

誤検出傾向に着目した走行環境適応型歩行者検出器の構築に関する検討

鈴木 悠暉[†] 出口 大輔^{††} 井手 一郎[†] 村瀬 洋[†]

[†] 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

^{††} 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: †{suzuki}@murase.m.is.nagoya-u.ac.jp, †{ddeguchi,ide,murase}@is.nagoya-u.ac.jp

あらまし 近年、車載カメラを用いた安全運転支援技術の研究開発が盛んに行われている。その中でも歩行者検出は重要な課題であり、さまざまな手法が提案されている。車載カメラで撮影した歩行者の見えは、天候や時刻、走行場所などによって多様に変化する。そのため、単一の検出器で多様な見えを含む全ての走行環境に対応することは困難である。そこで本報告では、走行環境の違いにより観察される誤検出傾向に着目し、環境毎に適した歩行者検出器を構築する手法を検討する。走行環境を撮影した画像に通常の汎用的な歩行者検出器を適用した際、走行環境特有の誤検出が得られる。この環境特有の誤検出傾向を利用して、複数の歩行者検出器を構築し、それらを用いて歩行者検出の検出精度向上を図る。提案手法の有効性を確認するために、公開データセットを用いた評価実験を行った。その結果、従来の手法を用いて構築した歩行者検出器と比べ、提案手法によって検出精度が向上することを確認した。

キーワード 適応型学習, 誤検出傾向, 歩行者検出

A study on construction of pedestrian detectors adaptive to driving environment referring to false detection tendency

Yuki SUZUKI[†], Daisuke DEGUCHI^{††}, Ichiro IDE[†], and Hiroshi MURASE[†]

[†] Graduate School of Information Science, Nagoya University

^{††} Information and Communications Headquarters, Nagoya University

Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

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

Abstract Recently, safety driving assistance systems using an in-vehicle camera have been studied actively. In particular, since pedestrian detection by image processing is an important task, various methods have been proposed. In real environment, appearance of pedestrians varies largely due to the difference of driving environments such as weather condition, time, driving scene, etc. Therefore, it is difficult to handle all driving scene by a single detector. This report focuses on the false detection tendency according to various driving environments, and constructs pedestrian detectors adaptive to each environment. By applying a conventional pedestrian detector to an in-vehicle camera image, false detections specific to each driving environment are obtained. By referring to these false detections, the proposed method constructs pedestrian detectors adaptive to a driving environment, and they are used for improving the detection accuracy. We conducted experiments using a public dataset to evaluate the proposed method. As a result, accuracy of the proposed method was improved in comparison with a conventional pedestrian detector.

Key words adaptive learning, false positive tendency, pedestrian detection.

1. まえがき

警視庁の統計 [1] によると、日本の交通事故死者数は減少傾向にあるものの、平成 26 年においてもまだ 4114 人もの尊い命が失われている。その中でも自動車と歩行者の衝突事故の割合が高く、そのほとんどは運転手による歩行者の見落としが主な原因である。そこで近年では、自動車と歩行者の衝突そのものを防ぐ技術が注目を集めており、自動車の走行環境下にいる歩行者の検出などの安全運転支援技術の研究開発が盛んに行われている。特にその中でも、歩行者の検出においては車載カメラを用いた手法が数多く提案されている。

車載カメラ画像を用いた歩行者検出の代表的な手法として、輝度勾配の情報を利用した Histogram of Oriented Gradients (HOG) 特徴量と Support Vector Machine (SVM) 識別器を組み合わせた手法 [2] (以降、HOG+SVM 検出器と呼ぶ) がある。この手法では、歩行者検出器の学習に歩行者画像と背景画像 (歩行者を含まない走行環境を撮影した画像) から HOG 特徴量を抽出し、識別器を構築する。

一般に、撮影した季節や撮影時刻による明るさの変化、走行環境の背景などの複数の要因によって、車載カメラ画像の見えは多様に変化する。また歩行者自身においても、歩行者の向きや所有物などにより多様な見えの変化が生じる。このような多様な見えの変化に対応するためには、多様な見えの変化をもつ大量の画像を用いて歩行者検出器を学習する必要がある。しかし、単一の検出器で多様な見えの変化をもつ全ての走行環境に対応することは困難である。

