

反復式生成型学習法を用いた歩行者検出

○ 吉田英史[†], 出口大輔[‡], 井手一郎[†], 村瀬洋[†], 町田貴史^{††}, 小島祥子^{††}

○ Hidefumi YOSHIDA[†], Daisuke DEGUCHI[‡], Ichiro IDE[†], Hiroshi MURASE[†]
Takashi MACHIDA^{††} and Yoshiko KOJIMA^{††}

[†]: 名古屋大学 大学院情報科学研究科, {yoshidah,ide,murase}@murase.m.is.nagoya-u.ac.jp

[‡]: 名古屋大学 情報連携統括本部, ddeguchi@nagoya-u.jp

^{††}: 株式会社豊田中央研究所

近年, 車載カメラで撮影された画像から歩行者を検出する技術に関する研究が注目されている. 画像中の歩行者の見えを特徴として用いる従来の検出手法では, 多様な環境下で撮影された大量の学習用画像を事前に人手で収集し, それらを用いて検出器を構築する必要があった. この問題を解決するために, 本発表では検出に有効な歩行者画像を反復的に生成する“反復式生成型学習法”を提案する. 本手法では, まず少数の歩行者画像を用いて形状やテクスチャ, 背景の生成モデルを構築する. 次に, 実際の歩行者画像から誤検出しやすい歩行者の生成パラメータを求め, その近傍で多数の学習用歩行者画像を生成する. 実験の結果, 従来手法と比べて検出精度は大きく向上し, 提案手法の有効性を確認した.
<キーワード> 歩行者検出, 生成型学習法, HOG, SVM

1. はじめに

自動車運転支援技術の1つとして, 車載カメラを用いて歩行者を検出する技術に期待が寄せられている. 画像から歩行者を検出する手法の多くは, 歩行者とその周辺のエッジや濃淡などの見えの情報を利用したものである [1]. このような手法の1つとして, HOG 特徴と SVM を用いた歩行者検出手法が広く用いられている [2]. この手法では, 多様な見えの歩行者を精度よく検出するために, 様々な環境下で撮影された歩行者画像を数多く用意し, それらを歩行者検出器の学習に用いる必要がある. しかし, 図 1 に示すように, 歩行者の見えは多様であり, これらを人手で収集する労力は大きい.

この問題を解決するために, 人工的に多様な見えをもつ歩行者画像を生成し, 学習に用いる枠組みが提案されている. この枠組は“生成型学習法”と呼ばれ, 対象の実環境における見えの変化をモデル化することで, 少数の原画像から変化に富んだ多数の学習用画像を生成することができる [3]. このとき構築されるモデルを生成モデルと呼び, 対象に合わせて生成モデルを適切に選択することで, 歩行者検出をはじめとする様々な研究に応用



図 1 多様な見えをもつ歩行者画像例

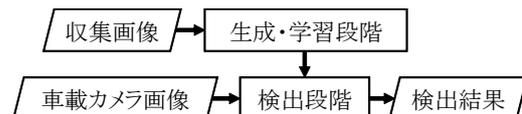


図 2 提案手法の処理の流れ

されている [3, 4, 5, 6, 7]. しかし, 見えの変化を制御するパラメータの適切な設定が難しく, 多くの場合, 変化のパラメータは経験的に決定される. 特に歩行者検出への応用を考えると, 不適切な学習用画像が生成されることで, 検出器の性能が低下することすら考えられる.

そこで本研究では, 実際の歩行者画像から学習が不十分な画像特徴をもつものを見つけ, それに似た特徴をもつ学習用画像を重点的に生成する. このようにすることで, 学習に有効なサンプルを生成するための生成パラメータを, 経験的ではな

