、 $\mathbf{i} = [i, j]^T$ は平行移動ベクトルである。このとき、明るさの補正を考えて、明度変化に影響されない類似度として次の NCC を利用する。

$$R(\mathbf{x}_i, \mathbf{i}) = \frac{\sum_{\mathbf{u} \in C(\mathbf{x}_i)} [y_k(\mathbf{u}) \times y'_b(\mathbf{u} + \mathbf{i})]}{\sqrt{\sum_{\mathbf{u} \in C(\mathbf{x}_i)} y_k(\mathbf{u})^2 \times \sum_{\mathbf{u} \in C(\mathbf{x}_i)} y'_b(\mathbf{u} + \mathbf{i})^2}} \quad (5)$$

ただし、 $C(\mathbf{x})$ は \mathbf{x} の近傍領域である。

実際には、 $y_k(\mathbf{x}_i)$ は RGB フルカラーであるので、RGB の各チャンネルを同時に用いる。

さらに、 $\mathbf{i} = [-1, -1]^T, [-1, 0]^T, [-1, 1]^T, [0, -1]^T, [0, 0]^T, [0, 1]^T, [1, -1]^T, [1, 0]^T, [1, 1]^T$ における $R(\mathbf{x}_i, \mathbf{i})$ を用いて、2 次曲面 $z(i, j) = ai^2 + bi + cj^2 + di + ej + f$ のパラメータの最小二乗解を求めることができる [7]。2 次曲面のパラメータがわかれば、ローカルモーションのパラメータ dx, dy を以下の式で求めることができる。

$$a < 0, c < 0, D < 0 \text{ のとき} \quad (6)$$

$$\begin{cases} dx = \frac{2cd - be}{D} \\ dy = \frac{2ae - bd}{D} \end{cases} \quad (7)$$

ただし、 $D = b^2 - 4ac$ である。一方、条件 (6) が満たされないとき、2 次曲面は上に凸な楕円放物面ではないので、その点の位置合わせは不正確であると考え、超解像処理には利用しない。

ここで、 $|dx| \leq 0.5, |dy| \leq 0.5$ であれば、求めた dx, dy によってモーションを補正して超解像処理に利用する。 $|dx| > 0.5$ または $|dy| > 0.5$ であるときは、位置合わせ誤差が大きく再推定されたモーションの信頼度が低いと考え、この画素は超解像処理に利用しない。また、ローカルモーションで補正した点における類似度は

$$z_{max} = \frac{ae^2 + cd^2}{D} + f \quad (8)$$

と推定できる。推定類似度が低い画素はオクルージョンが生じていると考えられるので、超解像処理に利用しない。

以上より、以下の条件

$$z_{max} \geq \kappa_0 \quad (9)$$

$$\begin{cases} |dx| \leq 0.5 \\ |dy| \leq 0.5 \end{cases} \quad (10)$$

がすべて満たされると、 $y_k(\mathbf{x}_i)$ を超解像処理に利用する。また、ローカルモーションの再推定結果を用いてモーションを修正する。条件 (9)、条件 (10) のいずれかが満たされない場合、 $y_k(\mathbf{x}_i)$ は超解像処理に利用しない。ただし、 κ_0 は調整パラメータである。

画素選択とモーションの修正を伴う超解像処理を行うために、評価関数 $E[\mathbf{z}]$ を以下のように再定義する。

$$E[\mathbf{z}] = \|\text{diag}(\mathbf{q})(\mathbf{y} - \mathbf{A}'\mathbf{z})\|^2 + f(\mathbf{z}) \quad (11)$$

ただし、 $\mathbf{A}' = [\mathbf{A}'_1^T \mathbf{A}'_2^T \dots \mathbf{A}'_K^T]^T$ 、 $\mathbf{A}'_k = \mathbf{D}_k \mathbf{B}_k \mathbf{M}'_k$ である。 \mathbf{M}'_k はローカルモーション再推

定結果を用いて修正されたモーシオン行列である。 $\text{diag}(\mathbf{q})$ は \mathbf{q} を対角要素にもつ対角行列を表し、 $\mathbf{q} = [\mathbf{q}_1^T \mathbf{q}_2^T \dots \mathbf{q}_K^T]^T$ の各要素 \mathbf{q}_k は $q_k(\mathbf{x}_i)$ の画素値をラスタスキャン順に並べた縦ベクトルである。 $q_k(\mathbf{x}_i)$ は、対応する画素 $y_k(\mathbf{x}_i)$ が超解像処理に利用されるならば 1、利用されないならば 0 が代入されるマスク画像である。