そこで、特定の環境下に適応した識別器を学習する、適応型学習 [3] が提案されている。適応型学習は、画像の見えに応じて学習画像をクラスタリングし、学習画像クラスタ毎に識別器を構築する手法である。特定の環境下に適応することにより、識別器構築時と検出時の走行環境が一致する場合には高精度な検出が可能である。鈴尾ら [3] [4] は場所ごとに走行環境の見た目が異なる点に着目し、車載カメラ画像の見えが似た走行環境と場所を関連付けて複数の識別器を構築する手法を提案している。しかし、鈴尾らの研究では走行環境のクラスタリングと歩行者識別器の構築で異なる特徴量を用いている。そのため、識別器の特性と学習サンプルのクラスタリング結果に整合性がなく、誤検出の低減が必ずしも得られないという問題がある。

そこで本報告では、識別器の構築に用いる特徴量と環境のクラスタリングに同じ特徴量を用いる。そして、識別器の特性として現れる誤検出傾向に着目した走行環境分類手法を提案する。誤検出は歩行者識別器を通して得られるため、識別器の特性を反映していると考えられる。そのため、誤検出傾向に着目して学習サンプルをクラスタリングすることによって、検出器の特性に合った走行環境の分類を図る。

以降、2 節で提案手法のアイデアについて説明し、3 節で提案手法の詳細、4 節で提案手法の有効性を確認するために行った実験の結果を示す。

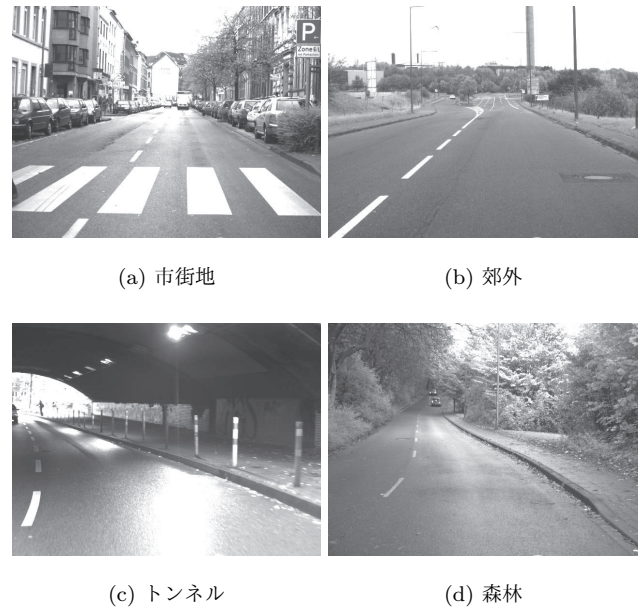


図 1 車載カメラ画像の見えの違い



図 2 歩行者として検出された誤検出物体の例

2. 提案手法のアイデア

学習型の識別器を利用する場合、一般に車載カメラ画像の多様な見えに起因して多くの誤検出が生じる。この問題は、HOG+SVM 検出器を用いた歩行者検出でも同様である。実際の車載カメラ画像系列の例を図 1 に示す。図 1 からわかるように、それぞれの走行環境ごとに特有の見えが存在する。走行環境適応型検出は、特有な見えを持つ走行環境ごとに学習サンプルをクラスタリングし、各環境に適した歩行者識別器を構築する手法である。各環境特有の見えを学習することにより、誤検出を削減することが可能である。

鈴尾らは場所ごとに走行環境の見た目が異なる点に着目し、走行環境の背景画像をクラスタリングしている。その際、走行環境のクラスタリングと歩行者識別器の構築では異なる特徴量を用いている。しかし、走行環境のクラスタリングにおいては、人が似ていると感じる走行環境の見えと歩行者識別器が似ていると感じる走行環境の見えは必ずしも一致しない。識別器の特性に合わせて環境を分類して歩行者検出器を構築すれば、さらなる誤検出が削減が期待できる。

