

監視カメラ映像における人物の学習型超解像

Exemplar-Based Super-Resolution of Human Body Image in Surveillance Video

○西堀 研人¹ 高橋 友和^{1,2} 出口 大輔¹ 井手 一郎¹ 村瀬 洋¹
Kento Nishibori¹, Tomokazu TAKAHASHI^{1,2}, Daisuke DEGUCHI¹,
Ichiro IDE¹ and Hiroshi MURASE¹

¹ 名古屋大学 ² 岐阜聖徳学園大学

○ nishiborik@murase.m.is.nagoya-u.ac.jp

概要: 本研究では、防犯や事件の解決を目的として設置されている広域監視カメラシステムのための低解像度人体画像の学習型超解像について検討を行った。服装や姿勢の変化による人物の見えの変動は大きく、全てを網羅する高解像度の学習データを用意するのは困難であるが、予め人物を様々な方向から高解像度で撮影した学習用画像を実際の監視カメラ映像に適用することで、人物画像の画質を向上させることができた。

<キーワード> 監視カメラ, 学習型超解像, 人体画像, 高周波成分

1 はじめに

近年、空港や駅、人が多く集まる街中のような公共の場において、防犯を目的とした監視カメラの設置が増えてきており、不審者の検出や人物照合等に用いられている [1, 2, 3]. 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 [4, 5, 6]. しかし、人体の画像は姿勢や服装のように見えの変動が大きく超解像が困難である。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。

超解像の主な手法の1つである複数枚超解像は、複数の低解像度画像をサブピクセル以下で高精度に位置合わせを行うことで高解像度画像を生成する手法である。多くの監視カメラシステムは長時間撮影するため保存容量の制約からフレームレートを下げ、画像データを圧縮している。そのため、動いている人物のフレーム間の位置ずれが大きく、高精度に位置合わせを行うことは困難である。そのため、もう一つの主な手法である大量の人物画像からなる学習用画像を用いて、1枚の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



図 1: 色や明度の多様性への対応

式を適用し、人体画像の超解像について検討を行う [7].

2 人体画像の超解像

本研究で用いる学習型超解像は、超解像の性能が学習に用いる画像に依存する。しかし、人体のように姿勢や服装による見えの多様性が大きい場合には、膨大な学習用画像の用意が困難となる。そこで、図 1 に示すように明度と色情報を分離し、さらに明度の高周波成分を超解像に用いることで色や明度の多様性に対応する。

図 2 に高周波成分を用いた学習型超解像の処理を示す [8]. まず、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した低解像度画像との差を求め高解像度の高周波成分を得る。さらに低

