

## 低解像度人体画像の学習型超解像に関する検討

西堀 研人<sup>†</sup> 高橋 友和<sup>††</sup> 出口 大輔<sup>†</sup> 井手 一郎<sup>†</sup> 村瀬 洋<sup>†</sup>

<sup>†</sup> 名古屋大学 大学院情報科学研究科 〒464-8601 愛知県名古屋市千種区不老町  
<sup>††</sup> 岐阜聖徳学園大学 経済情報学部 〒500-8288 岐阜県岐阜市中鷺1丁目38番地  
<sup>†††</sup> 名古屋大学 情報連携統括本部 〒464-8601 愛知県名古屋市千種区不老町

E-mail: <sup>†</sup>nishiborik@murase.m.is.nagoya-u.ac.jp,

<sup>†</sup>{ide,murase}@is.nagoya-u.ac.jp, <sup>††</sup>ttakahashi@gifu.shotoku.ac.jp, <sup>†††</sup>ddeguchi@is.nagoya-u.ac.jp

**あらまし** 防犯対策を目的として監視カメラ映像から人物照合を行う技術が求められている。本研究では、低解像度の人体画像を高解像度化することで、人が見ても機械が処理をしても人物照合をより正確に行うことができる高解像度の人体画像を生成することを目的とする。大量の人物画像からなる学習用画像を用いて、低解像度画像から高解像度画像を生成する学習型超解像方式を適用し、人体画像の超解像について検討を行った。

**キーワード** 学習型超解像, 人体画像, 高周波成分, 監視カメラ, 人物照合, 画質評価

## A Study on Exemplar Based Super Resolution for Low Resolution Images of Human Body

Kento NISHIBORI<sup>†</sup>, Tomokazu TAKAHASHI<sup>††</sup>, Daisuke DEGUCHI<sup>†</sup>, Ichiro IDE<sup>†</sup>, and  
Hiroshi MURASE<sup>†</sup>

<sup>†</sup> Graduate School of Information Science, Nagoya University  
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601 Japan

<sup>††</sup> Faculty of Economics and Information, Gifu Shotoku Gakuen University  
Nakauzura 1-38, Gifu, Gifu, 500-8288 Japan

<sup>†††</sup> Information and Communications Headquarters, Nagoya University  
Furo-cho, Chikusa-ku, Nagoya, Aichi, 464-8601 Japan

E-mail: <sup>†</sup>nishiborik@murase.m.is.nagoya-u.ac.jp,

<sup>†</sup>{ide,murase}@is.nagoya-u.ac.jp, <sup>††</sup>ttakahashi@gifu.shotoku.ac.jp, <sup>†††</sup>ddeguchi@is.nagoya-u.ac.jp

**Abstract** The purpose of this study is to generate a human body image in LR to a high-resolution (HR) image for enabling both a human and a computer system to conduct the identification process more accurately. As a result of experiments, we confirmed the efficiency of the proposed super-resolution method.

**Key words** Exemplar-based super-resolution, Human body image, High-frequency component, Surveillance system, Image quality assessment

### 1. ま え が き

防犯対策や事件の解決のための監視カメラシステムの利用への期待が高まっており、数多くの監視カメラが空港、駅、繁華街のような人の往来が頻繁な公共の場所に配置され、不審者の検出や人物照合等に用いられている [?, ?, ?]. このような広域を監視するカメラを用いた場合、人物は低解像度で撮影されることが多く人物の照合が困難となる。低解像度画像を高解像度化する技術として、超解像技術が注目されている。超解像技術

は、低解像度画像の失われた高周波成分を復元することで、鮮明な高解像度画像を生成する手法であり、近年様々な手法が提案されている [?, ?, ?, ?, ?]. 人物照合のために、顔画像の超解像に関する報告もなされており、顔の各パーツの位置や形状を手がかりとした顔画像の超解像が行われている [?, ?, ?, ?, ?]. しかし、監視カメラにより人物の顔を適切に撮影することは難しい。そのため、顔だけでなく、人物全体の外観から照合する事が必要となる。しかし、人体の画像を対象とした場合は、姿勢や服装のように外見の変化が大きく超解像のための手がかり

