一般社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
IEICE Technical Report
PRMU2013-187(2014-3)

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

鈴尾　大地†　出口　大輔††　井手　一郎†　村瀬　洋†
石田　皓之†††　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　〒 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@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技術が注目されている．一般に，車載カメラ画像に含まれる走行環境は多様に変化するため，特定の走行環境を対象とするような従来の研究では高精度に歩行者を検出することが困難であった．そこで本研究では，走行環境に適応することにより高精度な歩行者検出を行う手法を提案する．これは，多様に変化する走行環境に対して複数の識別器を構築し，これを切り替えながら検出に利用する手法である．そのために，車載カメラ画像系列を画像特徴のクラスタリングにより複数の走行環境に分類し，各環境を走行場所および進行方向と関連付ける．そして，各走行環境に応じた識別器を構築し，検出に利用する．提案手法の有効性を確認するため，実画像に対して環境適応型の歩行者検出を適用し，精度評価を行った．その結果，提案手法の検出精度は環境適応を行わない手法と比べて向上し，その有効性を確認した．
キーワード　歩行者検出，適応的検出器，ITS

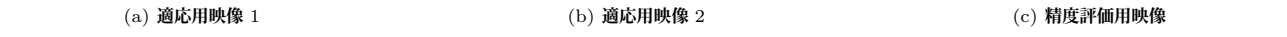
Accuracy improvement of pedestrian detection from in-vehicle camera images based on environment adaptation using location information

Daichi SUZUO†, Daisuke DEGUCHI††, Ichiro IDE†, Hiroshi MURASE†,
Hiroyuki ISHIDA†††, and Yoshiko KOJIMA†††

† Graduate School of Information Science, Nagoya University
†† Information and Communications Headquarters, Nagoya University
Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464–8601 Japan
††† Toyota Central Research and Development Laboratories, Inc.
41–1 Yokomichi, Nagakute-shi, Aichi, 480–1192 Japan
E-mail: †suzuod@murase.m.is.nagoya-u.ac.jp, †,††{ddeguchi,ide,murase}@is.nagoya-u.ac.jp

Abstract　Recently, pedestrian detection techniques from in-vehicle camera images is being focused. However, since the environment observed in the in-vehicle camera images varies largely, it is difficult to detect pedestrians accurately by previous works is trained for a specific environment. Therefore this paper proposes an accurate pedestian detection method that can adapt to various environments. The concept consists of construction of environment-specific classifiers and detection of pedestrians using an environment-specific classifier. To achieve this, in-vehicle camera images are categorized into multiple environments that are classified by clustering of image features, where each environment is associated with locations and directions. Then, environment-specific classifiers are constructed and pedestrians are detected by selecting an optimal classifier. To evaluate the effectiveness of the proposed method, an experiment was conducted by applying the proposed method to actual in-vehicle camera images. As a result, the accuracy of the proposed method was remarkably improved compared with the non-adaptive methods, and the effectiveness of the proposed method was confirmed.
Key words　Pedestrian detection, adaptive detector, ITS

This article is a technical report without peer review, and its polished and/or extended version may be published elsewhere.
Copyright ©2014 by IEICE



(a) 適応用映像 1　(b) 適応用映像 2　(c) 精度評価用映像

図 6　実験に用いる映像データセットにおける画像の例．同一地点で撮影された画像を示す．

づく背景領域推定により背景サンプルを収集する．同一地点を異なる時刻に撮影した画像の画像間差分を求めると，動的な物体である歩行者や車両などを除外した背景領域を求めることができる．これを実現するために，提案手法では適応処理への入力映像に加え，同一経路を異なる時刻に走行して撮影した参照映像を用いる．具体的には，あるフレーム $I(\boldsymbol{e}_i)$ と参照映像中の同一地点において撮影された画像 $I(\boldsymbol{e}_j)$ の画像間差分によって求める．$I(\boldsymbol{e}_i)$ と $I(\boldsymbol{e}_i)$ の対応付けは，式 (9) により行う．そして，推定された背景領域からランダムに画像を切り出すことで，背景サンプルを収集する．

一般に，画像間差分を求めるには双方の画像が厳密に位置合わせされている必要がある．しかしこれは困難なため，位置ずれに対して影響を受けにくい画像間の変化領域推定手法 [10] を用いる．これは，局所特徴量に基づいて画像間の対応を求め，対応が存在しない領域に差分があると推定する手法である．この処理の結果の例を図 5 に示す．