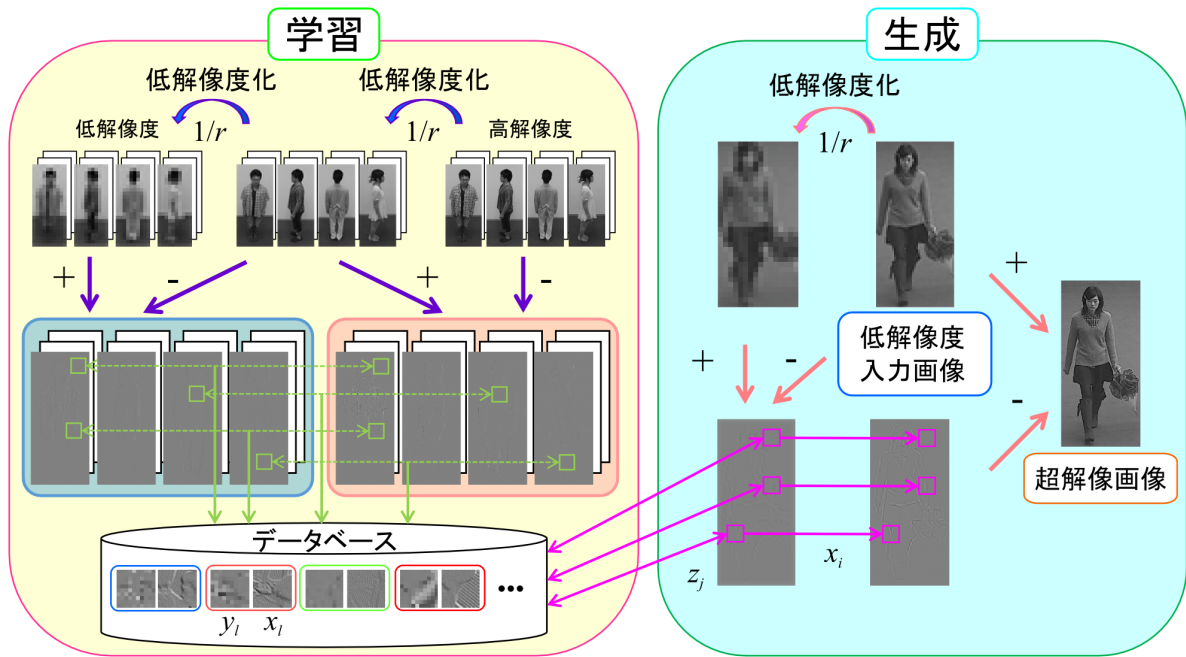


図 2: 高周波成分を用いた超解像

解像度度画像とそれらを低解像度化して得られる低解像度画像との差により低解像度の高周波成分を得る。高解像度と低解像度の高周波成分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。低解像度の入力画像においてもさらに低解像度化した画像との差により低解像度の高周波成分を算出し、それ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。探索に最も類似し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、この処理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。

学習画像と入力画像の撮影条件が異なる場合に、それぞれから切り出されたパッチ画像の高周波成分のばらつきがパッチ画像間の照合や低解像度パッチ画像を高解像度パッチ画像に置き換える際に影響を与える可能性が考えられる。そこで、学習画像と入力画像の高周波成分のパッチ画像に対して、学習画像の低解像度パッチ画像を基準とした正規化と学習画像の高解像度パッチ画像を基準とした正規化の2通りの手法を適用する [9]。

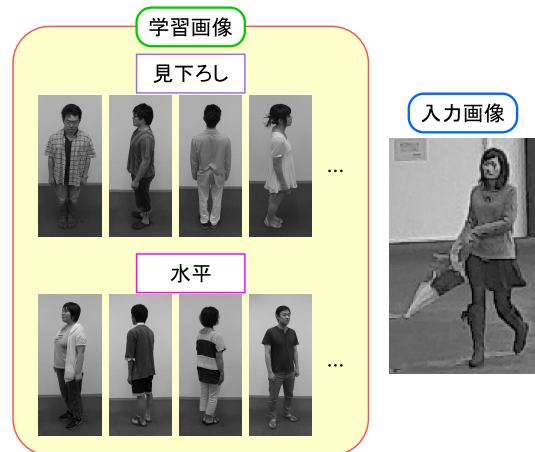


図 3: 学習画像と入力画像

3 実験

提案手法の有効性を確認するため、学習画像とはカメラや照明等が異なる条件で撮影された監視カメラ映像中の人物の超解像を行う。図 3 に示すように、図 3 に示すように、学習用画像には入力画像と異なるカメラや照明条件において、延べ 47 名の人物がカメラに対して 45 度ずつ向きを変えた 8 方向、および人物を見下ろす視点と水平方向の

視点の2つの視点から撮影した画像の計752枚を使用する。

ここで、人物の入力画像には(I)監視カメラ映像中の比較的高解像度の画像を低解像度化することにより生成した場合、(II)監視カメラ映像中の低解像度の画像を使う場合の2通りの実験を行った。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る。近似最近傍探索には、OpenCV ライブラリのFLANN(Fast Library for Approximate Nearest Neighbors)を利用する。拡大率は3倍、パッチサイズは7×7画素とした。

(I) 入力画像を低解像度化により生成した場合

画像の特徴量として、次の8つの手法を比較する。(i)最近傍補間、(ii)Bicubic、(iii)明度、(iv)高周波成分、明度の正規化に(v)低解像度パッチを用いた場合、(vi)高解像度パッチを用いた場合、明度の正規化に(vii)低解像度パッチを用いた場合、(viii)高解像度パッチを用いた場合である。生成された超解像画像の画質評価には、PSNR(Peak signal-to-noise ratio)および画像のノイズに対して人の視覚認識に近いとされるSSIM(Structural Similarity)を用いる[10]。画質評価のための参照画像として元の高解像度画像を用いる。

表1は、各手法によって生成された画像をPSNRにより画質評価した結果である。高周波成分の正規化を学習画像の低解像度パッチを基準として行った超解像結果のPSNR値が最も高くなっている。表2は、各手法によって生成された画像をSSIMにより画質評価した結果である。PSNRによる評価に比べ、高周波成分を用いた超解像結果が最も高い値を示している。PSNRでは画素ごとのパッチの明度差を算出しているのに対し、SSIMでは画素間の構造も考慮しているためと考えられる。