誤検出に着目して識別器の特性を考えると、同じ識別器であ

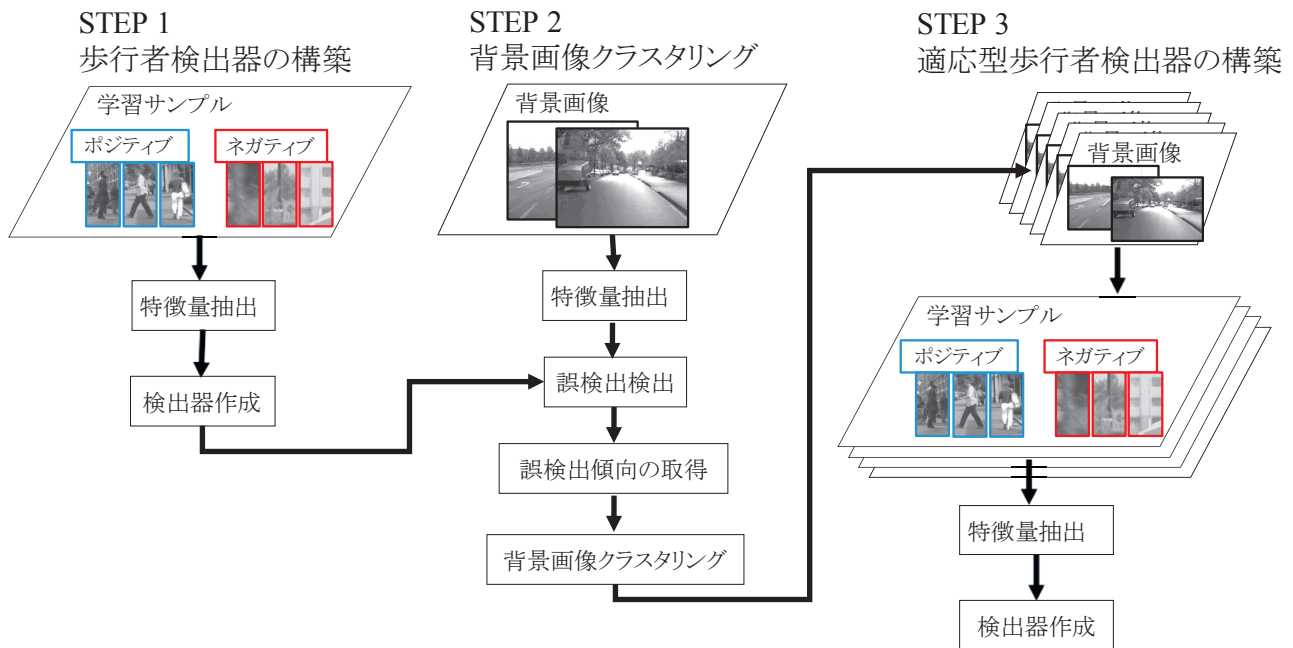


図 3 環境適応型検出器の構築手法の概要

れば誤り方は常に同じである。そこでこの考えに基づき、本稿では誤り方の傾向にあわせて走行環境を分類する手法を提案する。

車載カメラ画像に HOG+SVM 検出器を適用して得られた誤検出の例を図 2 に示す。図 2 における誤検出の例として、電柱や木などの直線的な物体、信号および標識、車のタイヤなど、といった円形状の物体の周りに誤検出が見られる。一方、人が見た場合に歩行者とは間違えようのない木の葉の影やビルの窓などの誤検出も多く見られる。このように、人から見て歩行者と似ていると感じる感覚と歩行者検出器が似ていると判断した結果は必ずしも一致しない。また、歩行者検出によって得られた誤検出には、走行環境のどこにでも存在しうる物体から得られる誤検出と走行環境ごとに特有の誤検出が同時に存在することもわかる。このような誤検出の傾向は走行環境を特徴づける情報だと考えられる。そこで、誤検出をクラスタリングすることで類似した誤り方のパターンを求める。本稿では、これを誤検出傾向と呼び、これを利用して歩行者識別器を構築する手法を提案する。

3. 誤検出傾向に着目した走行環境適応型歩行者検出

本稿では、誤検出傾向を利用して走行環境適応型の歩行者識別器を構築する手法を提案する。提案手法では、歩行者検出器を（歩行者の写っていない）背景画像に適用した際、歩行者として誤検出された特徴量のクラスタリング結果を誤検出傾向として利用する。これは、識別器の誤り方を分類していると考えられることができる。そして、この誤検出傾向に基づいて走行環境を分類する。この誤検出傾向を利用した歩行者識別器の構築の流れを図 3 に示す。提案手法は、歩行者識別器の構築、誤検出