3.3 明度の補正

まず、照明についての観測モデルを仮定し、各画素の明度を求める。次に、注目画素の明度を対応画素の明度で置き換えることによって明度補正を行う。明度推定と明度補正について、具体的な方法を以下で述べる。

照明についての画像の観測モデルを式 (12) のように仮定する。

$$y(\mathbf{x}_i) = P(\mathbf{x}_i)L(\mathbf{x}_i) \quad (12)$$

ただし、 $y(\mathbf{x}_i)$ は観測画像、 $P(\mathbf{x}_i)$ は撮影シーンの反射率分布画像、 $L(\mathbf{x}_i)$ は撮影シーンに与えられた光源の分布画像である。

明度 $L(\mathbf{x}_i)$ は空間的に滑らかであることを仮定し、次式で求める。

$$L(\mathbf{x}_i) = \frac{\sum_{\mathbf{u} \in C_L(\mathbf{x})} y_V(\mathbf{x}_i + \mathbf{i})}{N(C_L(\mathbf{x}))} \quad (13)$$

ただし、 $y_V(\mathbf{x}_i)$ は $y(\mathbf{x}_i)$ を画素ごとに RGB 空間から HSV 空間へ色変換したときの V の画像、 $C_L(\mathbf{x})$ は \mathbf{x} の近傍領域、 $N(C_L(\mathbf{x}))$ は $C_L(\mathbf{x})$ 内の画素数を表す。ここでは、RGB の比が変化しない明度として、HSV 空間の V を用いている。RGB から V を求めるには、次式を計算すればよい。

$$V = \max\{R, G, B\} \quad (14)$$

V を V' に変化させたときの RGB の変化は次式で表される。

$$\begin{cases} R' = R \times \frac{V'}{V} \\ G' = G \times \frac{V'}{V} \\ B' = B \times \frac{V'}{V} \end{cases} \quad (15)$$

つまり、 V を変化させても RGB の比は不変である。

明度の推定結果を利用して、入力フレーム $y_k(\mathbf{x}_i)$ の明度 $L_k(\mathbf{x}_i)$ を、入力フレームに位置合わせした基準フレーム $y'_b(\mathbf{x}_i)$ の明度 $L'_b(\mathbf{x}_i)$ で置き換えることによ

て明度を補正する。次式で明度補正後の入力フレーム $y'_k(\mathbf{x}_i)$ を求める。

$$y'_k(\mathbf{x}_i) = y_k(\mathbf{x}_i) \times \frac{L'_b(\mathbf{x}_i)}{L_k(\mathbf{x}_i)} \quad (16)$$

明度補正を伴う超解像処理を行うために、評価関数 $E[\mathbf{z}]$ を以下のように再定義する。

$$E[\mathbf{z}] = \|\text{diag}(\mathbf{q})(\mathbf{y}' - \mathbf{A}'\mathbf{z})\|^2 + f(\mathbf{z}) \quad (17)$$

ただし、 $\mathbf{y}' = [y'_1{}^T y'_2{}^T \dots y'_K{}^T]^T$ である。 y'_k は $y'_k(\mathbf{x}_i)$ の画素値をラスタスキャン順に並べた縦ベクトルである。

4 実験結果

4.1 画素選択の精度評価実験

超解像処理に利用可能な画素を正確に判定できているか確認するため、シミュレーション画像を用いて実験を行った。図 3 に、実験に用いたシミュレーション時系列画像を示す。画像サイズは QVGA (320×240) であり、120 枚の画像を利用した。シミュレーション画像なので位置合わせ情報、オクルージョンの生じている領域は既知である。オクルージョンとして木箱を配置した。さらに、固定光源と移動光源を配置して明るさの変化を加えた。固定光源が右下にあり、移動光源が左下から右上へ移動している。固定光源によって木箱の左上に濃い影が発生しているのが確認できる。120 枚すべての画像に対し、真値を利用して位置合わせを行い、200×130 の領域を切り取り、画素選択を行った。ただし、奇数番号のフレームには真値に対して 1 画素のずれを画像全体に加えた。なお、類似度計算時の近傍領域は 15×15、明度推定時の近傍領域は 5×5 に設定した。 κ_0 は、超解像処理に利用可能な画素が正確に判定される割合が各判定方法で等しくなるように調整した。