図4に高解像度化した画像の比較を示す。学習画像とは異なる条件で撮影された画像において、図4(a)は5名の人物の低解像度入力画像であり、図4(b)に示す超解像画像の画質が入力画像に比べ向上していることがわかる。図4(d)SSIMマップは図4(c)に示す元の高解像度画像を参照画像として、超解像画像との差異をSSIMマップにて表示したものである。色が薄い場所ほど類似度が高く、色の濃い場所は差異が大きいの。細かい柄ほど差異が大きいのことがわかる。

(II) 入力画像に監視カメラ映像を用いた場合

図5は監視カメラ映像中の5名の低解像度人物画像に対して高周波成分を用い、超解像した結果である。低解像度入力画像の最近傍補間に比べて、見た目の画質が向上していることがわかる。

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、低解像度人物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。今後の課題としては、学習型超解像に時系列情報を用いてより正確な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。

参考文献

- 1) 岩下 友美, Adrian Stoica, 倉爪 亮: 影に着目した歩行画像からの個人識別, 画像の認識・理解シンポジウム (MIRU 2010), IS2-13, 2010.
- 2) 鷗沼 正人, 長峰 望: 異常行動検知のための画像処理手法の開発, 鉄道総研報告, Vol. 21, No. 11, pp. 17-22, 2007.
- 3) Lutz Goldmann, Mustafa Karaman, J. T. Saez Minquez, and Thomas Sikora: Appearance-Based Person Recognition for Surveillance Applications, 7th Int'l Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2006), 2006.
- 4) Junjun Jiang, Ruimin Hu, Zhen Han, Tao Lu, and Kebin Huang: Position-Patch Based Face Hallucination via Locality-Constrained Representation, Int'l Conf. Multimedia and Expo (ICME 2012), pp. 212-217, 2012.
- 5) 田中 正行, 沼田 宗敏, 奥富 正敏, 奥水 大和: 超空間解像・超階調解像技術, 精密工学会誌, Vol.77, No.12, pp. 1099-1103, 2011.

表 1: PSNR による画質評価 (単位 : dB)

人物	最近傍補間	Bicubic	明度	高周波成分	明度の正規化		高周波成分の正規化	
					低解像度	高解像度	低解像度	高解像度
1	28.75	29.35	28.66	28.94	28.70	28.96	29.18	28.54
2	29.97	30.42	29.97	30.56	30.14	30.59	30.68	30.21
3	30.07	30.89	30.23	30.52	30.07	30.72	30.62	30.25
4	23.44	23.62	23.47	23.67	23.20	23.00	23.87	23.45
5	26.90	27.56	26.93	27.49	26.89	26.82	27.61	27.21
平均	27.83	28.37	27.85	28.24	27.80	28.02	28.39	27.93

表 2: SSIM による画質評価

人物	最近傍補間	Bicubic	明度	高周波成分	明度の正規化		高周波成分の正規化	
					低解像度	高解像度	低解像度	高解像度
1	0.758	0.763	0.729	0.769	0.743	0.741	0.749	0.757
2	0.787	0.788	0.759	0.802	0.772	0.771	0.787	0.799
3	0.787	0.793	0.760	0.794	0.770	0.770	0.768	0.781
4	0.624	0.618	0.593	0.643	0.599	0.594	0.626	0.633
5	0.703	0.708	0.660	0.727	0.670	0.665	0.709	0.720
平均	0.732	0.734	0.700	0.747	0.711	0.708	0.728	0.738

- 6) 陳 延偉 : 学習型画質改善・画像復元・超解像技術, Optics Electronics, Vol. 32, No. 3, pp. 270–274, 2010.
- 7) Daniel Glasner, Shai Bagon, and Michal Irani: Super-resolution from a single image, In Proc. 12th Int'l Conf. Computer Vision, pp. 349–356, 2009.
- 8) 西堀 研人, 高橋 友和, 出口 大輔, 井手 一郎, 村瀬 洋: 低解像度人体画像の学習型超解像に関する検討, 信学技法 PRMU, pp. 213–218, 2013.
- 9) 西堀 研人, 高橋 友和, 出口 大輔, 井手 一郎, 村瀬 洋: 監視カメラシステムのための低解像度人体画像の超解像に関する検討, 計測自動制御学会 システム・情報部門 学術講演会 (SSI 2013), 2013.
- 10) Zhou Wang, Bovik, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P. Simoncelli: Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Processing, vol. 13, no. 4, pp. 600–612, 2004.

