

誤検出マイニングに基づくシーン適応型歩行者検出*

鈴木悠暉** 出口大輔*** 川西康友** 井手一郎** 村瀬 洋**

Pedestrian Detection by Scene Adaptation Based on False Positive Mining

Yuki SUZUKI, Daisuke DEGUCHI, Yasutomo KAWANISHI, Ichiro IDE and Hiroshi MURASE

In recent years, demands for Advanced Driving Assistance Systems (ADAS) is increasing, and pedestrian detection has become one of the most important and popular technologies in this system. In the case of pedestrian detection using an in-vehicle camera, since the road environment varies widely according to difference in lightning, weather, etc., it is very difficult to handle them with a single classifier, and numerous false positives are detected. To overcome this problem, this paper proposes a novel pedestrian detection method by scene adaptation based on false positive mining. When we observe the appearance of false positives in in-vehicle camera images, those with similar features are found even in different road environments. The proposed method focuses on the appearance of the detected false positives, and considers it as a scene that the classifier is not good at. By analyzing such a false positive tendency in each scene, the proposed method associates the false positive tendency to each scene and then associates them to each training image. Then, classifiers are constructed so that they can cope with false positives observed in each scene. To evaluate the effectiveness of the proposed method, experiments were conducted on the Caltech Pedestrian Detection Benchmark datasets. Its results showed that the proposed method outperforms the method without adaptation.

Key words: false positive mining, scene adaptation, pedestrian detection, in-vehicle camera

1. 緒 言

近年, Advanced Driving Assistance Systems (ADAS)¹⁾に関する研究開発が盛んに行われている。車両と歩行者の衝突を防ぐための技術の1つとして歩行者検出への注目が非常に高まっており, 多くの研究グループと自動車業界がこの問題の解決に取り組んでいる。車載カメラは安価でかつ高い解像度の画像が得られることから, 車両の周囲環境の情報を得るために重要なセンサとなっている。その中でも, 車載カメラを用いた歩行者検出への期待は高い。

車載カメラ画像からの歩行者検出にあたっては, 機械学習に

基づく検出アプローチが広く利用されている。例えば, DalalらはHistogram of Oriented Gradients (HOG)特徴量とSVM識別器の組み合わせによる歩行者検出を提案している²⁾。更にDalalらの手法を発展させ, Felzenszwalbらは歩行者のさまざまな見え(姿勢の変形など)に対処するため, Latent SVMを用いたパーツベースの歩行者検出手法を提案している³⁾。またDollarらは, HOG特徴量とLUV色空間の情報を組み合わせて利用するAggregated Channel Feature (ACF)特徴とAdaBoostを組み合わせた歩行者検出手法を提案している⁴⁾。近年では, 深層学習(Deep Learning)に基づく歩行者検出手法も多数報告されている⁵⁾。このような機械学習に基づく歩行者検出手法の性能改善を目的として, データマイニングの技術と組み合わせる試みもなされている。Malisiewiczらは物体の姿勢に注目し, アスペクト比別に構築した複数の識別器を組み合わせることで歩行者検出を行うことで性能の改善を図っている⁶⁾。

一方, 機械学習に基づく歩行者検出では, 多様な見えに対処するために数多くの学習データが必要である。しかし, 一般に学習データの増加にともなって検出対象とその背景の組み合わせのバリエーションが増加し, それに比例して誤検出が増加することが知られている。車載カメラを用いた歩行者検出の場合, 場所や輝度の変化, 天候の変化などによって歩行者とその背景の見えが大きく変動するため, 単一の識別器のみで多様な走行環境すべてに対応することは非常に難しい。

図1は, Dollarらが提案したACFに基づく歩行者検出器を実際の車載カメラ画像に適用した際に得られた誤検出の例を示している。このように, 木, 標識, 車の窓やランプといった様々な物事において誤検出が発生することが分かる。

このような誤検出を削減する研究の1つとして, Suzuoらは画像全体の見えに基づいて走行環境を分類し, 走行環境毎に検出器を構築することで検出対象と背景の組み合わせのバリエー

