

車載カメラを用いた走行環境適応型歩行者検出器に関する予備的検討

鈴尾 大地[†] 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 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 情報連携統括本部情報戦略室 〒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 愛知県長久手市横道 41-1

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

あらまし 近年、車載カメラ画像から歩行者を検出する研究が注目されている。しかし、時刻や天候、季節、走行場所などの環境要因によって歩行者の見えは大きく変動するため、高精度に検出することは難しい。画像中の歩行者の見えを特徴とした従来の検出手法では、事前に大量の学習用歩行者画像および非歩行者画像を人手で収集することで、このような変動に対応していた。しかし、あらゆる見えの変動を想定して学習用画像を用意することは難しい。本研究では、車載カメラ映像から学習用画像を収集して歩行者検出器を走行環境に適応させる枠組みを提案し、この問題の解決を図る。本報告では検出精度に影響を与える主要な環境要因である走行場所の変化に着目し、走行場所毎に歩行者画像および非歩行者画像を収集することで走行環境適応型歩行者検出器を構築する手法を提案する。理想的に適応が行われた歩行者検出器の精度を検証したところ、従来手法と比較して大幅な精度向上が見られ、有効性を確認した。

キーワード 歩行者検出、適応的検出器、HOG、SVM、ITS

A preliminary study on a pedestrian detector adaptive to the driving environment using an in-vehicle camera

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

[†] Graduate School of Information Science, Nagoya University Furo-cho, Chikusa-ku, Nagoya-shi, Aichi,
464-8601 Japan

^{††} Strategy Office, Information and Communications Headquarters, Nagoya University Furo-cho,
Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

^{†††} Toyota Central Research and Development Laboratories, Inc. Yokomichi 41-1, Nagakute-shi, Aichi,
480-1192 Japan

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

Abstract Recently, pedestrian detection from in-vehicle camera images is being focused. However, it is difficult to detect pedestrians accurately due to the variety of their appearances caused by various environmental factors such as time, weather, season, and location. To deal with changes of pedestrians' pose and environment, most existing methods that makes use of their appearances require the preparation of a lot of pedestrian images manually. However, it is difficult to provide training images considering all possible environmental factors. To tackle this problem, we propose a method to make a pedestrian detector adaptive to a driving environment by collecting training images from in-vehicle camera images. We focused on the variety of the appearance per location that primarily affects the accuracy of pedestrian detection, and built an environment adaptive pedestrian detector by collecting pedestrians images and non-pedestrian images in each driving location. In evaluation experiments assuming ideal adaptation to the environment, we confirmed the effectiveness of the method compared to existing methods.

Key words Pedestrian detection, Adaptive detector, HOG, SVM, ITS

1. まえがき

平成 22 年の交通事故死者数 4,863 人のうち、歩行者は全体の 35% (1,714 人) と大きな割合を占めている [1]。そのため、運転者へ歩行者の存在を知らせるなどの自動車運転支援を目的とした高精度な歩行者検出技術が求められており、これまでに車載カメラ画像を対象とした歩行者検出手法が広く研究されている。

Dalal らは Histogram of Oriented Gradient (HOG) 特徴量と Support Vector Machine (SVM) を用いた歩行者検出手法を提案し [2]、現在さまざまな応用で利用されている。しかしながら、高精度な歩行者検出を実現するためには、大量の歩行者画像および非歩行者画像が識別器の学習に必要である。

これに対して、歩行者を含む映像中より学習用画像を自動収集するというアプローチが提案されている。Wöhler らは、歩行者検出と歩行者追跡の組み合わせにより、検出された歩行者を学習サンプルとして動的に収集する枠組みを提案している [3]。また、歩行者検出の結果から学習用画像を集める際、誤検出による検出精度低下を防ぐため Wang らは画像情報を用いて正しい歩行者のみを収集する手法を提案し [4]、服部は時系列情報を用いて正しい歩行者のみを収集する手法を提案している [5]。

これらの手法は検出器構築のコストを低減させることを目的としており、走行環境の違いによる車載カメラ画像の見え方の違いを考慮して効果的に適切な学習用画像を収集するものではない。そのため、走行環境の違いによる車載カメラ画像の見え方の変化へ対応する手法とはなっていない。図 1 は、季節による服装変化に伴う歩行者画像の見え方の変化、走行場所の変化に伴う非歩行者画像の見え方の変化を表したものである。図 1 からわかるように、走行環境によって見え方は多様に変化するため、単に車載カメラ映像から学習用画像を収集し、それらを学習に用いるだけでは、それぞれの走行環境で精度良く歩行者を検出することは難しい。

