

生成型学習法を用いた傘をさした歩行者の高精度な検出に関する検討

吉田 英史[†] 出口 大輔[†] 井手 一郎[†] 村瀬 洋[†]

町田 貴史^{††} 小島 祥子^{††}

[†] 名古屋大学 〒464-8601 愛知県名古屋市千種区不老町

^{††} (株)豊田中央研究所 〒480-1192 愛知県愛知郡長久手町大字長瀬字横道 41-1

E-mail: †yoshidah@murase.m.is.nagoya-u.ac.jp, †ddeguchi@nagoya-u.jp, †{ide,murase}@is.nagoya-u.ac.jp

あらまし 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研究が注目されている。特に、傘をさした歩行者は自動車の接近に気付きにくいいため正しく検出してドライバに警告する必要がある。しかし、傘をさした歩行者は傘によって体の一部が隠れてしまうため、従来の技術では検出が難しい。この問題を解決するため、本研究では車載カメラ画像からの歩行者検出において、特に傘をさした歩行者を検出する手法を提案する。提案手法では、これまで提案してきた生成型学習法を用いた歩行者検出手法を発展させ、傘を合成した歩行者画像を多数生成し、それらを検出器の学習に利用する。これにより、傘をさした歩行者の高精度な検出の実現をねらう。車載カメラ画像を用いた実験の結果、従来手法に比べて提案手法の検出精度は14%高く、そ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, 雨天, 所持品, 傘, HOG, SVM,

A study on a method for high-accuracy detection of a pedestrian holding an umbrella with generative learning

Hidefumi YOSHIDA[†], Daisuke DEGUCHI[†], Ichiro IDE[†], Hiroshi MURASE[†],

Takashi MACHIDA^{††}, and Yoshiko KOJIMA^{††}

[†] Nagoya University Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601 Japan

^{††} Toyota Central Research & Development Laboratories, Inc. Nagakute, Aichi, 480-1192, Japan

E-mail: †yoshidah@murase.m.is.nagoya-u.ac.jp, †ddeguchi@nagoya-u.jp, †{ide,murase}@is.nagoya-u.ac.jp

Abstract Recently, pedestrian detection from in-vehicle camera images has become an interesting topic for researchers. Especially, since a pedestrian holding an umbrella may not recognize an approaching vehicle, the system is expected to detect them and alert the driver. However, it is difficult to detect a pedestrian holding an umbrella by existing detection methods since an umbrella usually occludes some parts of the body. To tackle this problem, we propose a method to detect a pedestrian holding an umbrella from an in-vehicle camera image. In this paper, we propose a method that synthesizes umbrellas onto generated pedestrian images, and then learns them for improving the accuracy of the detection of pedestrians holding umbrellas. Experimental results showed that the detection accuracy of the proposed method outperformed existing methods by 14%.

[Note] This document is an informal handout distributed at an IEICE TC-PRMU workshop.

Key words Pedestrian detection, generative learning, rainy day, belongings, umbrella, HOG, SVM

1. ま え が き

近年、自動車運転支援技術の1つとして、車載カメラによって撮影された画像から歩行者を検出する技術に期待が寄せられている。自動車の走行環境において、雨天時に傘をさした歩行者や快晴時に日傘をさした歩行者は自動車の接近に気付きにく

いという点で、非常に危険である。そのため、このような歩行者を検出し、事故を未然に防ぐ安全運転支援システムが強く望まれている。

車載カメラを用いて撮影された画像から歩行者を検出する研究は盛んに行われており、それらの多くが歩行者とその周囲のエッジや濃淡といった見えの情報を利用している。近年では、

HOG 特徴と線形 SVM を用いた手法が広く歩行者検出に用いられている [1]. HOG 特徴とは、画像の局所領域の輝度勾配方向ヒストグラムであり、照明変化や影の影響を受けにくく、かつ局所的な幾何学変化に頑健である。また、SVM は高い汎化性能を持つ識別器である。この手法では、歩行者の多様な見えの変化に対応するために様々な環境下で歩行者画像を撮影し、それらを数多く収集して学習する必要がある。しかし、歩行者の見えは歩行動作や所持品、更には歩行者の周辺環境によって大きく変化する。これらの見えの変化の組み合わせは膨大なため、人手でこれらの歩行者画像を収集することは非常にコストが高い。