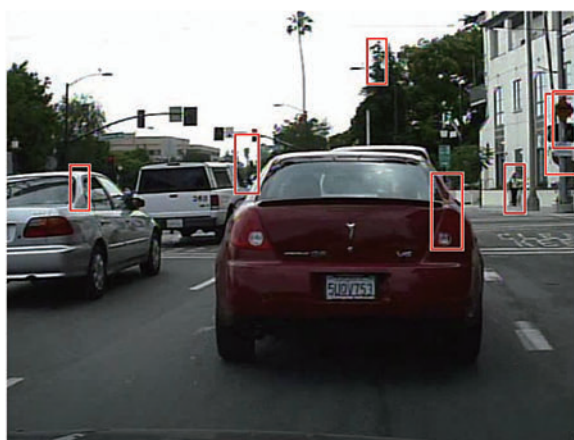


Fig. 1 Example of false positives detected by a pedestrian detector

* 原稿受付 平成 28 年 5 月 9 日

掲載決定 平成 28 年 9 月 6 日

** 名古屋大学大学院情報科学研究科 (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 (愛知県名古屋市千種区不老町)

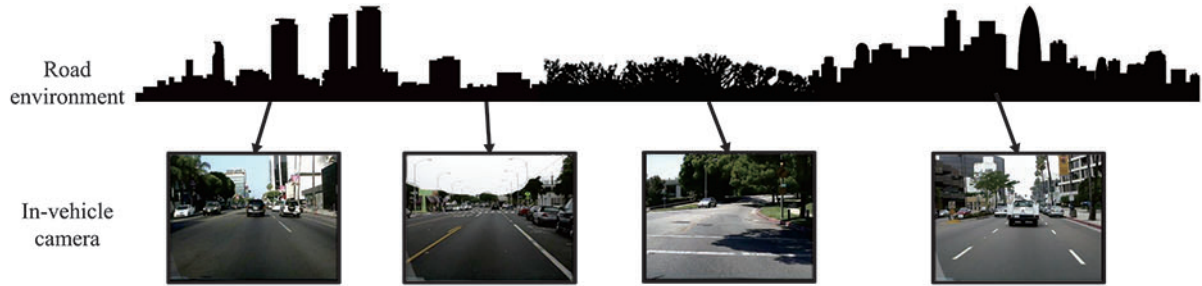


Fig. 2 Illustration of relationships between driving environments and their in-vehicle camera images



Fig. 3 Illustration of similar false positives obtained from in-vehicle camera images having different appearances

ジョンに対応する技術を提案している⁷⁾。また Siva らは、入力画像と似た背景画像をデータベースから抽出して物体検出に用いる手法を提案している⁸⁾。これらの手法を用いることにより、走行環境が比較的単純なシーンで構成されている場合は精度良く誤検出の削減が可能であることが示されている。しかし、これらの手法は画像全体の見えに基づいて環境の分類を行うため、局所的に複数のシーンから構成されるような複雑な走行環境においては、適切に誤検出を削減することができないという問題がある。

そこで本論文では、局所的に複数のシーンから構成されるような複雑な走行環境であっても誤検出の低減が可能なシーン適応型の歩行者検出手法を提案する^{*1)}。本論文の具体的な貢献は以下の2つである。

- (1) 識別器が苦手とする誤検出の傾向を抽出し、その傾向に基づいてシーンを分類するという新しい概念を提案する。従来は画像全体の見えに基づいて走行環境を分類していたため、複雑なシーンへの対応が難しかったが、この新しい分類方法によりこの問題の改善を図る。
- (2) 誤検出の傾向毎に識別器を構築し、それらを組み合わせることで単一の識別器では対応が困難な複雑なシーンに対する歩行者検出の性能の改善を図る。

以降、2. で本論文の基本的なアイデアについて述べ、3. で提案するシーン適応の具体的な手法について述べる。そして、4. で提案手法の有効性を評価するための実験と考察について述べ、最後に5. で本論文をまとめる。

2. シーン適応型歩行者検出のアイデア

図2は、車両が走行する走行環境と得られる車載カメラ画像の関係を示したものである。図に示すように、車載カメラ画像全体の見えは市街地や郊外、森といったいくつかの走行環境に