これに対し、学習用画像を走行環境毎に分類して収集し、実際に歩行者検出を行う走行環境に適した画像のみを学習に用いることで、検出精度の向上が期待できる。つまり、歩行者検出器を走行環境に適応させることを意味する。これは、ある走行環境に適応させた歩行者検出器を構築するために、その走行環境を撮影した映像のみから学習用画像を収集することで実現することが可能である。これにより、走行環境の違いによる歩行者の見え方の変化に対して頑健な歩行者検出を達成できると考えられる。

そこで本報告では、走行環境毎の映像から学習用画像を収集し、歩行者検出器を構築することで、その走行環境における検出精度の向上が可能であることを定量的な評価を通して示す。

以降、2 節で歩行者検出器を走行環境に適応させるための枠組みを提案する。3 節において提案手法に基づいた歩行者検出の環境適応の有効性の検証方法について説明し、4 節において検証実験の結果および考察を述べる。最後に、5 節で本報告をまとめる。



図 1 走行環境の違いによる見え方の変化の例。

2. 走行環境の変化と歩行者検出精度の関係

歩行者検出器の精度に影響を与える走行環境の変化の例として、次の 3 つが挙げられる。

- 走行場所の変化による背景の見え方の変化
- 時刻の変化による照明の変化
- 天候・季節の変化による歩行者の服装の変化

このうち、時刻や天候・季節による歩行者の見え方の変動パターンは限定的であり、それらへの対処は比較的容易である。一方、走行場所の違いによる変動範囲は極めて大きく、事前に全ての変動を想定し学習用画像を収集して対応することは不可能である。そこで本報告では、走行場所の違いに着目した環境適応型歩行者検出器の構築を行う。

3. 走行場所へ適応した歩行者検出器の構築

提案手法は学習段階と検出段階からなる。

学習段階は、歩行者検出器を走行場所に適応させる処理である。事前に人手で集められた学習用画像から構築した歩行者検出器を入力とし、走行場所に適した歩行者検出器を構築する。このために、車載カメラ映像から各走行場所の学習用画像を収集し、事前準備された学習用画像に追加し再び学習を行う。

検出段階は、各走行場所に適応した歩行者検出器を用いて、自車の現在位置に適応した歩行者検出器を選択し歩行者検出を行う処理である。

なお提案手法では、学習段階および検出段階ともに自車位置の取得が必要となる。このためには、GPS などを用い車載カメラ映像の撮影時に位置情報を併せて記録することなどが考えられる。

以降、学習段階および検出段階の詳細について述べる。

3.1 歩行者検出器の走行場所への適応

歩行者検出器は、走行中に通過する各場所のそれぞれに対し

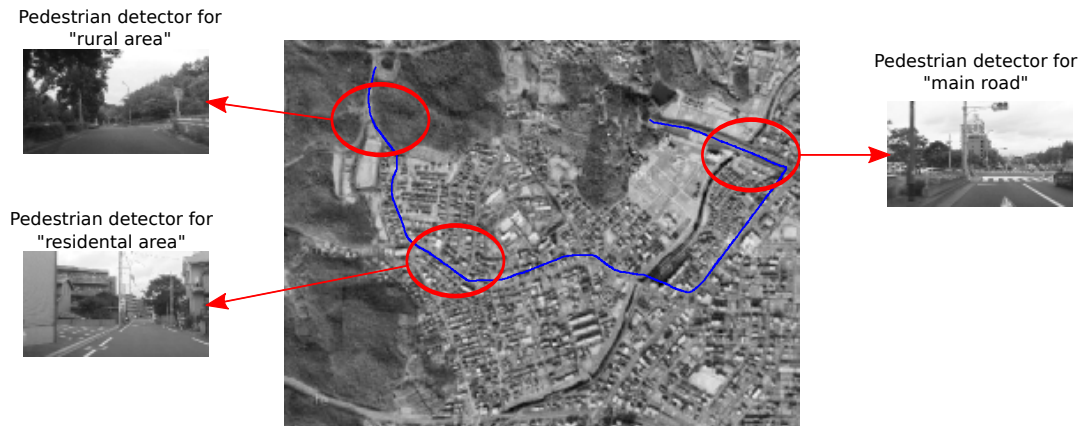


図2 走行場所へ適応した歩行者検出器の概念. 走行場所ごとの車載カメラ画像の見えの変化に着目し, それぞれに適応した歩行者検出器を構築する.