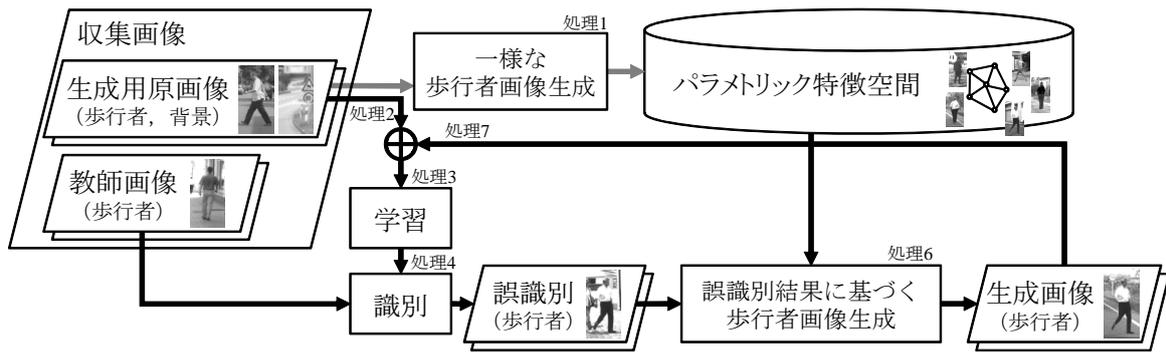


図 3 生成・学習段階の処理の流れ

く自動的に求めることができる。

本研究では、統計的形状モデルを用いた歩行者画像生成手法 [4] を生成モデルとして歩行者画像を生成する。その際、モデル化された歩行者の形状やテクスチャ、背景から検出に有効な生成パラメータを推定する。このようにして生成した歩行者画像を検出器の学習に用いることで、歩行者の検出精度向上を図る。

以降、2 節で提案手法による歩行者検出器の構築方法について説明する。3 節では、実際に車載カメラによって撮影された画像を用いた提案手法の評価実験について述べ、考察を加える。最後に、4 節でまとめる。

2. 提案手法

提案手法の処理の流れを図 2 に示す。提案手法は生成・学習、検出の 2 段階からなる。まず生成・学習段階では、収集した少数の歩行者画像および背景画像を用いて、検出に有効な特徴をもつ歩行者画像を反復的に生成し、歩行者識別器を繰り返し構築する。次に検出段階では、構築した歩行者識別器を用いて車載カメラ画像を走査し、車載カメラ画像から歩行者を検出する。以降の節でこれらを順に説明する。

2.1 生成・学習段階

生成・学習段階の処理の流れを図 3 に示す。処理を始める前に、入力である収集画像を“生成用原画像”と“教師画像”の 2 群に分ける。ここで、“生成用原画像”は歩行者画像生成に用い、“教師画像”は生成パラメータ推定に用いる。生成・学習段階は、以下に示す順で歩行者画像の生成および識別器の構築を反復的に行う。

- 処理 1. 生成モデルのパラメータを一様に変化させながら歩行者画像を生成し、“パラメトリック特徴空間”を構築する (2.1.1 節)。
- 処理 2. 生成用原画像を学習用画像とする。
- 処理 3. 歩行者識別器を構築する (2.1.2 節)。
- 処理 4. 処理 3 で構築した識別器を用いて教師画像を識別する。
- 処理 5. 誤識別があれば処理 6 に進み、なければ処理 8 に進む。
- 処理 6. 誤識別結果に基づき、“パラメトリック特徴空間”を用いて歩行者画像を生成する (2.1.3 節)。
- 処理 7. 生成画像を学習用画像に追加し、処理 3 に戻る。
- 処理 8. 教師画像を学習用画像に追加し、最終的な歩行者識別器を構築する (2.1.2 節)。

これら一連の生成・学習処理を反復式生成型学習法と呼ぶ。生成・学習段階の各処理について以降で詳しく説明する。

2.1.1 “パラメトリック特徴空間”の構築

本節では、収集した少数の歩行者画像および背景画像から生成モデルを構築し、生成パラメータを一様に変化させることで多数の歩行者画像を生成する。そして、歩行者画像を生成するパラメータと画像特徴の対応を与える“パラメトリック特徴空間”を構築する。生成・学習段階において、本処理は最初に一度だけ行われる。歩行者画像生成には、生成型学習法を用いた歩行者検出 [4] で提案されている手法を用いる。この手法を用いた歩行者画像生成の詳細な手順については、付録 A に掲載する。