この問題に対処するために、人工的に多様な見えの歩行者画像を生成する枠組みが提案されている。この枠組みは“生成型学習法”と呼ばれており、様々な対象に応用されている [2]~[8]. 生成型学習法を用いることで、実際の環境で起こりうる様々な見えの変化をモデル化し、このモデルを用いて少数の原画像から多様な変化を含む多数の生成画像を体系的に生成することができる。これにより、低コストに多様な変化を含む画像を学習に使用できるため、低コストで精度の良い検出器の構築が可能となる。

生成型学習法を用いた歩行者検出は、生成モデルの良し悪しが検出性能にも大きく寄与する。歩行者画像を生成する手法には、3次元 CG を用いて作成した仮想環境を生成モデルとして利用する手法 [7], [8] と、実際に撮影された画像から統計的形狀モデル [9] を作成し、歩行者画像を多数生成する手法 [5], [6] が提案されている。3次元 CG を用いてモデルを作成する手法では、多様な見えの画像を生成するにあたり、モデルに対して変化を加えることが容易である。例えば、作成する画像の見えを変化させたい場合、モデルのマテリアルやテクスチャを変えることによって光の反射の仕方や質感を変化させ、見えの異なる画像を生成することができる。しかしながら、生成した画像が本当に現実の見えに即しているかどうかを評価することは難しい。一方で、収集した画像から得られる統計情報に基づいて統計的形狀モデルを作成する手法では、生成画像が実際の見えに即しているという特徴がある。ところが、生成画像に与えることのできる見えの変化は実際に収集された画像に依存するため、生成の基となる原画像の収集には注意が必要である。本研究では、より実際の歩行者の見えに即している統計的形狀モデルを用いた Enzweiler らの手法 [5] をベースとして用いる。この手法では歩行者の形状、テクスチャ、背景をモデル化することで、少ない歩行者原画像から多様な見えを持つ歩行者画像を生成し、学習に用いる。

しかしながら、これまでに述べたいずれの手法も歩行者の持ち物を考慮しておらず、持ち物によって通常とは見えの異なる歩行者を精度よく検出することが難しい。歩行者の持ち物の中でも、傘はさし方によって歩行者の見えを大きく変化させる。例えば、図 1 右端の画像のように傘を背面に傾けてさした歩行者を後ろから見た場合、その歩行者の頭部は完全に傘に隠れてしまい、傘をさしていない歩行者と見えが大きく異なる。そのため、これらの従来手法を用いて晴天時に日傘をさしている



図 1 傘によって見えが変化した歩行者の例

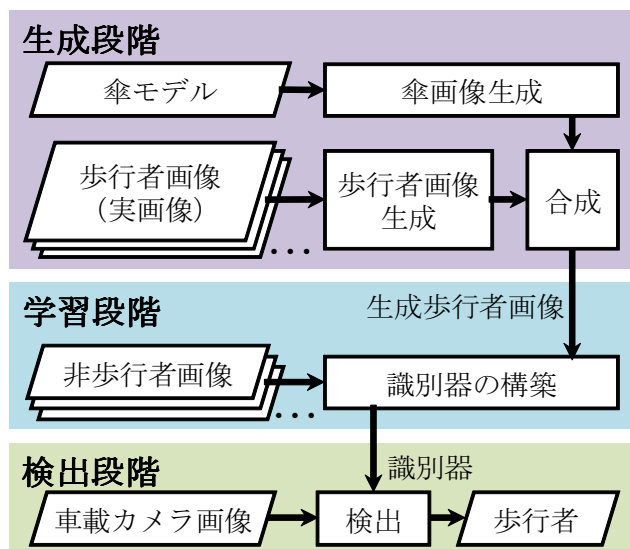


図 2 提案手法の処理の流れ

歩行者や、雨天時に傘をさしている歩行者を検出することは困難である。このような部分的に隠れた歩行者を検出する方法として、Enzweiler らは隠れていない部分の識別結果を統合するパーツベースの手法 [10] を提案している。この手法では、画像の見えの情報に加えて深度情報や動き情報を利用することで、部分的に隠れた歩行者の検出を実現している。これに対して本研究では、傘を含んだ歩行者全体の画像から、見えの情報のみを学習に使用して検出を行う。