分類することができる。それらの各環境における車載カメラ画像には、木の並びやビルなどの建物、空、交通環境に対応する道路、標識、車両といった複数の局所的な見えが含まれ、一般にそれらの組み合わせとして車載カメラ画像は観察される。ここで、それぞれの車載カメラ画像に対して発生する誤検出は一般に異なり、木や道路、ビルなどの異なる物体が誤検出として現れる。これらの誤検出の増加を低減するための解決方法の1つは、歩行者検出手法をこのような車載カメラ画像それぞれで異なる変化に適応できるようにすることである。一方、走行環境中に様々な見え（ビル群や木の集合など）が含まれる場合、車載カメラ画像を走行環境のどれか1つに分類できないため、Suzuo ら⁷⁾のように画像全体を分類に用いる手法では誤検出の低減が難しいという問題がある。そこで本論文では誤検出傾向を抽出することにより、このような画像の全体的な見えに基づく環境分類の問題の解決を図る。

ここで、車載カメラ画像から歩行者検出を行った結果を注意深く観察すると、それぞれの車載カメラ画像から特有の誤検出（木や標識、電柱など）が得られることが分かる。また、車載カメラ画像上の局所的な見えが異なっていたとしても、特徴空間上の距離が近い場合は、車載カメラ画像上で同じように誤検出として現れるという性質がある。この性質に着目すると、たとえ車載カメラ画像上での局所的な見えが異なっていたとしても、特徴空間上での距離が近い誤検出はクラスタリングによって1つにまとめることが可能であることに気付く。これは、誤りのパターンを表す誤検出傾向に対応していると考えられる。

一方、図3は異なる環境（車庫と木立）において、同じような誤検出（標識）が現れる例を示している。画像中では木が大きな面積を占めるような環境であったとしても、誤検出はそれとは無関係に発生することが分かる。しかし、これらの誤検出は特徴空間上では非常に近い位置に存在すると考えられることから、これら2枚の車載カメラ画像には同じ誤検出傾向が含まれていると考えられる。

そこで、このような誤検出傾向を画像中に含まれる一種のシーンと考え、それぞれのシーンに適応した識別器を構築する。そして、入力車載カメラ画像に含まれる誤検出傾向について分析し、それに対応する識別器を複数個組み合わせることで誤検出の低減を図る。