3.1.3　識別器および識別器パラメータマップの構築

それぞれの走行環境 $\mathcal{E}_i$ に対して収集した背景サンプルを用いて，走行環境 $\mathcal{E}_i$ 用の識別器 h_i を構築する．なお，識別器の構築に必要な歩行者サンプルは，背景サンプルと比較すると環境の変化の影響を受けにくいことから，事前に人手で用意したものを用いる．これをベースライン学習サンプルと呼ぶ．また，提案手法では識別器の構築に Deformable Part Model [6] [13] を用いる．

ここまでの処理により，任意の走行環境パラメータ $\boldsymbol{e}_x$ が与えられると，式 (5) により走行環境 $\mathcal{E}_i \in \mathbb{E}$ を選択することができる．そして，$\mathcal{E}_i$ に対応する識別器 h_i が得られる．式 (4), (5), (6) の関係を表した式 (7) を識別器パラメータマップと呼ぶ．ここで得られた識別器パラメータマップが適応処理の最終出力であり，検出処理ではこれを用いて歩行者検出を行う．

3.2　検 出 処 理

検出処理は，入力された車載カメラ画像に対して最適な識別器を用いて歩行者検出を行う処理である．車載カメラ画像に付与された走行環境パラメータ $\boldsymbol{e}_x$ をもとに，式 (7) で表される識別器パラメータマップを参照し，最適な識別器 h_i を選択する．提案手法では，式 (5) を求めるためのアルゴリズムとして，k 近傍探索（k-NN）を用いる．

続いて，選択された識別器 h_i を用いて，入力画像 $I(\boldsymbol{e}_x)$ に対する歩行者検出を行う．

4.　評 価 実 験

本節では, 提案手法における走行環境への適応とそれに基づく歩行者検出手法の検出精度評価実験について述べる．

4.1　データセット

データセットとして，各フレームに GPS 情報が付与された映像データを用いた．解像度は $1,920 \times 1,080$ 画素であり，24P モード（23.976fps，プログレッシブモード）で撮影を行った．また，焦点距離は 30.4mm 相当（35mm 版換算）とした．

映像の撮影は，同一経路を異なる日時に 3 回走行して行い，2 本を適応用映像として，残る 1 本を評価用映像として検出処理に用いた．図 6 に各映像における同一地点で撮影された画像の例を示す．本実験では映像を左右に分割し，図 4 に示すような左向きおよび右向きカメラから撮影された画像系列を生成した．

また，これらとは異なる地域で撮影された画像をベースライン学習サンプルとして用いる．これは，歩行者サンプル $3,250$ 枚，非歩行者画像 5 枚（ここから背景サンプルをランダムに多数切り出す）を含む．

4.2　比 較 手 法

提案手法の有効性を確認するため，以下の 2 手法と比較した．なお，いずれの手法においても識別器の構築には Deformable Part Model [13] を用いた．

通常の学習型検出器: ベースライン手法

環境適応を行わない，通常の学習型歩行者検出である．ベースライン学習サンプルのみを利用して構築された識別器を用いる．

進行方向を利用しない手法: 環境適応（GPS のみ）

進行方向の変化による環境の変化を考慮せず，自車位置のみを用いた走行環境適応 [5] である．

4.3　実 験 結 果

図 7 に検出結果の一例を示す．図から，提案手法によって高精度な検出が実現されることを確認した．特に提案手法では，過検出（False Positives）を低減する効果があることを確認した．最大 F 値によって検出精度を評価したところ，提案手法では 0.480 であったのに対し，環境適応（GPS のみ）では 0.431，ベースライン手法では 0.393 であり，提案手法が最も高い値を示した．図 8 に示した精度評価では，同等の FPPF においても比較手法に対して検出率の向上が得られていることがわかる．特に FPPF = 1.0 においては，提案手法の Detection Rate はベースライン手法と比較して 18.2% 向上している．これは，提

図 7　各手法における検出結果の例.

案手法における過検出を抑制する効果により，同等の過検出率（FPPF）において，検出率が大きく改善したことを表す.

4.4　走行環境数 k と検出精度の関係の評価

走行環境数 k を変えた場合の検出精度の評価結果を図 9 に示す．ここで，$k = 10$ の例は図 8 における提案手法に等しい.