本研究では、まず傘のモデルを作成して多様な見えの傘画像を生成する。このとき、傘は剛体であることから 3次元 CG で比較的簡単にモデリングできるため、傘のモデリングに 3次元 CG を採用する。そして、生成した傘画像と従来の生成型学習法 [5] を組み合わせ、傘をさした歩行者画像を生成する。こうすることで、従来の多様な形状、テクスチャ、背景に加え、傘による見えの変化を持った歩行者画像を生成することができる。これらの画像を用いて歩行者検出器を構築することにより、傘をさした歩行者の検出精度向上を図る。

以降、2 節で傘画像の生成手順、傘をさした歩行者画像の生成方法について説明する。3 節では実際に車載カメラにより撮影された画像を用いた比較実験について述べ、考察を加える。最後に 4 節で本報告をまとめる。

2. 提案手法

図 2 に提案手法の流れを示す。提案手法は生成段階、学習段階、検出段階の 3 段階から構成される。生成段階では、まず傘の 3D モデルを作成し、多様な見えを持つ傘画像を生成する。

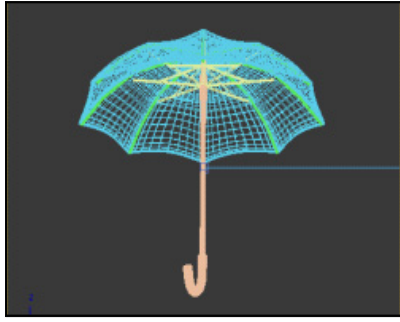


図 3 傘モデルのワイヤフレーム表示

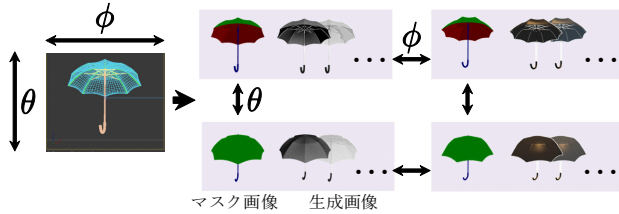


図 4 傘画像生成の流れ

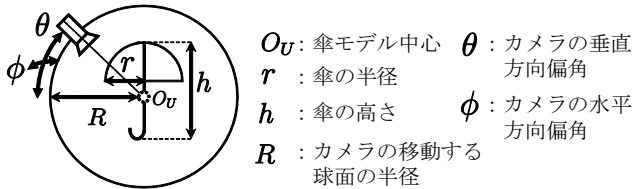


図 5 傘モデルパラメータ及びカメラパラメータ

そして、従来の生成型学習法によって生成された多様な姿勢変化を含む歩行者画像と傘画像を合成して、多様な見えをもつ傘をさした歩行者画像を生成する。学習段階では、生成した歩行者画像に非歩行者画像を加え、2クラス識別器を構築する。検出段階では、学習段階で構築した識別器を用いて車載カメラ画像中から歩行者を検出する。

2.1 傘画像生成

傘をさした歩行者画像の生成に際し、歩行者の前景画像、背景画像、傘画像を用意する。これらの画像を合成し、傘をさした歩行者画像を生成する。まず最初に、本節では傘モデルを作成し、そのモデルから多様な見えをもつ傘画像を生成する傘画像生成処理について述べる。

ここでは、3次元CGを用いて傘モデルを作成する。作成したモデルを図3に示す。また、図4に傘画像生成処理の流れを示す。傘モデル作成時の各種パラメータを図5に示す。これらの座標系は傘モデルの中心 O_U を原点とする。

- 傘モデルパラメータ
高さ h 、半径 r 、マテリアルの種類 M
- カメラパラメータ
傘モデルの中心からの距離 R
垂直方向偏角 θ 、水平方向偏角 ϕ

ここで、カメラは図5に示すように O_U を原点とした半径 R の球面上を移動し、カメラは常に O_U を画面中心に捉えるように設定する。



図 6 生成した歩行者の前景画像の例

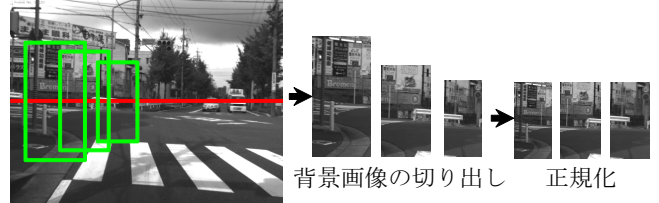


図 7 背景画像収集の流れ

モデルから傘画像を生成する際に、傘画像に加えて図4に示したマスク画像を生成する。マスク画像を用いることで、傘をさした歩行者画像の生成時に歩行者と傘の前後関係を考慮した生成を実現する。マスク画像の傘の内側は赤色に塗られており、外側は緑色に塗られている。このとき、マスク画像の赤色に塗られた領域が歩行者の背面に合成される部分を表し、緑色に塗られた領域が歩行者の前面に合成される部分を表す。