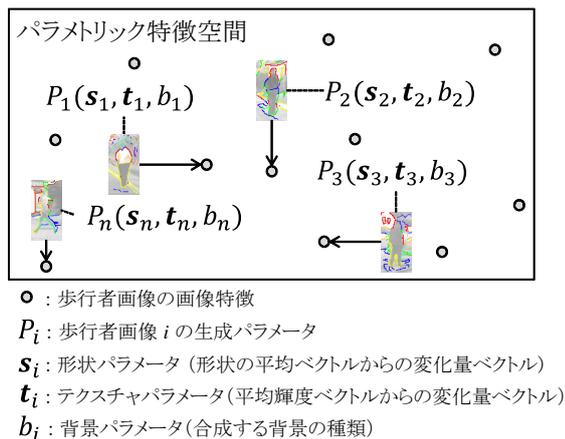


図 4 パラメトリック特徴空間の概念図

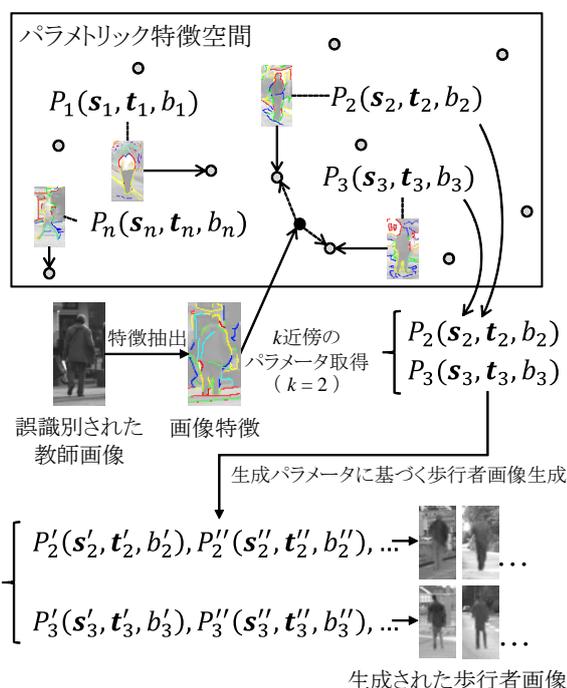


図 5 識別結果を反映した歩行者画像生成の処理の流れ

提案手法では、まず図 3 に示すように、収集した歩行者画像を 2 群に分割する。一方の歩行者画像及び背景画像を生成モデル構築のための生成用原画像とし、統計的形状モデル (Statistical Shape Model: SSM) [8] を用いて多様な形状、テクスチャ、背景をもつ多数の歩行者画像を生成する。ここで、生成に用いた形状の変化量やテクスチャの輝度値の変化量、背景の種類が生成パラメータとなる。

次に、生成された歩行者画像から画像特徴を抽

出し、画像特徴と生成パラメータが 1 対 1 に対応付けられた特徴空間を作る。本発表では、この特徴空間を“パラメトリック特徴空間”と呼ぶ。パラメトリック特徴空間の概念図を図 4 に示す。

2.1.2 歩行者識別器の構築

まず、学習用画像の歩行者画像および背景画像から画像特徴をそれぞれ抽出する。次にそれらを学習して、入力画像が歩行者であるか否かを判別する 2 クラス歩行者識別器を構築する。

2.1.3 誤識別結果に基づく歩行者画像の生成

2.1.2 節で構築した歩行者識別器を用いて教師画像を識別する。ここで、誤識別された歩行者画像はその識別器では識別困難な画像特徴を持っていると考えられる。よって、誤識別された歩行者画像を生成するパラメータを、2.1.1 節で構築したパラメトリック特徴空間を用いて推定し、得られたパラメータの付近で新たに歩行者画像を生成し、学習用画像に加える。

識別結果を反映した歩行者画像生成の処理の流れを図 5 に示す。ここではまず、2.1.2 節で構築した歩行者識別器によって誤識別された歩行者画像から画像特徴を抽出する。得られた特徴を 2.1.1 節で構築したパラメトリック特徴空間にプロットし、 k 近傍の特徴を求め、それらの特徴に対応する生成パラメータを得る。次に、これらの生成パラメータの周辺で再度歩行者画像を複数生成する。