傾向を利用した背景画像クラスタリング、背景画像クラスターごとの歩行者識別器の構築、の 3 段階からなる。

以降では、それぞれの段階について詳細を述べる。

3.1 歩行者識別器の構築

図 3 に示す歩行者識別器の構築の流れの step1 に従い、全ての環境の学習データを用いて HOG+SVM 検出器 [2] を構築する。まず、歩行者画像と背景画像からランダムに切り出した背景画像を学習サンプルとして入力する。そして、それぞれの学習サンプルから HOG 特徴量を抽出し、SVM を用いて歩行者と背景を識別する識別器を構築する。背景画像からランダムに切り出した画像から得られる特徴量は、必ずしも歩行者と背景の識別に有効であるとは限らない。そのため、ランダムに切り出した背景サンプルを用いて学習した歩行者識別器では多くの誤検出が生じる。そこで、歩行者識別器の誤検出低減を目的として、背景サンプルの再収集を行う。具体的には、上記で構築した HOG+SVM 検出器によって生じる誤検出を新たに加えて、識別器を再学習する。歩行者識別器の再学習の際、図 4 に示すように線形 SVM の分離平面とサポートベクトル間のマージンに含まれる誤検出のみを新たにネガティブサンプルに加える。

3.2 適応型歩行者検出器の学習

次に、図 3 の step2 に示す誤検出傾向を利用した歩行者識別器の構築を行う。まず、図 3 の step1 で構築した歩行者識別器を用い、走行環境に現れる誤検出から誤検出傾向を求める。さらに、誤検出傾向に合わせて走行環境を分類し、それぞれの走行環境に適した HOG+SVM 識別器を構築する。これにより、走行環境に適した複数の HOG+SVM 識別器が得られる。以降で誤検出傾向に着目した歩行者識別器の構築について詳細に説明する。

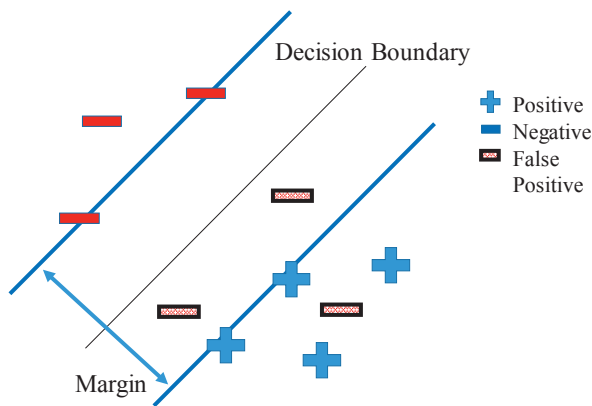


図 4 線形 SVM の分離平面およびマージンの概要

3.2.1 誤検出傾向に基づく走行環境の分類

2 節で、歩行者検出器の類似した誤り方のパターンを誤検出傾向とすることを述べた。検出器から見て歩行者のよう見える木のパターンや信号のパターンといったように、それぞれの誤検出クラスタから誤検出傾向を求める。つまり、誤検出クラスタだけ誤検出傾向は存在する。そして図 5 に示すとおり、走行環境の分類に用いる誤検出傾向を得た後、誤検出傾向に従って走行環境を分類する。

まず、各背景画像に環境適応なしの検出器を適用して誤検出を得る。得られた誤検出に k -means クラスタリング手法を適用し、各誤検出クラスタ中心を求める。HOG 特徴量は高次元であるため、誤検出の傾向を反映したクラスタリングができない可能性がある。そこで、Principle Component Analysis (PCA) を用いて誤検出クラスタを適切に分離できる空間に写像する。

次に、誤検出傾向と走行環境の背景画像に現れる誤検出を用いて走行環境を分類する。具体的には、走行環境がどの誤検出傾向を多く含むのかによって走行環境を分類する。各走行環境に現れた誤検出に最も近い誤検出クラスタ（誤検出傾向）を求め、各誤検出クラスタに属する誤検出数を算出する。走行環境に最も多く含まれる誤検出傾向に対して、一定割合以上の誤検出数をもつクラスタと背景画像を結びつける。これにより、各誤検出クラスタに対して複数の走行環境画像がひもづけられる。以上が、環境適応型歩行者識別器の構築に用いる、背景画像（ネガティブサンプル）の収集方法である。