2.2 歩行者画像生成

歩行者の前景画像はEnzweilerらの手法[5]により生成する。まず少数の歩行者画像から歩行者の輪郭形状を抽出し、統計的形状モデルを用いて多様な形状を生成する。次に、Delaunay三角形分割とアフィン変換を用いて生成した形状にテクスチャを与えることで、歩行者の前景画像を生成する。これにより、多様な体型、姿勢及び服装の歩行者の前景画像を生成することができる。生成した歩行者の前景画像の例を図6に示す。

2.3 背景画像選択

次に、歩行者の写っていない車載カメラ画像から、合成に使用する背景画像を図7のように収集する。切り出す位置に制約を与えることで空や地面のみの写った画像が収集されることを抑制し、不自然な歩行者画像が生成されることを避ける。ここでは、切り出し枠の位置に関する制約として、垂直方向は枠の中心と画像の中心が一致するように設定する。水平方向には制約を与えない。切りだされたすべての背景画像は歩行者の前景画像と同じ大きさに正規化する。

2.4 画像合成

最後に、切り出した背景画像に生成した多様な形状、テクスチャを持つ歩行者の前景画像、2.1節で生成した傘画像を合成する。合成時に正規乱数を用いて、画像に微小な位置ずれや回転、拡大縮小を適用することで、多様な見えを持つ歩行者画像を生成する。傘画像及び歩行者画像の合成パラメータを以下に示す。

- 傘画像合成パラメータ
拡大縮小 $N(\mu_{us}, \sigma_{us})$ 、位置ずれ $N(\mu_{up}, \sigma_{up})$ 、
回転 $N(\mu_{ur}, \sigma_{ur})$
- 歩行者画像合成パラメータ
拡大縮小 $N(\mu_{ps}, \sigma_{ps})$ 、位置ずれ $N(\mu_{pp}, \sigma_{pp})$

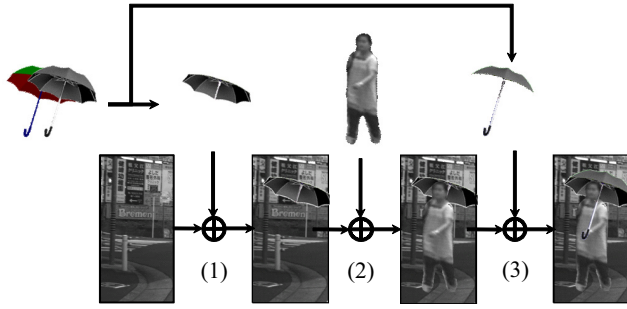


図 8 画像合成の流れ



図 9 歩行者画像の例



図 10 非歩行者画像の例

また、傘画像の合成を行うときに、同時に生成したマスク画像を用いて予め傘画像を歩行者の前面になる部分と背面になる部分に分割しておく。マスク画像の緑色に塗られた部分を前面、赤色に塗られた部分を背面として分離する。合成は以下の順に行なう。合成の流れを図 8 に示す

- (1) 背景画像に傘の背面部分を合成
- (2) (1) に歩行者の前景画像を合成
- (3) (2) に傘の前面部分を合成

3. 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

識別器の構築に必要な学習用画像及び評価用画像系列は、車両に車載カメラを搭載し、市街地を走行して収集した。歩行者画像生成の種となる原画像には、晴れの日収集した車載カメラ画像から手で歩行者画像を 60 枚切り出して使用した。更に、歩行者の写っていない車載カメラ画像にスケーリングを加え、歩行者画像と同じ大きさとランダムに切りだした 10,000 枚の画像を学習用非歩行者画像として使用した。用意した歩行者画像の例を図 9、非歩行者画像の例を図 10 にそれぞれ示す。

また、雨の日収集した車載カメラ画像を評価用画像とした。このとき、50 ~ 480 画素の高さの歩行者を 1 人以上含む区間のみを評価の対象とし、延べ 1,813 人の歩行者を含む計 1,000 枚の画像を使用した。評価画像の例を図 11 に示す。