生成の手順は 2.1.1 節と同様、文献 [4] で提案されている手法を用いた。ただし、形状生成およびテクスチャ生成では、 k 近傍で得られた生成パラメータとその近傍のパラメータ間を補間するようにして生成する。これは、背景の変化に比べて、形状やテクスチャの変化が歩行者の特徴をよく表すと考えられるからである。

2.2 検出段階

検出段階では、車載カメラ画像上で矩形領域を走査し、構築した歩行者識別器を用いてその領域に歩行者が含まれているか否かを識別する。この操作を矩形領域のサイズを変えながら車載カメラ画像全体に対して行い、様々な大きさの歩行者を検出する。

表 1 各手法の概要

手法	画像生成	学習用画像枚数		
		歩行者画像		背景画像
		収集	生成	
提案	反復生成	200	5,634	5,000
比較 1	単純生成	200	5,634	5,000
比較 2	なし	200	-	5,000

3. 評価実験

本節では提案手法の評価実験について述べる。

3.1 データセット

収集画像および評価用画像系列として、Daimler社より提供されているグレースケール画像のベンチマークデータセット¹を利用した。歩行者画像は200枚をランダムに収集し、生成用原画像や教師画像として使用した。また背景画像として、データセット内で歩行者の存在しない車載カメラ画像200枚を選択し、歩行者画像と同じ大きさにランダムに5,000枚を切り出して使用した。

精度評価には、述べ1,110人の歩行者を含む1,016枚の車載カメラ画像系列を用いた。これらは、学習に用いた画像系列とは異なるものである。

3.2 識別器の仕様

学習およびパラメトリック特徴空間の構築に使用する画像特徴には、歩行者検出に広く用いられているHOG特徴を使用し、学習にはSVMを使用した。SVMにはcuSVM[9]を利用し、そのカーネルにはRBFカーネルを使用した。

3.3 実験条件

各手法の実験条件を表1に示す。提案手法は、収集した200枚の歩行者画像の内、90枚を生成用原画像として使用し、110枚を教師画像として使用した。そして、これらの画像から5,634枚の歩行者画像を生成した。つまり、計5,834枚の歩行者画像を最終的な検出器の学習に用いた。パラメトリック特徴空間から生成パラメータを求める際

¹http://www.gavrila.net/Research/Pedestrian_Detection/Daimler_Pedestrian_Benchmark_D/Daimler_Mono_Ped_Detection_Be/daimler_mono_ped_detection_be.html

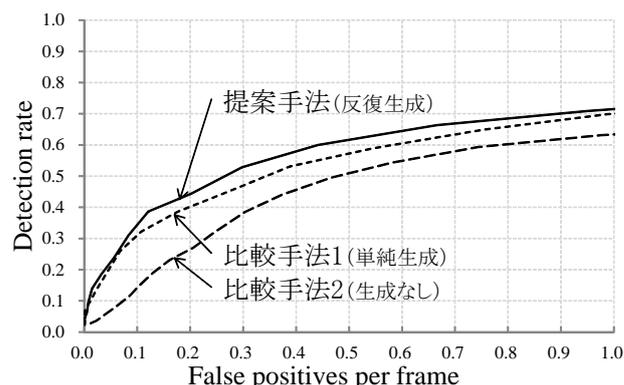


図 6 各手法における歩行者検出精度

の近傍数は $k = 25$ とし、1近傍当たりの周辺画像の生成枚数は9枚とした。

比較手法1は、パラメトリック特徴空間からランダムにサンプリングした生成パラメータを用いて、歩行者画像5,634枚を生成した（反復処理は行わない）。ここに、提案手法と同様に、収集した200枚を加えた計5,834枚の歩行者画像を最終的な検出器の学習に用いた。

比較手法2は2.1節の生成・学習段階を経ずに、収集した歩行者画像200枚のみを学習に用いた。

最終的な識別器構築時の背景画像には、すべての手法で同じ5,000枚を用いた。パラメトリック特徴空間構築時の背景画像には、この内600枚を使用した。これらの手法を用いて、評価画像系列に対し歩行者検出を行い、データセットに含まれている正解データを用いて検出精度を算出した。