この結果から，大きな k において検出精度が向上したことを確認できる．これは，各走行環境に含まれる画像の見えの変化の範囲が小さくなり，高精度な識別器の構築が可能になったためだと考えられる．しかし，k を大きくすることによる精度向上には制限があると考えられる．適切な k の値は適応用映像に含まれる走行環境の種類に依存し，一般に未知である．しかし，提案手法では k を定数のパラメータとして与える必要がある．従って，k を推定する手法や k に依存しないクラスタリング手法などについて検討する必要がある.

4.5　走行環境毎の検出精度の評価

本節では，クラスタリングされた各走行環境における検出精度を評価した結果について考察する．図 10 に，主要な 3 種類の走行環境に対する精度評価結果を示す．具体的には，提案手法の各環境 $\mathcal{E}_i$ に対応する識別器 h_i とベースライン識別器との精度比較の結果である．この結果から，環境毎に専用の識別器を構築することで検出精度の改善が可能であることがわかる.

4.6　未検出の傾向と対策

提案手法においては，背景サンプルの自動収集により過検出を抑制する効果が確認できた．一方，提案手法では抑制できない過検出として，いくつかの傾向が見られた.

図 11(a) は，自動車上に現れる過検出の例である．これは，提案手法における背景サンプル収集手法が画像間差分に基づくことから，自動車などを含む背景サンプルを収集できなかったためと考えられる．また図 11(b) は，歩行者周辺に現れる多重検出の例である．これは，背景サンプルを収集する際に，歩行者の周辺であるが歩行者を含まないものを十分収集できなかったためと考えられる．これらの問題に対処するため，背景サンプル自動収集の手法の改善が必要である.

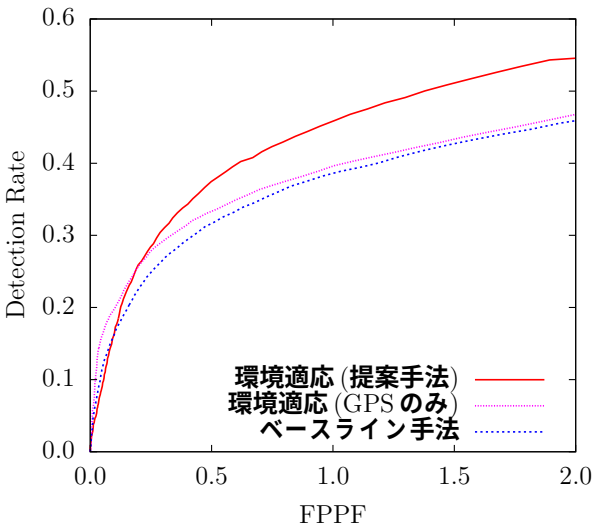


図 8　各手法における精度評価の結果.

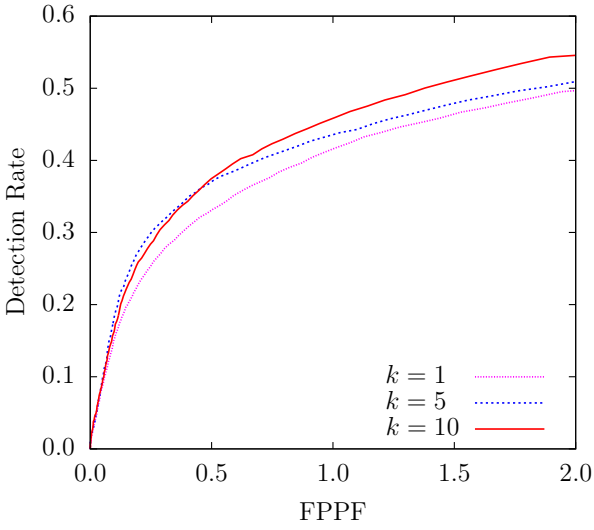
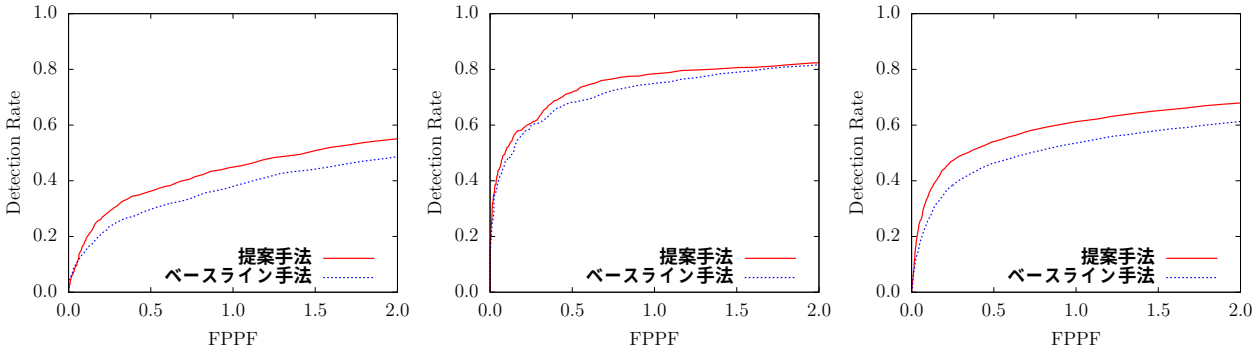