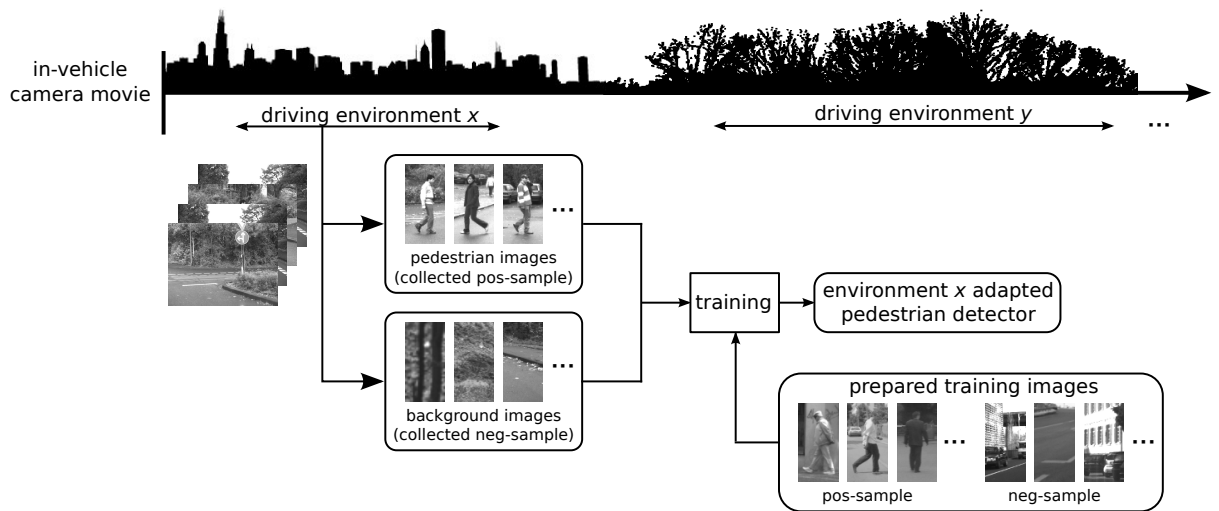


図3 走行場所の映像からの学習用画像の収集と歩行者検出器の学習.

て適応される. 従って, 走行場所が変化するとそれに適応した新たな歩行者検出器が構築されることとなる. つまり, 図2に示すように, 自車の走行位置それぞれに対して適応した歩行者検出器が関連付けられる.

歩行者検出器を走行場所に適応させる処理は, 大きく分けて走行場所の識別, 学習用画像の収集, それを用いた学習, の3段階からなる. 次節よりこれらについて説明を行う.

3.1.1 走行場所の識別

歩行者検出器を走行場所に適応させるための学習用画像を収集する際, その走行場所のみを撮影した映像が必要となる. そこで提案手法では, 映像の撮影された位置の情報を用いることで, 走行場所毎の映像を収集する. 各走行場所の映像は, 撮影地点の前後のある程度の区間を含むように切り出す.

3.1.2 学習用画像の収集

特定の走行場所を撮影した映像が得られた後, 学習用画像として歩行者画像および非歩行者画像の収集を行う.

歩行者画像は, 映像から検出した歩行者を切り出して収集する. また, 非歩行者画像は, 歩行者と重ならないように映像よりランダムな位置およびスケールの矩形を切り出し収集する. この処理を図3に示す.

3.1.3 歩行者検出器の再学習

本手法では既存の歩行者検出器を走行場所に適応させることを目的としているため, 学習用画像は事前に準備されたものが既にあるものとする. 前述の処理により収集された学習用画像は, 事前準備された学習用画像と併せて歩行者検出器の学習に用いられる. これにより, 走行場所に適応した歩行者検出器を構築する. この歩行者検出器は3.1.1節の処理で識別された走行場所に関連付けられ, 図2に示すような走行場所と歩行者検出器の対応関係を得る.