3.4 結果および考察

提案手法と従来手法による歩行者検出精度を図6に示す。図6は縦軸に歩行者の検出率、横軸にFalse Positives Per Frame (FPPF: フレームあたりの誤検出数)をとったグラフで、識別器の出力するスコア（確信度）に対する検出しきい値を変化させることで描画した。このグラフは曲線が左上に存在するほど精度が良いことを表す。グラフ上のFPPF=0.2の点における各手法の検出結果画像例を図7に示す。また、提案手法の反復処理による効果を図8に示す。図8は、反復回数に対する検出精度（最大F値）である。なお本実験での反復回数は3回であり、3回目で教師画像の誤識別がなくなり、生成・学習段階を終えた。

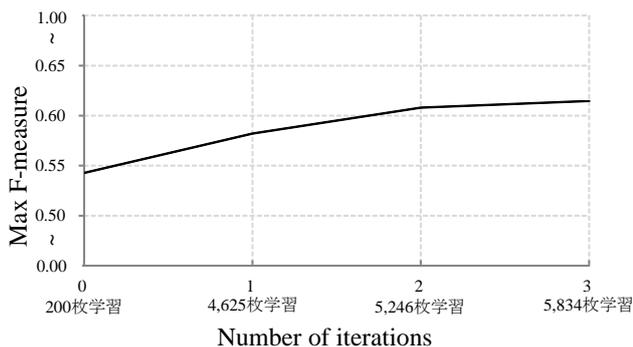


図 8 提案手法における反復回数と検出精度

図 6 から提案手法は比較手法に比べて精度よく歩行者を検出できていることが分かる。特に、FPPF が小さいとき、つまりしきい値を厳しく設定したときの検出率が最大で 15% 以上向上している。この結果から、厳しいしきい値を設定した場合に比較手法では検出できなかった歩行者であっても、提案手法では高いスコアで検出できたことを表している。これは生成・学習段階で、誤識別された教師画像に近い特徴を持つ歩行者画像を生成し、学習した効果であると考えられる。

また、図 7 の左から 1, 2 列目の検出結果画像を見ると、提案手法によって、比較手法では検出できなかった歩行者を検出できたことがわかる。一方、比較手法 1 では、左から 3 列目の画像のように未検出を減らすことはできたが、同時に誤検出が増えてしまった。この原因としては、パラメータ推定をせずに歩行者画像を生成したことが考えられる。

最後に図 8 から、提案手法は反復を繰り返すことで検出精度が向上することが示された。

4. むすび

本研究では、反復的に検出に有効な歩行者画像を生成・学習して歩行者検出器を構築する反復式生成型学習法を提案した。具体的には、少数の歩行者画像から一様に生成した学習用画像を用いてパラメトリック特徴空間を構築し、実際の歩行者画像を用いて誤検出しやすい歩行者画像の生成パラメータを求めた。そして、そのパラメータの近傍で学習用画像を生成し、学習に用いることで検出精度が向上することを確認した。

今後の課題として、より大規模な評価用画像系列を用いた提案手法の精度評価が考えられる。

謝辞

本研究の一部は、JST 戦略的創造研究推進事業 CREST および科学研究費補助金による。また、本研究では画像処理に MIST ライブラリ (<http://mist.murase.m.is.nagoya-u.ac.jp/>) を使用した。

付録 A – 歩行者画像生成

歩行者画像の生成の流れを図 9 に示す。歩行者画像の生成は、形状生成、テクスチャ生成、背景合成の順に行う。まず、少数の歩行者画像に対して、歩行者の輪郭上に一定数の制御点をプロットする。次に、すべての歩行者画像上の対応する制御点座標の平均と分散を求め、統計的形状モデル (Statistical Shape Model : SSM) [8] を構築する。そして、構築された SSM を用いて新たな歩行者の形状を一様に生成する。生成された歩行者形状の例を図 9 の右上に示す。