3.3 適応型歩行者検出器の学習

誤検出傾向を利用して分類した走行環境の背景画像をネガティブサンプルとして用いて、誤検出クラスタ毎に歩行者識別器を構築する。各クラスタの歩行者識別器構築に用いるポジティブサンプルには、環境適応前の検出器構築に利用した歩行者画像を用いる。誤検出傾向毎に歩行者識別器を構築することは、それぞれの環境に適応した歩行者検出器を構築することに等しい。環境適応型歩行者検出器を構築する際、誤検出の抽出に用いた HOG+SVM 検出器と同様に検出器の再学習を行う。これらの環境適応後の歩行者検出器では、走行環境を写した入力画像における誤検出傾向が背景画像クラスタと一致する場合

に高精度な歩行者検出が可能である。

3.4 歩行者検出

複数サイズの矩形の検出窓を用いて、セル単位で検出窓を移動させながら歩行者の検出を行う。各検出窓から抽出した特徴量と歩行者識別器の重みによって得られる識別スコアがしきい値以上の場合に、検出窓が示す範囲を歩行者として検出する。その結果、歩行者の周りには検出窓が複数現れることがある。このような検出窓は、Mean-shift クラスタリングを用いて統合する。そして、統合後の検出矩形を最終的な歩行者として出力する。

提案手法は、誤検出クラスタ毎に複数の誤検出傾向に合った検出器を構築し、それぞれ統合した後の検出結果を用いる。さらに 3.1 節で構築した検出器の検出結果を加え、最終的な検出結果は全ての検出矩形を重み付き多数決法で求める。まず、各検出器から得られた検出結果が全検出器の過半数から得られない場合、その検出結果は歩行者でないと判断する。各検出矩形の重みは、図 4 で示した線形 SVM の分離平面からの距離で与えられる。また、検出結果を統合したときの重みは、統合された各検出結果の重みの総和である。一般に、全ての検出器において歩行者の周辺には多くの検出矩形が現れるため、検出矩形を統合するときに重みが増す。一方で、誤検出傾向に合った検出器の検出結果では、誤検出傾向が似た誤検出は現れないため重みは低くなる。そのため、誤検出と比べて歩行者を示す検出矩形の重みが高くなり、歩行者が検出されやすくなる。

4. 評価実験

提案手法の有効性を確認するために、評価実験を行った。現在広く用いられている Dalal ら [2] の検出手法と 3 節で述べた提案手法を比較した。以降、実験に用いたテストデータの説明と検出器の構築パラメータについて順に述べる。