3.2 走行場所に適応した歩行者検出器による歩行者検出

歩行者検出器の走行場所への適応により図2に示すような走行場所と歩行者検出器の対応関係が得られる. 歩行者検出を行う際には, この対応関係を利用して自車の走行位置と対応付けられた歩行者検出器を選択し, 車載カメラ画像を入力する. これにより, 高精度な歩行者検出が期待できる.

4. 実験

提案手法の有効性を確認するための実験を行った. 本実験では, 少数の学習用画像を用いて構築された歩行者検出器を提案手法により複数の走行環境にそれぞれ適応させた場合の検出精

表 1 映像から分類した走行環境の一覧

走行環境	開始時間	終了時間	フレーム数	延べ歩行者数
1	5分7秒	7分3秒	1,590	329
2	7分11秒	9分26秒	1,839	190
3	13分12秒	14分32秒	911	269
4	15分10秒	17分49秒	2,201	524
5	19分45秒	21分32秒	1,261	262

度の変化を検証した。

4.1 実験データ

実験に用いる車載カメラ映像データとして、Daimler 社に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 [8] を用いる。テスト用映像は 21,790 フレームで構成されており、各フレームに歩行者ラベルが与えられている。このテスト用映像は郊外から市街地に至るまでの走行環境の変化を含んでいる。学習用画像としては、15,560 枚の歩行者画像および 6,745 枚の非歩行者画像が与えられている。なお歩行者検出器の学習の際、非歩行者画像は歩行者画像と同じサイズで与えられる必要があるため、ランダムな位置およびスケールで切り出されたものを線形補間により拡大縮小したものをを用いる。

4.2 歩行者検出器の環境適応

本実験では、提案手法のうち撮影位置の分類による走行環境の識別を人手で行った。風景が大きく変化しない範囲毎にテスト用映像を人手で分割し、それぞれを異なる走行場所で撮影された映像として扱った。郊外や市街地など、走行場所ごとに映像を人手で分割したものを表 1 に示す。

各走行環境の映像からの学習用画像の収集は、データセットに与えられている歩行者ラベルを用いて自動的に行った。はじめに、各走行環境の映像に対して歩行者を含む画像群と歩行者を含まない画像群に分割した。このうち、歩行者を含まない画像群の中からランダムな位置およびスケールの矩形で切り出した画像を学習用の非歩行者画像とした。また、歩行者を含む画像群の中からランダムに選んだ 100 枚を精度評価用のテスト画像とし、残りの中から歩行者画像を切り出し、各環境における学習用の歩行者画像とした。

このようにして収集した学習用画像を、Daimler データセットの学習用画像の中からランダムに選択した歩行者画像および非歩行者画像と併せて学習に用いることで、走行環境に適応した歩行者検出器を構築した。なお歩行者検出器には、特徴量として HOG、識別器として libsvm [9] による線形 SVM を用いた。

走行環境ごとに分割した映像の概要を表 1 に示す。異なる走行環境の映像間で重複はない。図 4 および図 5 は、走行環境 1 と走行環境 5 におけるテスト用画像の例、収集された学習用画像の例を示す。

4.3 実験条件

提案手法の有効性を確認するため、表 2 に示す各手法について、さまざまな走行環境のテスト用映像を入力し、精度評価を行った。

従来手法 1 は、Daimler データセット中の学習用画像のみを



図 4 走行環境 1 の画像例と収集された学習用画像の例。



図 5 走行環境 5 の画像例と収集された学習用画像の例。

用いて学習した歩行者検出器である。提案手法は、従来手法 1 における学習用画像に加え、それぞれの走行環境から収集した学習用画像を追加して走行環境に適応した歩行者検出器を構築した。従来手法 2 は、提案手法における学習用画像の総数と同数の Daimler データセットの学習用画像を用いて学習した歩行者検出器である。