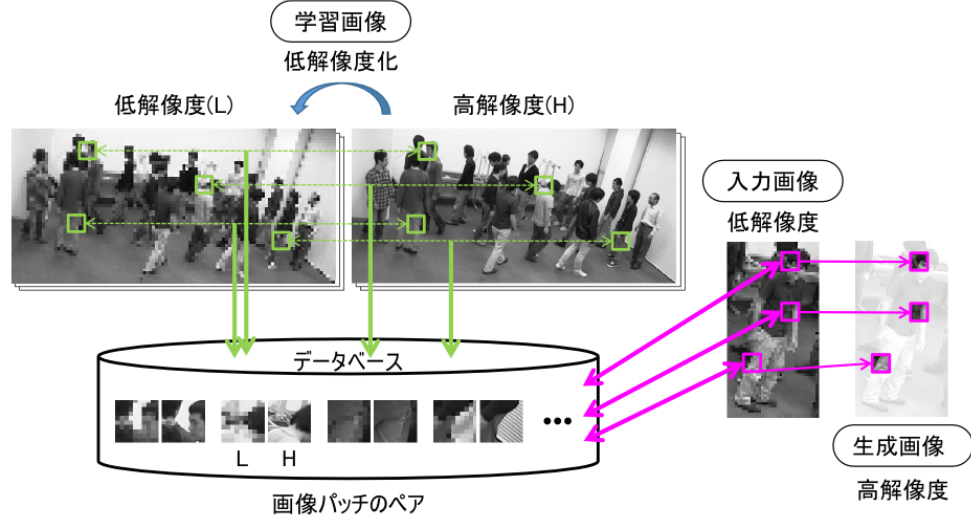


図1 学習型超解像による高解像度画像の生成

を用いるのが困難である。そこで、本研究では低解像度の人体画像を高解像度化することで、人が見ても機械が処理をしても人物照合をより正確に行うことができる高解像度の人体画像を生成することを目的とする。ここでは、大量の人物画像からなる学習用画像を用いて低解像度画像から高解像度画像を生成する学習型超解像方式を適用し、人体画像の超解像について検討を行う。

## 2. 人体画像への超解像

### 2.1 学習型超解像

図??に、本研究で用いた学習型超解像の基礎となる高解像度画像の生成手順について示す[?]. 学習型超解像は以下の手順で行う。

(i) 学習用の高解像度画像を複数用意し、それらを低解像度化することで低解像度画像を生成する。

(ii) 高解像度画像と低解像度画像から位置関係を保ったままパッチ領域をそれぞれ切り出し、高解像度パッチ画像と低解像度パッチ画像のペアとしてデータベースに登録する。

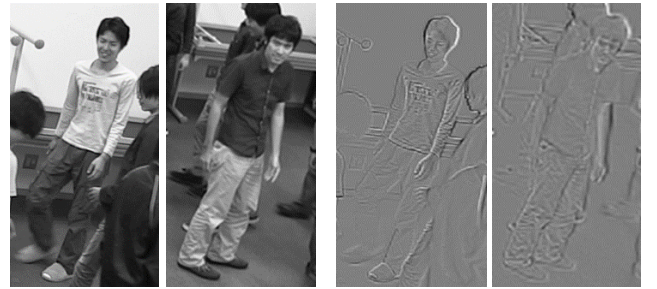
(iii) 入力画像である低解像度人体画像から切り出したパッチ画像とデータベース内で最も類似する低解像度パッチ画像を探索する。

(iv) 入力画像である低解像度人体画像から切り出したパッチ画像を (iii) において探索された低解像度パッチ画像とペアとなる高解像度パッチ画像と置き換え、これを入力画像のすべてのパッチ領域に行い超解像画像を生成する。

### 2.2 高周波成分を用いた超解像

図??は、高周波成分を用いた学習型超解像手法である。低解像度画像の失われた高周波成分を復元するために学習データの高周波成分を用い、高解像度画像を生成する。

(i) 学習用画像を用意し、順次低解像度化高周波成分  $X_k$  は、高解像度画像  $I_k^{(0)}$  と低解像度画像  $I_k^{(1)}$  の差から次のように求められる。