次節で、具体的なシーン適応型歩行者検出の処理手順について述べる。

3. 誤検出マイニングに基づくシーン適応型歩行者検出

図4に誤検出マイニングに基づくシーン適応型歩行者検出の処理手順を示す。2. で説明した手法の方針に従い、誤検出をク

^{*1)} 本論文は動的画像処理実用化ワークショップ (DIA) 2016 にて発表した内容⁹⁾ を発展させてまとめたものである。

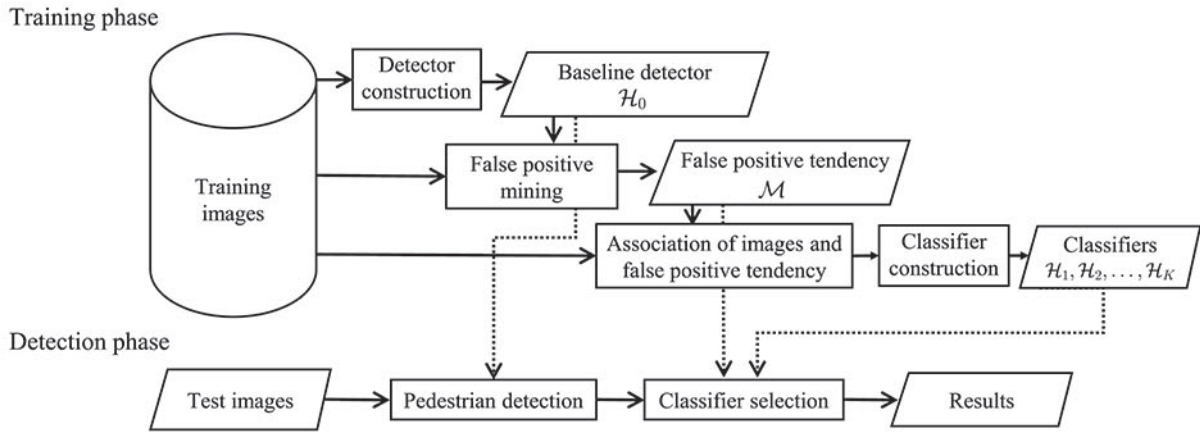


Fig. 4 Flowchart of the proposed method

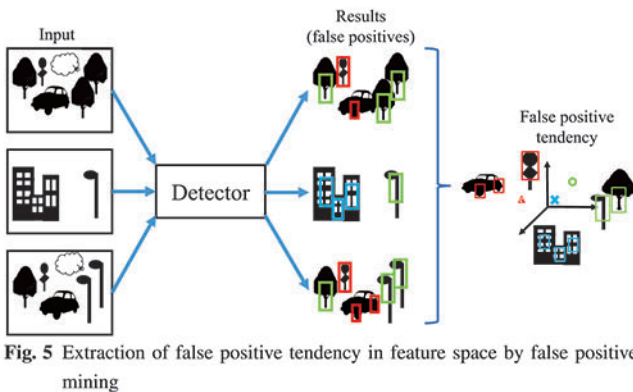


Fig. 5 Extraction of false positive tendency in feature space by false positive mining

クラスタリングすることで誤検出傾向を得る。提案手法は、学習段階と検出段階の2段階からなり、学習段階は以下の4つの処理からなる。

- (1) ベースライン検出器 \mathcal{H}_0 の構築。
- (2) 誤検出をクラスタリングすることで誤検出傾向 \mathcal{M} を獲得。
- (3) 学習画像と誤検出傾向を関連付け。
- (4) 関連付けられた学習画像を用いて誤検出傾向毎に識別器 \mathcal{H}_k を構築。

検出段階は、学習段階で構築した複数の識別器の識別スコアの重み付け和を求め、最終的な歩行者識別スコアとする。以降、各段階について詳しく説明する。

3.1 学習段階

3.1.1 ベースライン歩行者検出器の構築

まず初めに、全ての学習画像を用いてベースライン検出器 \mathcal{H}_0 を構築する。以降、このベースライン検出器を用いて誤検出傾向の抽出などを行う。提案手法は任意の検出器と組み合わせることができるが、本論文ではこのベースライン検出器ならびに次節以降で説明するシーン適応型検出器として Dollar が提案した ACF+AdaBoost⁴⁾ を用いる。

3.1.2 誤検出マイニングによる誤検出傾向の抽出

3.1.1 で構築したベースライン検出器 \mathcal{H}_0 を用いて誤検出を収集し、それらのクラスタリングを行うことで誤検出傾向を求める。図5は、誤検出傾向の抽出の概念を示した図である。

まず、誤検出を収集するために、学習画像全てに対してベースライン検出器 \mathcal{H}_0 を適用する。ここで、検出結果の中から歩行者画像と10%以上重なりがないものを誤検出とみなし、これらを収集する。

次に、全ての学習画像から得られた誤検出それぞれから特徴量を抽出し、特徴空間に射影する。ここで誤検出が K 種類の誤検出傾向に分類できると仮定し、特徴空間上で k -means クラスタリングを行うことで K 個の誤検出クラスタを取得する。そして、それぞれの誤検出クラスタの中心 \mathbf{m}_k を求め、それらの集合

$$\mathcal{M} = \{\mathbf{m}_k\}, k = 1, 2, \dots, K \quad (1)$$

を得る。本論文では、 \mathbf{m}_k が各誤検出傾向を代表すると仮定し、以降の処理で用いる。ただし、ここで用いる特徴量は高次元であるため、誤検出のクラスタリングが適切に行われない可能性がある。そこで、事前に主成分分析 (PCA: Principal Component Analysis) を適用することで次元圧縮を行う。PCAによって写像された特徴空間において、誤検出のクラスタを作成する。