5. 実験結果および考察

5.1 環境適応による検出精度変化に関する調査

走行環境への適応による検出精度の変化を調べるため、それぞれの走行環境に適応した歩行者検出器にその走行環境のテスト用画像を入力した場合の評価を行った。図 6 では、従来手法 1、従来手法 2、提案手法における検出精度を示している。縦軸は Recall (再現率) を表し、横軸は画像 1 枚当りの False Positives 数である FPPF を表す。この結果より、走行環境適応を行うことで歩行者検出精度が大きく向上することがわかる。FPPF = 0.30 における Recall を見ると、環境適応を行わない従来手法 1 は 0.44、従来手法 2 では 0.49 であったのに対し、提案手法では 0.57 に向上している。従来手法 2 のように事前に準備した学習用画像の枚数を増やすことで多少の精度向上が見られるが、提案手法により各環境の学習用画像を同数収集したほうが効果的であることがわかる。

また、従来手法 1 での検出結果の例を図 7 に、提案手法での検出結果の例を図 8 に示す。

5.2 学習用画像の収集枚数と検出精度の関係の調査

各走行環境における学習用画像の収集枚数と検出精度の関係を調査した。

提案手法は 50 枚の歩行者画像および 500 枚の非歩行者画像を学習用画像として映像から収集している。この枚数を増やし、100 枚の歩行者画像および 1,000 枚の非歩行者画像を収集した場合の検出精度を評価した。この結果を、図 9 に示す。なお、図 9 は走行環境 3 へ適応した歩行者検出器の結果のみを示している。この結果より、学習用画像の収集枚数を増加させることで、検出精度が高くなることがわかる。これは、収集枚数を増やすことで各走行環境により特化した識別器が構築できるため

表 2 精度比較に用いる歩行者検出器の構築手法一覧

手法	学習用画像収集数		事前に用意した学習用画像数		学習用画像の総数	
	歩行者	非歩行者	歩行者	非歩行者	歩行者	非歩行者
従来手法 1	0	0	100	500	100	500
従来手法 2	0	0	150	1,000	150	1,000
提案手法	50	500	100	500	150	1,000

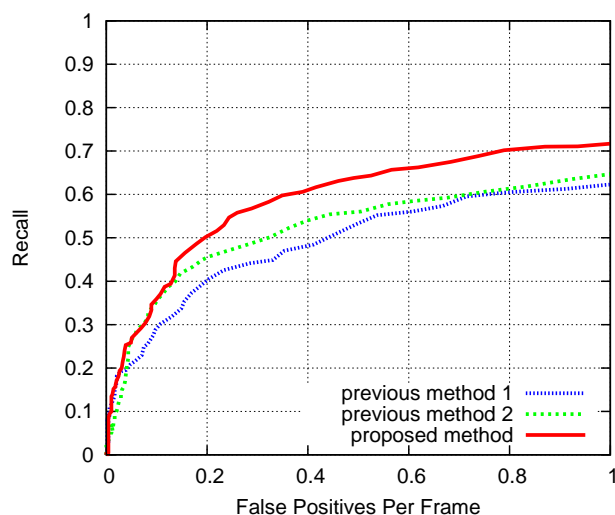


図 6 走行環境への適応による精度の変化。

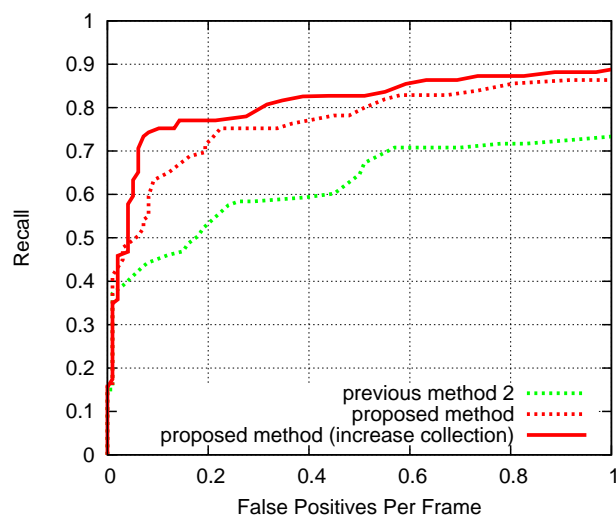


図 9 提案手法において学習用画像の収集枚数を増やした歩行者検出器の精度の変化。



走行環境 1 の例 走行環境 5 の例

図 7 従来手法 1 の検出結果。



走行環境 1 の例 走行環境 5 の例

図 8 提案手法の検出結果例。

だと考えられる。

特に、FPPF=0.09における Recall は、収集枚数が少ない場合に比べて 0.11、収集を行わない場合と比べると 0.30 向上しており、顕著な精度向上が確認できる。