図 4 に、基準画像と、位置合わせを行った入力画像を示す。フレーム 89 は、画像全体を 1 画素ずらしたため、位置合わせに誤差が含まれている。図 5 に、図 4 の各画像に対して明るさ補正と画素選択を行った結果を示す。斜線の部分が、画素選択の結果、超解像処理に利用されなかった画素を示す。上から順に、類似度による判定 (式 (9)) のみで画素選択を行った結果、位置ずれ量による判定 (式 (10)) のみで画素選択を行った結果、類似度、位置ずれ量の両方 (式 (9) と式 (10)) を評価して画素選択を行った結果である。

図 5(b) は、位置合わせに誤差の含まれる画像に対し、類似度による判定のみ (従来手法) で画素選択を行った結果である。本来すべての画素が超解像処理に利用不

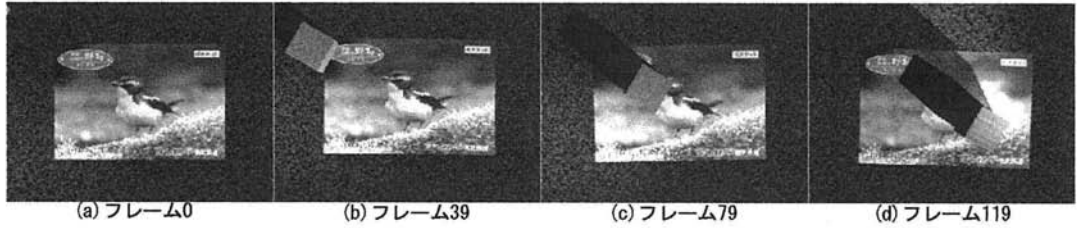


図 3: 画素選択の精度評価実験. 入力時系列画像.

可能な画素だが、位置ずれの起きている領域を判定できていないことがわかる。同図 (d) より、位置合わせ誤差による判定の結果、位置合わせが不正確な領域を比較的正確に判定できていることがわかる。しかし、同図 (c) より、位置合わせ誤差による判定のみではオクルージョンである木箱の一部を判定できていないことがわかる。同図 (e),(f) より、類似度による判定と位置合わせ誤差による判定を組み合わせる（提案手法）と、位置合わせに誤差が含まれる領域とオクルージョンの生じている領域の両方を判定できていることが確認できる。

120 枚すべての画像に対して画素選択を行った結果、超解像処理に利用不可能な画素が正確に判定された割合を以下に示す。従来手法である類似度による判定のみでは 45.1%，位置ずれ量による判定のみでは 98.2%，提案手法である類似度と位置ずれ量の両方による判定では 98.9% が正確に判定され、超解像処理に利用されなかった。各判定方法とも、超解像処理に利用可能な画素が正確に判定された割合は 88% である。この結果からも、類似度の評価と位置ずれ量の評価を併用することによる効果が確認できる。

4.2 実画像実験

オクルージョンが生じている実画像に対して提案手法を適用した結果を示す。図 1 に、実験に用いた時系列画像を示す。画像サイズは 160×120 であり、30 枚の画像を利用した。 40×30 の領域を切り取り、 4×4 倍の超解像処理を行った。類似度計算時の近傍領域は 15×15 ，明度推定時の近傍領域は 5×5 ， $\kappa_0 = 0.993$ に設定した。位置合わせには勾配法 [8] を利用した。本棚の本に注目したとき、手によってオクルージョンが生じていることが確認できる。

提案手法の超解像処理結果を図 6 に示す。また、比較のために一部分を拡大した画像を図 7 に示す。高解像度化処理には、田中らの手法 [5] を利用した。図 7 は上から順に、原画像をニアレストネイバー補間によって拡大した画像、オクルージョンの生じている領域や位置合わせの不正確な領域、明るさの変化を考慮しな

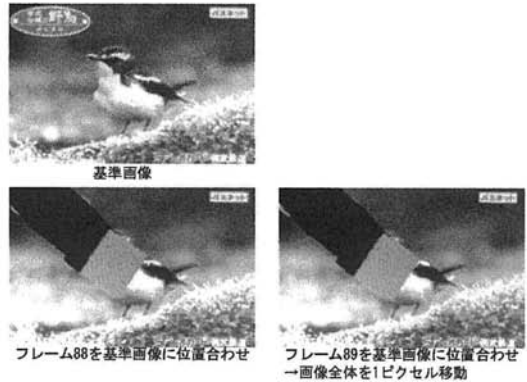


図 4: 画素選択の精度評価実験. 基準画像と位置合わせを行った入力画像.