3.1.3 誤検出傾向と学習画像の関連付け

学習画像と誤検出傾向を関連付ける処理の概念図を図6に示す。

まず、誤検出が含まれる各学習画像と3.1.2で求めた誤検出傾向 $\mathbf{m}_k (k = 1, 2, \dots, K)$ を関連付ける。具体的には、誤検出の特徴量 \mathbf{x} と誤検出傾向 \mathbf{m}_k の距離が最も近いもの同士を次式により対応付ける。

$$\arg \min_k \gamma(\mathbf{m}_k, \mathbf{x}) \quad (2)$$

ここで、 γ はPCAによって写像された特徴空間における入力特徴量 \mathbf{x} と誤検出傾向 \mathbf{m}_k の間の距離を定義する関数であり、本論文ではEuclidean距離を用いる。次に、各学習画像内に含まれる誤検出傾向の頻度ヒストグラムを作成し、頻度がしきい値以上の誤検出傾向を学習画像に関連付ける。ここで、最大頻度の80%をしきい値とする。これにより、1枚の画像に1つ以上の誤検出傾向が関連付けられる。

3.1.4 シーン適応型歩行者識別器の構築

3.1.3の処理により各誤検出傾向に関連付けられた学習画像を用い、誤検出傾向 $\mathbf{m}_k (k = 1, 2, \dots, K)$ 毎に識別器 \mathcal{H}_k を構築する。ここで学習に用いる正事例は、3.1.1のベースライン検出器 \mathcal{H}_0 の学習に利用した歩行者画像と同一である。

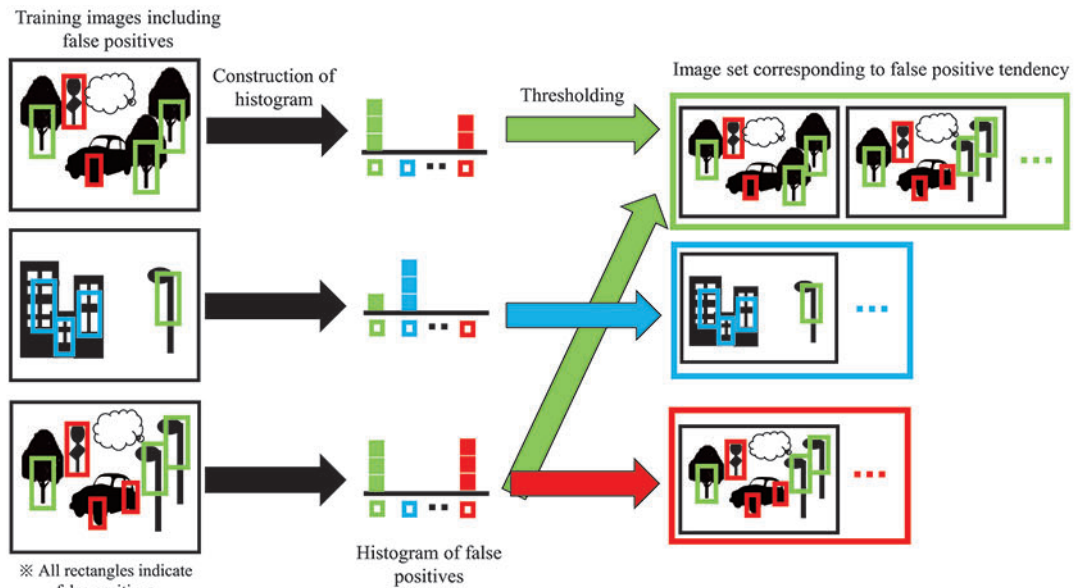


Fig. 6 Process flow for associating training images with their corresponding false positive tendencies

3.2 検出段階

提案手法では、ベースライン検出器の特徴量 \mathbf{x} と誤検出傾向に基づいて学習した複数の識別器 \mathcal{H}_k を組み合わせて歩行者検出を行う。ベースライン検出器 \mathcal{H}_0 の検出結果と誤検出傾向が一致した識別器を適用することで、誤検出の詳細な識別を行う。図 7 に、シーン適応型歩行者検出手法の概念を示す。