5.3 走行環境への適応性の調査

歩行者検出を実際に行う走行環境に適応することが最も効果的であることを確認するため、ある走行環境のテスト用映像を異なる走行環境に適応した歩行者検出器に入力した場合の精度評価を行った。その結果を図 10 に示す。ここで、走行環境 3 のテスト用映像に対して、その他の走行環境に適応した歩行者

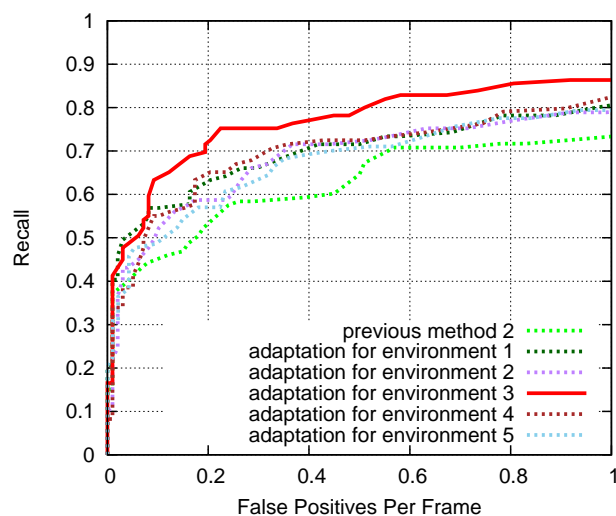


図 10 走行環境 3 のテスト映像に対し他の走行環境に適応した歩行者検出器を適用した場合の精度の変化。

検出器による検出精度も評価している。

FPPF=0.30 における Recall は、環境適応を行わない従来手法では 0.65、提案手法により走行環境 3 とは異なる走行環境へ適応した検出器では平均 0.66 であったが、走行環境 3 へ適応した検出器では 0.75 であった。このことから、実際に使用する走行環境に適応した歩行者検出器が最も高い精度を示すことを確認した。

6. む す び

本報告では、車載カメラ映像を用いた歩行者検出器の走行環境適応による精度向上について調査検討を行った。映像から歩行者画像および非歩行者画像を収集して歩行者検出器を学習することで、歩行者検出精度の向上が可能であることを実証した。

今後の課題として、歩行者検出の精度に対して影響を与える環境要因を調べるため、走行場所以外の要因に着目した環境適応についての検証などが挙げられる。例として、時計情報を用いることによる時刻の変化への適応、画像からの風景識別手法 [6] [7] などを用いることによる風景の変化への適応などが考えられる。また、GPS 等を用いた走行環境適応型歩行者検出器の実現手法なども今後の検討課題である。

謝辞 日頃より御指導および御討論をいただく名古屋大学村瀬研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。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] 警察庁交通局, “平成 22 年中の交通死亡事故の特徴及び道路交通法違反取締り状況について,” January 2011.
- [2] 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.
- [3] C. Wöhler, “Autonomous in situ training of classification modules in real-time vision systems and its application to pedestrian recognition,” Pattern Recognition Letters, vol. 23, no. 11, pp.1263–1270, September 2002.
- [4] M. Wang, and X. Wang, “Automatic adaptation of a generic pedestrian detector to a specific traffic scene,” Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.3401–3408, June 2011.
- [5] 服部 博憲, “固定カメラ映像を対象とした HOG 人物検出器のシーン適応手法,” 京都大学大学院情報学研究科知能情報学専攻修士論文, February 2010.
- [6] 藤田 隆二郎, “車載カメラ映像による風景特徴解析技術,” Forum on Information Technology 2005, 一般講演論文集, pp.137–138, August 2005.
- [7] Ryujiro Fujita, “A roadscape analysis system of vehicle mounted camera-image,” International Conference on Control Engineering 2006, pp.425–426, January 2006.
- [8] Daimler AG, “Daimler pedestrian detection benchmark dataset,” http://www.gavrila.net/Research/Pedestrian_Detection/Daimler_Pedestrian_Benchmarks/Daimler_Pedestrian_Detection_B/daimler_pedestrian_detection_b.html, October 2008.
- [9] C.C. Chang, and C.J. Lin, “LibSVM A library for support-vector machines,” <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.