(a) 人物の服装による外観の違い

(b) 人物画像の高周波成分

図2 異なる服装における高周波成分

$$X_k = I_k^{(1)} - I_k^{(0)} \quad (1)$$

$$I_k^{(1)} = D(I_k^{(0)})$$

高周波成分  $Y_k$  も同様に、低解像度画像  $I_k^{(1)}$  とそれをさらに低解像化した低解像度画像  $I_k^{(2)}$  の差から次のように求められる。

$$Y_k = I_k^{(2)} - I_k^{(1)} \quad (2)$$

$$I_k^{(2)} = D(I_k^{(1)})$$

ここで、 $k$  は学習用画像のフレーム番号であり、 $D(\cdot)$  は拡大率  $r$  による画像の低解像度化を表す。

(ii) 高周波成分のパッチペア  $x_l, y_l$  をサイズ  $rL \times rL$  [pixels],  $L \times L$  [pixels] で位置関係を保ったまま、高周波成分  $X_k, Y_k$  よりそれぞれ切り出し、データベースに登録する。ここで、 $l$  はパッチ番号とする。パッチペア  $x_l, y_l$  は、それぞれサイズが  $rL \times rL$  [pixels],  $L \times L$  [pixels] のガウシアンフィルタによって重み付けされ、データベースに登録される。

(iii) 入力画像とそれを低解像度化した画像間の差より高周波成分  $Z$  を算出高周波成分  $Z$  から切り出したパッチ  $z_j$  とデータベース内で最も類似するパッチ  $y_l$  を探索する。

$$Z = I_i^{(2)} - I_i^{(1)} \quad (3)$$

$$I_i^{(2)} = D(I_i^{(1)})$$

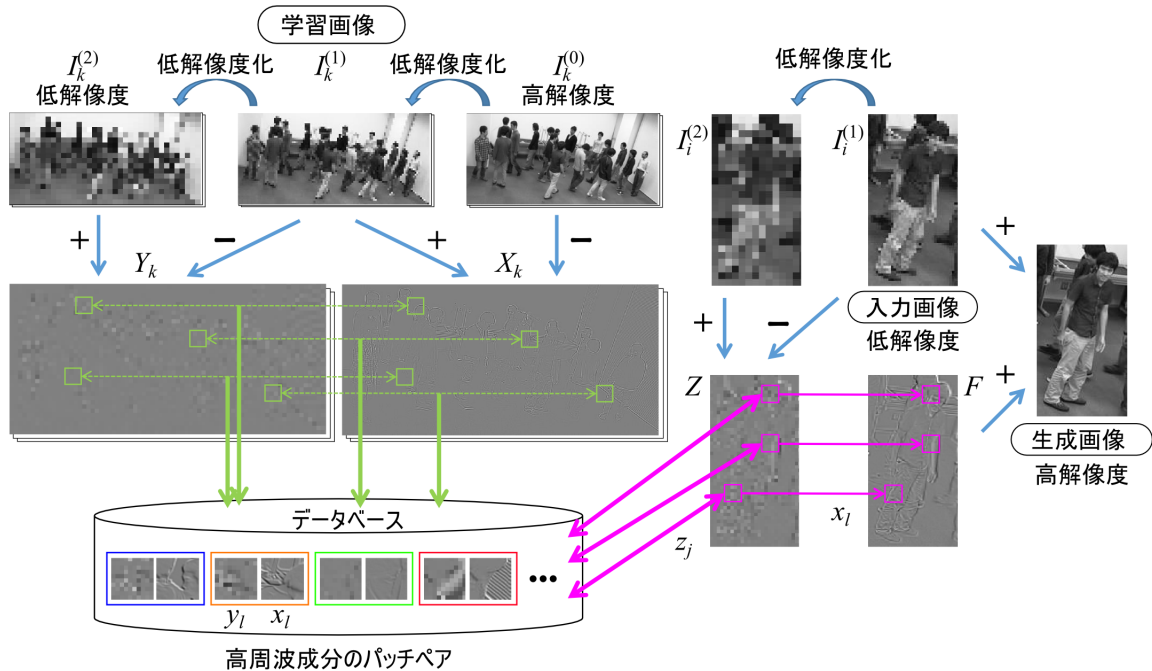


図3 高周波成分を用いた学習型超解像 (提案手法)