上記と同様の手順により、歩行者のテクスチャの画素値に対して SSM を構築することでテクスチャ生成を行い、生成した形状に歩行者のテクスチャを与える。ここで、生成した形状および生成用原画像の歩行者の輪郭には 1 対 1 に対応する制御点がプロットされている。そのため、原画像内の制御点 3 点で構成される三角形領域を、SSM により生成した形状の対応する三角形領域にアフィン変換を用いて貼り付ける。これにより、新たなテクスチャをもつ歩行者画像を得る。この時、三角形を構成する 3 点は、原画像にドロネー三角形分割を適用して得られる三角形の頂点を用いる。

最後に、得られた歩行者画像と背景を合成する。

参考文献

- [1] P. Dollár, C. Wojek, B. Schiele and P. Perona, “Pedestrian detection: An evaluation of the state of the art,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 34, no. 4, pp.743–761, Apr. 2012.
- [2] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp. 886–893, June 2005.



図 7 各手法の検出結果

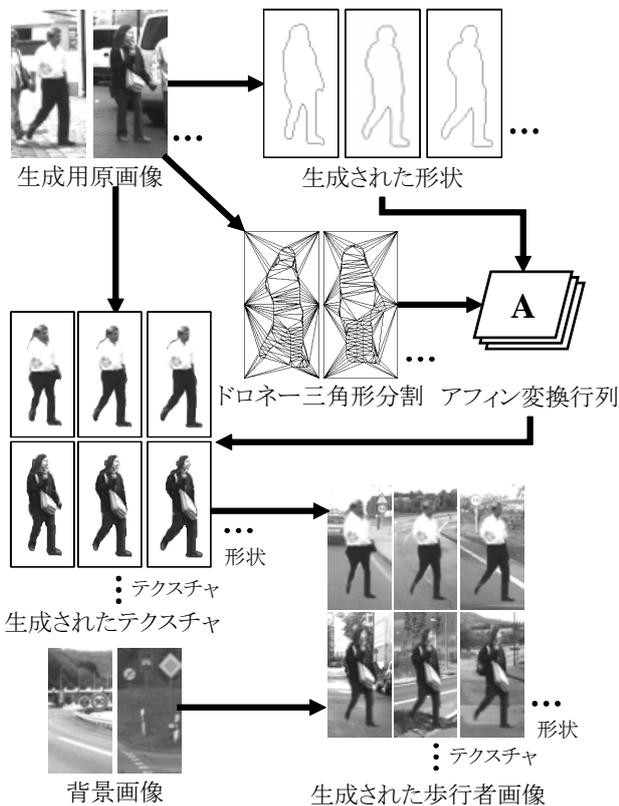


図 9 歩行者画像生成の流れ

[3] 村瀬洋, “画像認識のための生成型学習法,” 情報学論, vol.46, no.SIG15 (CVIM 12), pp.35–42, Oct. 2005.

[4] H. Yoshida, D. Deguchi, I. Ide, H. Murase,

K. Goto, Y. Kimura and T. Naito, “Integration of generative learning and multiple pose classifiers for pedestrian detection,” Proc. 7th Int. Conf. on Computer Vision Theory and Applications, vol.1, pp.567–572, Feb. 2012.

[5] J. Marín, D. Vázquez, D. Gerónimo and A. M. López, “Learning appearance in virtual scenarios for pedestrian detection,” Proc. 2010 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.137–144, June 2010.

[6] 道満恵介, 出口大輔, 高橋友和, 目加田慶人, 井手一郎, 村瀬洋, “色変動を考慮した生成型学習法による道路標識検出器の構築,” 信学論(D), vol.J93-D, no.8, pp.1375–1385, Aug. 2010.

[7] M. Noda, T. Takahashi, D. Deguchi, I. Ide, H. Murase, Y. Kojima and T. Naito, “Recognition of road markings from in-vehicle camera images by a generative learning method,” Proc. 11th IAPR Conf. on Machine Vision Applications, pp.514–517, May 2009.

[8] T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham, “Active shape models. Their training and application,” Computer Vision and Image Understanding, vol.61, pp.38–59, Jan. 1995.

[9] A. Carpenter, “cuSVM: A CUDA implementation of support vector classification and regression,” <http://patternsonscreen.net/cuSVM.html>