西堀研人: 2009 年名大・大学院博士後期課程修了, 2013 年より電子学会会員. 画像認識の基礎研究ならびにその応用に興味を持つ.

高橋友和: 平 9 茨城大・工・情報卒. 平 12 同大大学院理工学研究科博士前期課程了. 平 15 同研究科博士後期課程了. 博士 (工学). 同年より 2

年間名古屋大学大学院情報科学研究科 COE 研究員. 平 17 より 3 年間日本学術振興会特別研究員. 平 20 より岐阜聖徳学園大学経済情報学部准教授, 現在に至る. 画像認識の基礎研究ならびにその応用に興味を持つ. 電子学会会員.

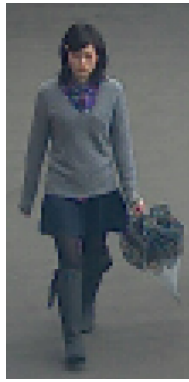
出口大輔: 2001 年, 名大・工・情報卒. 2006 年, 同大学院博士後期課程修了. 2004 年~2006 年まで日本学術振興会特別研究員. 2008 年より名古屋大学大学院情報科学研究科助教. 博士 (情報科学). 画像処理・パターン認識技術の開発とその ITS および医用応用に関する研究に従事.

井手一郎: 1994 年, 東大・工・電子卒. 2000 年, 同研究科電気工学専攻博士課程了. 博士 (工学). 同年国立情報学研究所助手. 2004 年, 名古屋大学大学院情報科学研究科助教, 2007 年より准教授. パターン認識技術の実応用や映像メディア処理全般に興味を持っている.

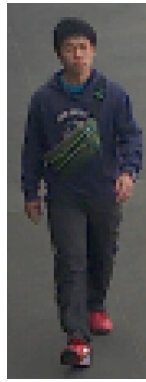
村瀬洋: 1978 年名大・工・電気卒. 1980 年同大大学院修士課程了. 同年, 日本電信電話公社 (現 NTT) 入社. 2003 年名古屋大学大学院情報科学研究科教授. 工学博士. 画像処理, パターン認識の研究に従事.



(i) 人物 1
47 × 114 画素



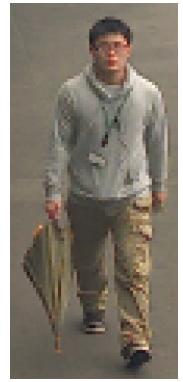
(ii) 人物 2
55 × 111 画素



(iii) 人物 3
44 × 116 画素

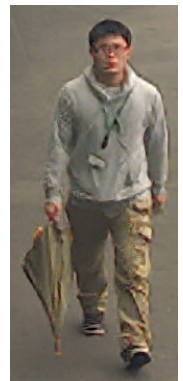
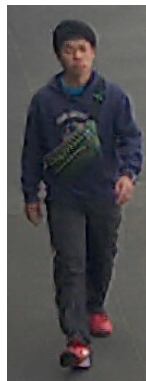
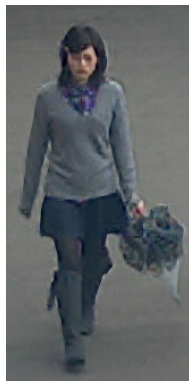


(iv) 人物 4
42 × 107 画素

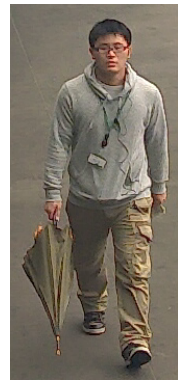
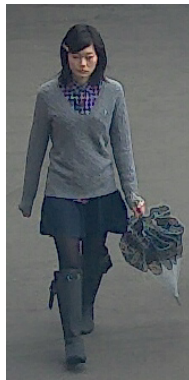


(v) 人物 5
54 × 121 画素

(a) 低解像度入力画像



(b) 超解像度画像



(c) 元の高解像度画像



(d) SSIM マップ

図 4: 入力画像に監視カメラ映像を低解像度化して用いた場合の超解像画像の画質評価



図 5: 入力画像に監視カメラ映像を用いた場合の超解像画像の画質評価