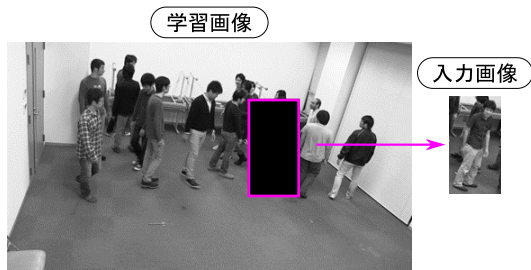


図4 学習画像と入力画像

(iv) 高周波成分のパッチ  $z_j$  を (iii) において探索されたパッチ  $y_l$  のペア  $x_l$  と置き換え、これを入力画像のすべてのパッチ領域に行い超解像画像を生成する。

### 3. 実験

#### 3.1 実験条件

提案手法の有効性を確認するために、異なる手法により生成した高解像度画像の比較を行う。本研究では、6名の低解像度人体画像に対し、1名ずつ高解像度画像を生成し画質を評価する。学習画像と入力画像には、照明による見えの変動を少なくするため、同一照明下に17人の人物が写った画像 (1920×1080 pixels) を10枚使用する。図??に示すように、入力画像にはある人物の領域画像を用い、学習画像には入力画像の人物を含まない残りの領域画像を用いる。データベースの大きさは、拡大率3倍、パッチのサイズを7×7 pixels のとき200万組である。データベース内の探索には、ランダム kd-tree による近似最近傍探索を用い、画像の特徴量として明度と提案手法である高周波成分を使用する。近似最近傍探索には、OpenCV ライブラリの FLANN (Fast Approximate Nearest Neighbor Search Library) を利用する。生成された超解像画像の画質を評価する

表1 SSIMによる画質評価

手法	低解像度入力	バイキュービック	明度	提案手法
SSIM	0.855	0.900	0.912	0.928

ために、SSIM (Structural Similarity) を用いる [?]. SSIM は画像のノイズに対して、画質評価尺度として用いられる他手法の平均二乗誤差 (MSE) やピーク信号対雑音比 (PSNR) に比べ、人の視覚認識に近いとされる。

#### 3.2 生成画像の画質評価

図??に低解像度人物画像に対する高解像度画像の生成実験を示す。低解像度人体画像に対する元の高解像度画像の拡大率は3倍、パッチサイズは7×7 pixels とする。図??(a)-(i) は元の高解像度画像であり、画質評価実験の真値とする。画像サイズは246×462 pixels であり、拡大率3倍で低解像度化することで低解像度入力画像 (82×154 pixels) を作成する。図??(a)-(ii) は、低解像度入力画像を単純に拡大した画像であり、図??(a)-(iii) は拡大するときにバイキュービック補間を行った結果である。図??(a)-(iv) はパッチを切り出すときに明度値を用い、図??(a)-(v) は高周波成分を用いた結果である。図??(a)における矩形で囲まれた領域を拡大し、図??(b)に示す。提案手法の画質が、他手法による高解像度化画像よりも優れていることがわかる。図??(c) は SSIM による画質評価値の分布であり、元の高解像度と類似度が高い領域ほど明るく表示される。

表??は、6名の低解像度人物画像を拡大率3倍で高解像度化したときに、単純な画像の拡大によって生成された画像、バイキュービック補間による画像の拡大によって生成された画像、パッチに明度、または提案手法である高周波成分を用いた超解像によって生成された画像を元の高解像度画像を真値として SSIM を求めたときの平均値である。



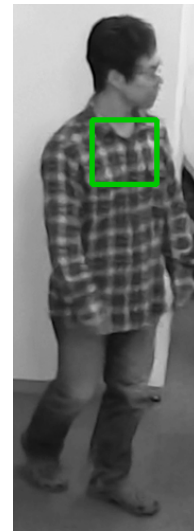
(i) 元の高解像度画像



(ii) 低解像度入力画像



(iii) バイキュービック  
(a) 人物画像



(iv) 明度



(v) 提案手法



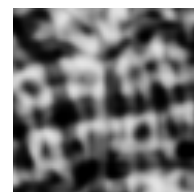
(b) 人物画像中の矩形領域の拡大図



(SSIM = 0.856)