3.2 識別器構築

学習に用いる画像特徴には、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 は、cuSVM[11] を利用した。なお、SVM のカーネルには RBF



図 11 評価画像の例

表 1 手法概要

手法	傘画像合成	学習/生成型学習に用いた入力歩行者画像枚数	識別器の構築に用いた画像枚数	
			歩行者	非歩行者
提案	✓	60	5,000	10,000
従来		60	5,000	10,000



図 12 生成した傘画像の例



図 13 生成した傘をさした歩行者画像の例

カーネルを使用した。

3.3 実験条件

表 1 に示した実験条件で従来手法と比較し評価した。

提案手法では、作成した傘モデルから 220 枚の傘画像を生成した。生成した傘画像の例を図 12 に示す。次に、収集した 60 枚の歩行者画像から、従来手法 [5] によって多様な見えの変化を含む 5,000 枚の歩行者画像を生成した。これらの歩行者画像の内 3,500 枚に生成した傘画像を合成し、傘画像を合成しない 1,500 枚とあわせて 5,000 枚の歩行者画像を学習に使用した。生成した傘をさした歩行者画像の例を図 13 に示す。また、傘をさした歩行者画像の生成に使用した各種生成パラメータを表 2 に示す。

従来手法は傘画像の合成を行わず、収集した 60 枚の歩行者画像から従来手法 [5] により生成した 5,000 枚の歩行者画像をそのまま学習に用いた。学習では、すべての手法で同一の非歩行者画像 10,000 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ない、人手で与えた正解データ

表 2 生成パラメータ

処理	パラメータ	値
傘画像生成	O_U	(0, 0)
	h	2.67
	r	1.45
	M	10
	R	6.00
	θ	$[-21.8^\circ, 21.8^\circ]$ を 4.36° ずつ等分割した 11 パターン
	ϕ	$0^\circ, 22.5^\circ$
傘画像合成	(μ_{us}, σ_{us})	(1.6, 0.03)
	(μ_{up}, σ_{up})	x 軸方向 (0.0, 10.0) y 軸方向 (0.0, 1.0)
	(μ_{ur}, σ_{ur})	(0.0, 10.0)
歩行者画像合成	(μ_{ps}, σ_{ps})	(0.9, 0.00005)
	(μ_{pp}, σ_{pp})	x 軸方向 (0.0, 1.0) y 軸方向 (6.0, 1.0)

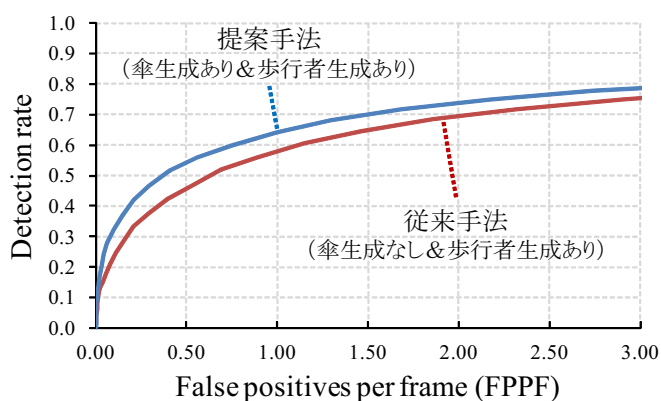


図 14 検出精度

表 3 各手法における歩行者の検出率 (FPPF = 0.5)

手法	傘をさした歩行者	傘をさしていない歩行者
提案手法	0.54 (684/1,279)	0.57 (304/534)
従来手法	0.40 (517/1,279)	0.59 (316/534)

を用いて検出精度を算出した。

3.4 実験結果と考察

提案手法と従来手法による歩行者検出の精度を図 14 に示す。図 14 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)をとったグラフである。検出によって描画された曲線が左上に位置するほど検出精度が高いことを表す。また、グラフ上の FPPF = 0.5 の点における各手法の検出率を表 3 に示す。これらは生成及び学習・検出を 3 回ずつ独立に行なった結果の平均である。同様に、FPPF = 0.5 の点における提案手法と従来手法の検出結果画像例を図 15 に示す。

図 14 から、提案手法は従来手法に比べて精度が高いことがわかる。この結果から、傘画像を生成し歩行者画像に合成することで、傘を考慮しない従来手法よりも精度良く雨天時の車載カメラ画像から歩行者を検出できることが分かった。さらに表 3 から、提案手法と従来手法における傘をさしていない歩行者