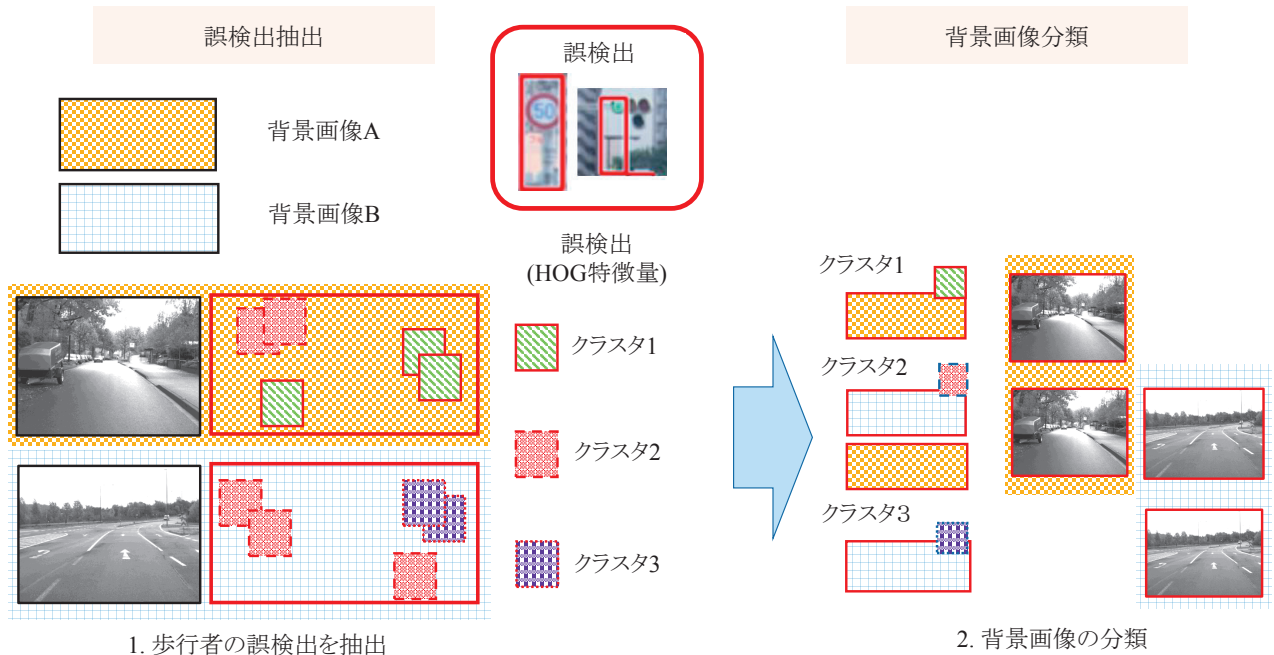
4.1 実験データ

実験データには、Daimler 社により提供されているベンチマーク用のデータセット [5] を用いた。Daimler データセットは、単眼の車載カメラを用いて町並みを撮影したグレースケール画像で構成されている。本実験では、テスト画像の中から 8000 フレームを用いた。また、各フレームに歩行者を囲む正解矩形が与えられている。データセットの特徴として、車載カメラ画像の撮影画角が狭く、画像の中心に道路が大きく映っている。

学習サンプルの収集では、歩行者ラベルのついた学習画像から自動的に歩行者画像と非歩行者画像を切り出し、それぞれポジティブサンプルとネガティブサンプルとした。ポジティブサンプルは、左右反転に加え、切り出し枠を平行移動したものをを用いた。一方、ネガティブサンプルは、歩行者を含まない非歩行者画像領域からランダムに切り出しを行い、ポジティブサンプルと同じスケールとなるように線形補間により拡大縮小したものをを用いた。

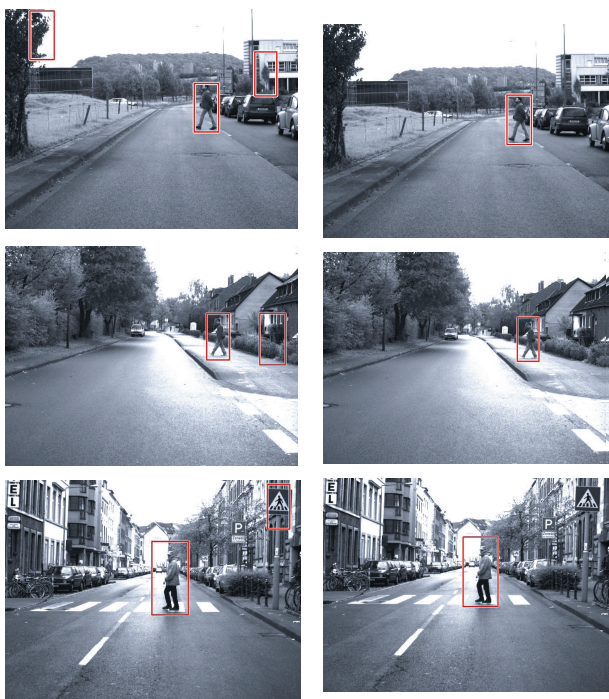
4.2 誤検出傾向を用いた歩行者検出器の構築パラメータ

本実験における、検出器の構築パラメータを記す。歩行者検出を行う際の検出窓の最小サイズは、横幅 48 画素、縦幅 96 画



背景画像の誤検出クラスタには、歩行者の情報が一切含まれていない

図 5 誤検出傾向を利用した背景画像クラスタリングの概要



(a) 従来手法

(b) 提案手法

図 6 各手法における歩行者検出例

素である。歩行者検出器の構築に用いる HOG 特徴量は、セルサイズ 6 画素でブロックサイズが 3 セルの 6084 次元の特徴量である。また、誤検出傾向として利用する誤検出クラスタ数は 5 とした。