(SSIM = 0.907)



(SSIM = 0.887)



(SSIM = 0.912)

(c) 画質評価における SSIM の分布

図 5 異なる手法により生成した高解像度画像の比較

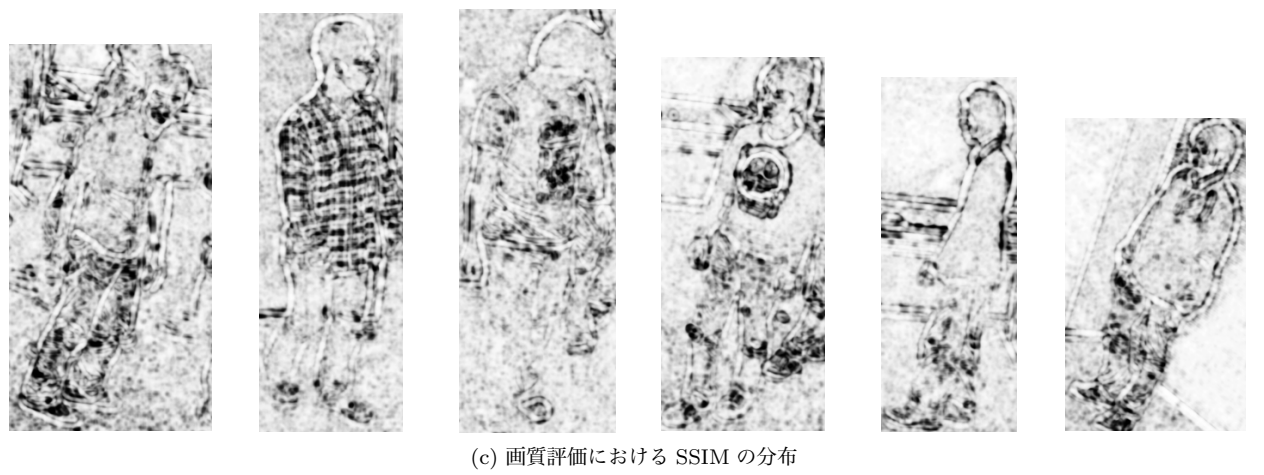
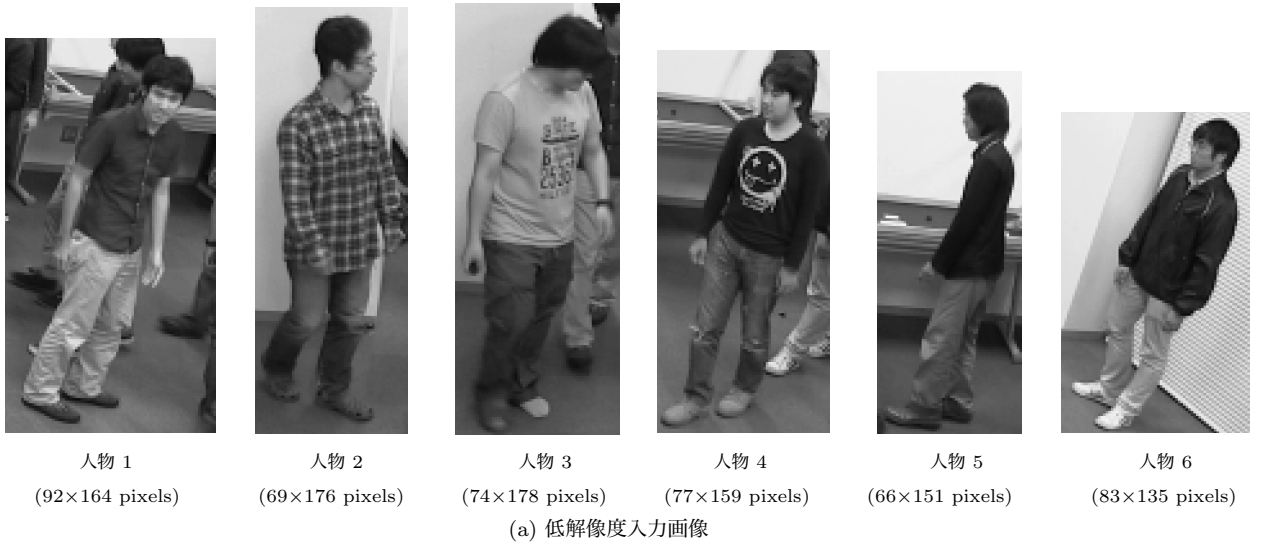