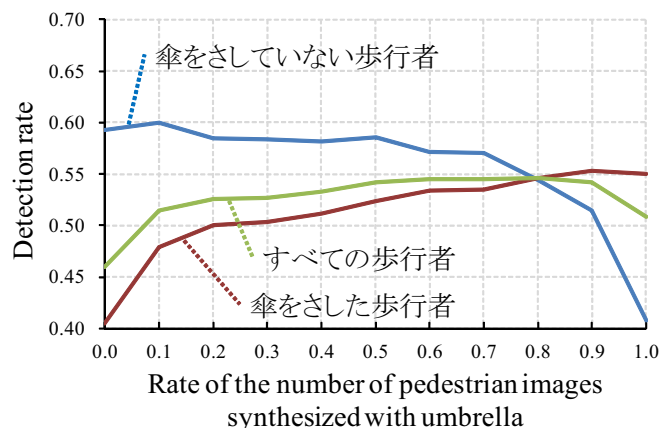


図 16 学習に使用する歩行者画像への傘画像を合成する枚数の割合と歩行者検出精度の関係

者の検出率の差はわずかに 2% であるが、傘をさした歩行者の検出率は提案手法の方が 14% 高いことがわかる。また、図 15 に示した提案手法と従来手法の検出結果を比べると、提案手法では傘をさした歩行者を正しく検出できたことがわかる。特に、提案手法では図 15 (a) 及び (d) のように傘によって完全に頭部の隠れてしまった歩行者であっても正しく検出することができた。一方で、同様のシーンに従来手法を適用した結果では、これらの歩行者を検出することができなかった。提案手法では傘画像を合成し生成した歩行者画像を学習したため、このような歩行者も検出することができたと考えられる。

本実験では、傘を合成した歩行者画像 70% (3,500 枚) と傘を合成しない歩行者画像 30% (1,500 枚) を提案手法の学習に使用した。図 16 に傘画像の合成割合と検出精度の関係を示す。この結果も各傘画像の合成比率に対して生成及び学習・検出を 3 回ずつ独立に行なった結果の平均である。図 16 から 70% の歩行者画像に傘画像を合成したとき、最も精度良く歩行者を検出していることがわかる。この割合は、評価画像に含まれる傘をさした歩行者の割合と一致する。一方で、すべての学習用歩行者画像 (5,000 枚) に傘画像を合成し学習したところ、FPPF = 0.5 のときに傘をさした歩行者は 55% 検出できたが、傘をさしていない歩行者は 41% しか検出できなかった。このことから、評価画像に含まれる傘をさした歩行者の割合に応じて、学習に用いる傘画像を合成した歩行者画像と合成していない歩行者画像の割合を調整する必要があると言える。よって、仮にすべての人が傘をさしているような状況で撮影された評価画像に提案手法を適用するのであれば、傘画像を合成した歩行者画像のみを用いて学習することが効果的である可能性がある。実際のアプリケーションでは、雨の強さをセンサから取得することで傘をさした歩行者の割合を予測し、傘画像を合成する歩行者画像の割合を変化させて学習した複数の識別器から、最適なものを選択することで高精度な歩行者検出が実現できると考えられる。