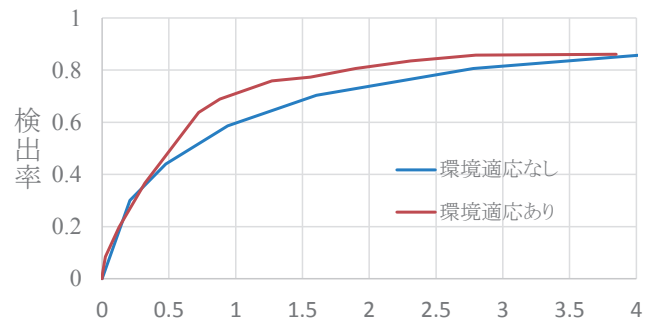


図 7 各手法による FROC 曲線

5. 実験結果および考察

誤検出傾向を利用した走行環境への適応による提案手法の有効性を確かめるために、本手法と環境適応なしの歩行者検出器を用いてテストデータ画像から歩行者検出を行った。歩行者検出の正解判断は PASCAL VOC Dataset [6] に準拠し、正解データの矩形と検出矩形の重なりがしきい値以上の検出矩形を歩行者とみなした。本報告では、しきい値が 0.3 以上の場合に歩行者とみなした。歩行者検出結果の評価指標には、縦軸に検出率、横軸に 1 フレームあたりの false positive である FPPF をとる FROC 曲線を用いた。FROC 曲線は、曲線が左上に近づくほど検出精度が高いことを示す。

Daimler データセットのトレーニング画像に対して比較手法の歩行者検出器を適応した際に得られた検出結果と提案手法による歩行者検出器を適応した際に得られた検出結果を図 6 に示



(a) 従来手法 (b) 提案手法

図 8 誤検出と検出漏れの関係

す。実験の結果、環境適応なしの歩行者検出器で見られた木や建物の周辺に現れた誤検出が環境適応ありの歩行者検出器では削減されている。

また、図 7 の FROC 曲線からわかるように、提案手法は環境適応なしの検出器に比べ FPPF の値が 1 のときの検出率が 0.1 向上している。このように、誤検出の削減ができ、検出精度の向上が見られたことから提案手法の有効性が確認された。

一方、FPPF が小さい時の検出率は低いことがわかる。これは、歩行者検出器の多数決手法において、誤検出傾向に合った検出器を用いて誤検出を削減したのと同時に、歩行者が検出されにくくなったためと考えられる (図 8)。

6. む す び

本報告では、車載カメラを用いた歩行者検出器の高精度化を目的とし、誤検出傾向に着目した走行環境適応による歩行者識別器の構築手法を提案した。誤検出傾向に基づいて走行環境の背景画像の収集とクラスタリングをし、各クラスタに対して歩行者識別器を構築した。Daimler データセットを用いた実験により、従来広く用いられている Dalal ら [2] の HOG+SVM 検出器の検出精度の向上が可能であることを確認した。今後の課題として、検出器の精度向上のために検出手法の改善を図る。例えば、環境適応型歩行者識別器の切り替えによる手法が考えられる。また、HOG+SVM 検出器と比較して検出精度の高い歩行者検出手法に本稿で書いた誤検出傾向に着目した走行環境適応型学習を適用する予定である。

謝 辞

村瀬研究室諸氏に深く感謝する。本研究の一部は、戦略的創造研究推進事業 CREST、科学研究費補助金による。

文 献

- [1] 警視庁交通局, “平成 26 年中の交通事故死者数について,” Jan. 2015.
- [2] N. Dalal et al., “Histograms of Oriented Gradients for Human Detection,” Proc. 2005 IEEE Computer Society Conf.

on Computer Vision and Pattern Recognition , pp.886–893, June, 2005.

- [3] 鈴尾大地, 出口大輔, 井手一郎, 村瀬 洋, 石田皓之, 小島祥子, “位置情報を用いた走行環境適応による車載カメラ画像からの歩行者検出の高精度化,” 信学技報, PRMU2014–113, Mar, 2014.
- [4] D. Suzuo et al., “Environment Adaptive Pedestrian Detection using In-vehicle Camera and GPS,” Proc. Int. Conf. on Computer Vision Theory and Applications 2014, pp.354–361, Feb, 2014.
- [5] M. Enzweiler et al., “Monocular Pedestrian Detection: Survey and Experiments,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.31, no.12, pp.2179–2195, Dec, 2009.
- [6] M. Everingham et al., “The PASCAL Visual Object Classes (VOC) Challenge,” Int. J. of Computer Vision, vol.88, no.2, pp.303–338, June, 2010.