図 6 提案手法によって生成した高解像度画像

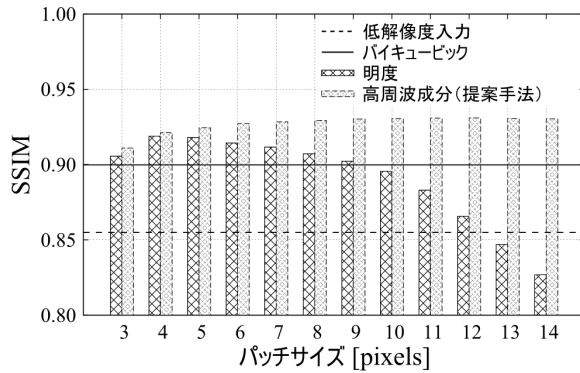


図7 パッチサイズの変化における SSIM

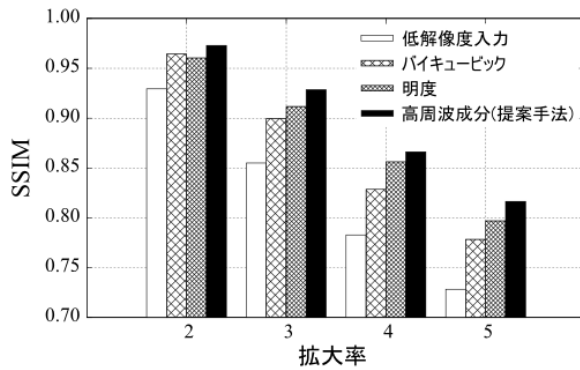


図8 拡大率の変化における SSIM

図??(a)は6名の低解像人物画像であり、図??(b)は高周波成分を用いた超解像画像である。図??(c)は高周波成分を用いた超解像画像の SSIM の分布を表しており、色が濃い場所ほど元の高解像度画像との差が大きい。人物の顔や輪郭、服装の柄のような高周波成分を含む領域において差が大きいことがわかる。

### 3.3 超解像における条件の変更

図??は、図??(a)の6名の低解像度人物画像をパッチサイズを $3 \times 3 \sim 14 \times 14$  pixelsとして生成された超解像画像のにより生成された高解像度画像 SSIM によって求めた平均値である。高周波成分を用いた超解像画像では、パッチサイズが変化しても、画質への影響が少ないことがわかる。

図??は、6名の低解像度人物画像に対して拡大率を2倍～5倍に変化させて高解像度化した場合の画質評価値の平均である。拡大率を変えてもバイキュービック補間に比べて、高周波成分を用いた超解像画像の画質が良いことがわかる。

## 4. まとめ

監視カメラ映像における人体画像を高解像度化することで、人と機械の双方が人物照合しやすい画像を生成することを目的に、低解像度人体画像に対する高周波成分を用いた学習型超解像について検討を行った。実験の結果、人物の服装や姿勢が変化しても、低解像度人体画像の高周波成分を復元でき、画質を向上させることができた。異なる拡大率(2倍～5倍)においても、キュービック補間法により生成した拡大画像よりも、提案手法により生成した超解像画像の方が SSIM による画質評価値

において良いことを確認した。

今後の課題としては、学習型超解像に時系列情報を利用してより正確で頑健な超解像を行うこと、空港、駅、街路等の実環境において撮影された監視カメラ映像中の低解像度人体画像の超解像を行うことである。

**謝辞** 日頃より熱心に御討論頂く名古屋大学村瀬研究室諸氏に深く感謝する。本研究の一部は、文部科学省先導的創造科学技術開発費補助金、および科学技術研究費補助金による。