また、図 15 の (e) を見ると、雨で濡れた路面の反射が誤検出されていることがわかる。図 17 に誤検出された路面の反射

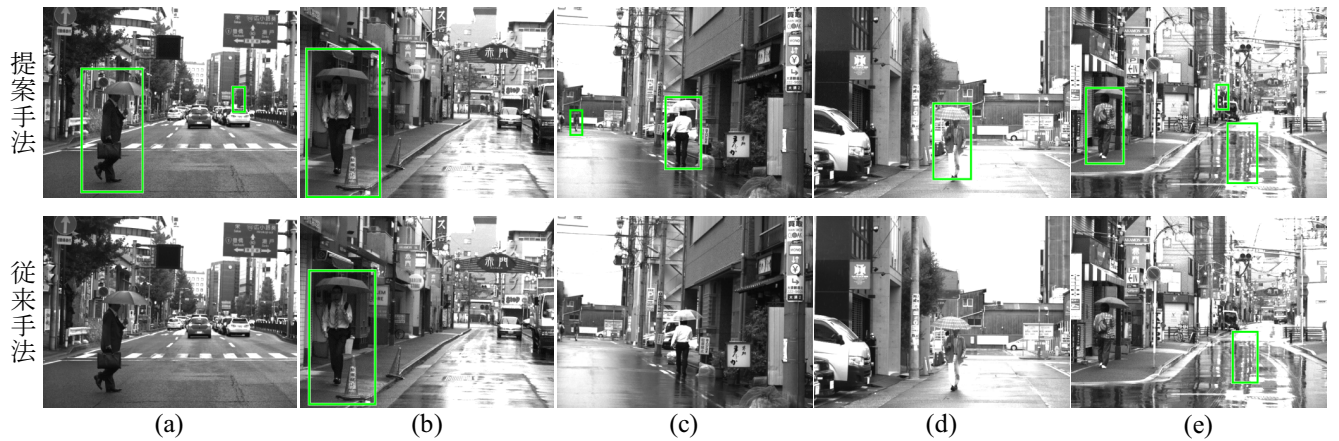


図 15 提案手法及び従来手法の検出結果の例

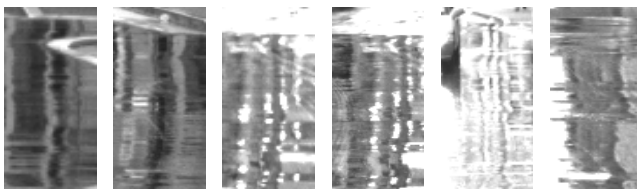


図 17 誤検出された路面の反射の例

の例を示す。これは雨天特有の状況であり、反射によって生まれる輝度勾配特徴が歩行者の特徴に似ていたことや、学習に使用した非歩行者画像にこのようなサンプルが含まれていなかったことが誤検出の原因として挙げられる。こうした雨天特有の非歩行者画像を学習に使用することで誤検出を抑制し、検出精度のさらなる向上が可能であると考えられる。

4. む す び

本報告では、生成型学習法を用いた傘をさした歩行者の高精度な検出手法を提案した。提案手法は、傘モデルから生成した傘画像を歩行者画像に合成し、学習に用いることで傘をさした歩行者の高精度な検出を実現した。

本報告では傘画像の生成に焦点を当てて報告したが、提案手法は傘だけでなく鞆など歩行者の見えを大きく変化させる物に対しても同様に適用することが可能である。傘をさした歩行者以外への適用例としては、例えばランドセルを背負った小学生の画像生成が考えられる。通学路など小学生の多い地域で、生成した画像を学習した識別器を用いて歩行者検出を行なうことにより、近年多発している通学途中の小学生を巻き込んだ交通事故を軽減することができると考えられる。よって、傘以外の持ち物に対する提案手法の適用が今後の課題として挙げられる。また、傘画像生成のパラメータと検出精度の関係についても詳しい調査が必要である。

謝辞 日頃より熱心な御指導と御鞭撻を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、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。

文 献

- [1] N. Dalal, and B. Triggs, “Histograms of oriented gradients for human detection,” Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.1, pp.886–893, June 2005.
- [2] 村瀬洋, “画像認識のための生成型学習,” 情処学論, vol.46, no.SIG 15 (CVIM 12), pp.35–42, Oct. 2005.
- [3] 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.
- [4] 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,” Proceedings of the 11th IAPR Conference on Machine Vision Applications, pp.514–517, May 2009.
- [5] M. Enzweiler, and D.M. Gavrila, “A mixed generative-discriminative framework for pedestrian classification,” Proceedings of 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.1–8, June 2008.
- [6] 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,” Proceedings of the 7th International Conference on Computer Vision Theory and Applications, vol.1, pp.567–572, Feb. 2012.
- [7] J. Marín, D. Vázquez, D. Gerónimo, and A.M. López, “Learning appearance in virtual scenarios for pedestrian detection,” Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.137–144, June 2010.
- [8] 草富省吾, 山内悠嗣, 藤吉弘亘, “人検出のための学習サンプルの自動生成と MILBoost を用いた学習法,” 画像センシングシンポジウム (SSII2011), IS1-15, no.188, pp.1–8, June 2011.
- [9] 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.
- [10] M. Enzweiler, A. Eigenstetter, B. Schiele, and D.M. Gavrila, “Multi-cue pedestrian classification with partial occlusion handling,” Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.990–997, June 2010.
- [11] A. Carpenter, “cuSVM: A CUDA implementation of support vector classification and regression,” <http://patternsonscreen.net/cuSVM.html>.