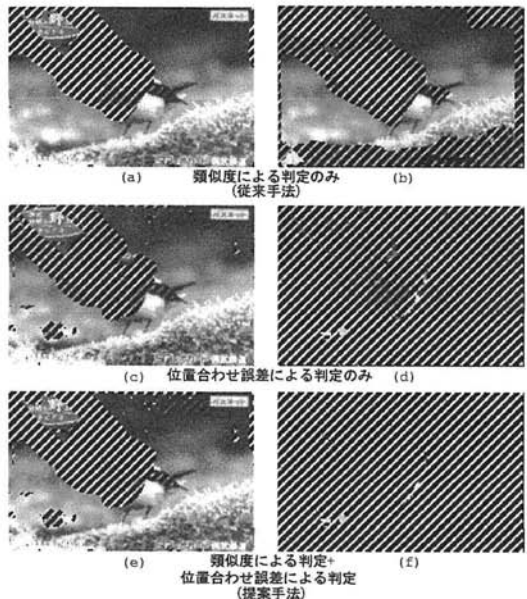


図 5: 画素選択の精度評価実験. 判定方法による画素選択結果の違い. 斜線部は画素選択の結果、超解像処理に利用されなかった領域を示す。

(a)(c)(e) フレーム 88. 木箱部分がオクルージョン領域.
(b)(d)(f) フレーム 89. 画像全体で位置合わせが不正確.

い従来の超解像処理を行った結果、提案手法による画素選択と明るさ補正を行った後に超解像処理を行った結果である。

画素選択を行わない従来の超解像では、手によるオクルージョンの影響によってノイズ状の誤差が発生している。画素選択を行う提案手法を用いると、原画像や従来の超解像処理では読めなかった部分が読めるようになり、全体の解像感も向上し、ノイズ状の誤差も生じていないことがわかる。

5 まとめ

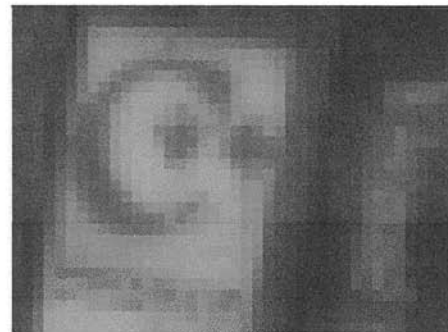
本論文では、入力画像と基準画像間の対応画素の周辺領域の類似度と、類似度分布から推定される位置合わせの誤差の大きさを評価し、超解像処理に使用する画素を選択する手法を提案した。また、推定された位置合わせ誤差を利用し、位置合わせ情報を修正する手法を提案した。画素の明度を基準画像上の対応画素に合うよう補正する手法を提案した。これらの画素選択、位置合わせ情報の修正、明るさの補正を適用した超解像処理を行うことで、得られる高解像度画像のノイズ状の誤差を軽減、画質を改善できることを実験によって確認した。

参考文献

- [1] Z. A. Ivanovski, L. Panovski, and L. J. Karam. "Robust super-resolution based on pixel-level selectivity". *Proceedings of SPIE*, Vol. 6077, p. 607707, 2006.
- [2] W. Zhao and H. Sawhney. "Is super-resolution with optical flow feasible?". *Proc. ECCV*, Vol. 1, pp. 599-613, 2002.
- [3] 戸田真人, 塚田正人, 井上晃. "レジストレーション誤差を考慮した超解像処理". *Proc. FIT 2006*, Vol. 1, pp. 63-64, 2006.
- [4] 張馴権, 清水雅夫, 奥富正敏. "領域選択を伴う 2 段階レジストレーション". *電子情報通信学会論文誌 D*, Vol. J90-D, No. 2, pp. 514-525, 2007.
- [5] 田中正行, 奥富正敏. "周波数領域最適化法による MAP 型超解像処理の高速化". *情報処理学会論文誌: コンピュータビジョンとイメージメディア*, Vol. 47, No. SIG10(CVIM15), pp. 12-22, July 2006.
- [6] 後藤知将, 奥富正敏. "単板カラー撮像素子の RAW データを利用した高精細画像復元". *情報処理学会論文誌*, Vol. 45, pp. 15-24, 2004.
- [7] C. Sun. "Fast algorithms for stereo matching and motion estimation". *Proc. of Australia-Japan Advanced Workshop on Computer Vision*, pp. 38-48, 2003.
- [8] S. Baker and I. Matthews. "Lucas-Kanade 20 Years On: A Unifying Framework". *International Journal of Computer Vision*, Vol. 56, No. 3, pp. 221-255, 2004.



図 6: 実画像実験. 提案手法の超解像処理結果.



元画像を拡大



通常の超解像処理



提案手法の超解像処理

図 7: 実画像実験. 4×4 倍に拡大した画像.