図 9　走行環境数 k と検出精度の関係.



(a) 走行環境 0 の検出精度　(b) 走行環境 2 の検出精度　(c) 走行環境 4 の検出精度

図 10　各走行環境における検出精度評価結果の抜粋.

(a) 自動車上の過検出

(b) 歩行者の多重検出

図 11　提案手法により低減できなかった過検出の例.

5. ま と め

本報告では，車載カメラ画像を用いた歩行者検出器を環境へ適応することで識別器を高精度化する手法を提案した．具体的には，車載カメラにより撮影される画像の見えを走行環境の集合と捉え，それぞれの走行環境に対して構築した識別器を切り替えながら歩行者検出を行う手法を提案した．また，走行場所および進行方向に着目して環境適応を行う手法を提案した．実験の結果，環境適応を行わない手法と比較して検出精度が大幅に向上することを確認した．また，進行方向を考慮しない環境適応と比較した場合でも検出精度の向上が得られることを確認した．今後の課題として，背景サンプル自動収集手法の更なる検討，走行環境クラスタリング手法の改善などが挙げられる．

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．

文　献

[1] 警視庁交通局，“交通事故統計（平成 25 年 11 月末）,” Nov. 2013.

[2] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” Proc. of the 2005 IEEE Computer Society Conf. 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, Sep. 2002.

[4] X. Wang, M. Wang, and W. Li, “Scene-specific pedestrian detection for static video surveillance,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 36, no. 2, pp. 361–374, Jan. 2014.

[5] 鈴尾大地，出口大輔，井手一郎，村瀬洋，石田皓之，小島祥子，“車載カメラと GPS を組み合わせた走行環境適応型歩行者検出器に関する検討,” 電子情報通信学会技術研究報告，PRMU2012-190，Mar. 2013.

[6] P. Felzenszwalb, D. McAllester, and D. Ramanan, “A discriminatively trained, multiscale, deformable part model,” Proc. of the 2008 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp. 1–8, June 2008.

[7] 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，“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,” 電子情報通信学会技術研究報告，PRMU2012-41，Sep. 2012.

[8] H. Yoshida, D. Suzuo, D. Deguchi, I. Ide, H. Murase, T. Machida, and Y. Kojima, “Pedestrian detection by scene dependent classifiers with generative learning,” Proc. of the 2013 IEEE Intelligent Vehicles Symposium, pp. 654–659, June 2013.

[9] National Geospatial-Intelligence Agency, “Department of Defense world geodetic system 1984, its definition and relationships with local geodetic systems third edition,” July 1997. http://earth-info.nga.mil/GandG/publications/tr8350.2/tr8350_2.html.

[10] R. Mitsumori, H. Uchiyama, H. Saito, M. Servières, and G. Moreau, “Change detection based on SURF and color edge matching,” in Proc. of the Workshop on Vision and Control for Access Space in the ACCV, 12p., Sep. 2009.

[11] H. Bay, T. Tuytelaars, and L. Gool, “SURF: Speeded up robust features,” Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346–359, Sep. 2008.

[12] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, “Visual categorization with bags of keypoints,” Proc. of the Workshop on Statistical Learning in Computer Vision in the ECCV, pp. 1–22, May 2004.

[13] C. Dubout and F. Fleuret, “Deformable part models with individual part scaling,” Proc. of the 24th British Machine Vision Conf., pp. 1–10, Sep. 2013.