まず、評価用画像に対してベースライン検出器 \mathcal{H}_0 を用いて歩行者検出を行い、検出結果の特徴量 \mathbf{x} に近い誤検出傾向 \mathbf{m}_k を求める。3.1.2 の誤検出クラスタリングと同様に、誤検出傾向のクラスタ中心 \mathbf{m}_k と検出結果の特徴量 \mathbf{x} の間の Euclidean 距離を求め、この距離に近い識別器を選択する。そして、これらの識別器に対して距離に応じた重み付けをすることで最終的な識別スコアを算出する。具体的には、検出結果から抽出した特徴量 \mathbf{x} と各クラスタ中心 \mathbf{m}_k ($k = 1, 2, \dots, K$) の距離が近い方から順に P 個の識別器を選択する。選択された各識別器の重み $\omega_k(\mathbf{x})$ を、

$$\omega_k(\mathbf{x}) = \frac{\exp(-\frac{\text{dist}(\mathbf{m}_k, \mathbf{x}) - \mu}{\sigma})}{\sum_{p=1}^P \exp(-\frac{\text{dist}(\mathbf{m}_p, \mathbf{x}) - \mu}{\sigma})}, \quad (3)$$

$$\mu = \frac{1}{P} \sum_{p=1}^P \text{dist}(\mathbf{m}_p, \mathbf{x}),$$

$$\sigma = \sqrt{\frac{1}{P} \sum_{p=1}^P (\text{dist}(\mathbf{m}_p, \mathbf{x}) - \mu)^2}$$

により算出する。ここで、 μ は P の識別器それぞれに対応する誤検出傾向 \mathbf{m}_k と特徴量 \mathbf{x} の距離の平均、 σ はその標準偏差である。なお、選択されなかった $K - P$ 個の識別器の重みは 0 とする。次に、選択した識別器を用いてベースライン検出器の検出結果を評価し、それぞれの識別スコア $f_k(Y = +1|\mathbf{x})$ を算出する。ここで、 Y は検出結果のラベルを表す。このとき、選択された識別器の中で最も大きい重みと同じ値をベースライン検出器に付与し、すべての重みの正規化を行う。ベース

ライン識別器を加えたこれらの識別器毎に算出した識別スコア $f_k(Y = +1|\mathbf{x})$ と識別器の重み $\omega_k(\mathbf{x})$ の積を計算し、その総和を、

$$f(Y = +1|\mathbf{x}) = \sum_{k=0}^K \omega_k(\mathbf{x}) f_k(Y = +1|\mathbf{x}) \quad (4)$$

により求め、最終的な識別スコアとする。もし $f(Y = +1|\mathbf{x})$ がしきい値未満であった場合、検出結果を棄却する。

4. 実験および考察

本論文で提案する誤検出マイニングに基づくシーン適応型歩行者検出手法の有効性を確認するため、公開データセットを用いて性能評価を行った。以降、識別器の構築に用いるデータセット、実験条件、実験結果および考察について順に述べる。

4.1 データセット

本実験では、一般に広く用いられている Caltech Pedestrian Detection Benchmark Dataset¹⁰⁾ (以降、Caltech データセットと呼ぶ) を用いた。Caltech データセットに含まれる車載カメラ画像の大きさは 640×480 画素であり、撮影のフレームレートは 30 fps である。本実験では、このデータセットが提供する学習用データセットから 4 フレーム毎に抽出した 32,077 枚を学習用画像として用いた。学習に用いた画像の一例を図 8 に示す。ここで、データセットには歩行者領域の外接矩形が付与されており、この矩形情報を元に歩行者領域を切り出して学習用の正事例とした。また負事例については、歩行者画像と 10% 以上の重なりがない領域を無作為に切り出して用いた。

一方、評価に関してはデータセットに含まれるすべての評価用データセット (4,024 枚) を用いた。

4.2 実験条件および結果

3.1.1 で述べたとおり、提案手法におけるベースライン検出器 \mathcal{H}_0 と各シーン適応型識別器 \mathcal{H}_k ($k = 1, 2, \dots, K$) として Dollar らが提案した ACF+AdaBoost 検出器⁴⁾ を用いた。また、検出対象の歩行者の大きさは 32×64 画素以上とした。ここで、ACF では 32×64 画素を縦横 1/2 倍に縮小した後に特徴量を算出するため、得られる特徴量は 5,120 次元である。なお、

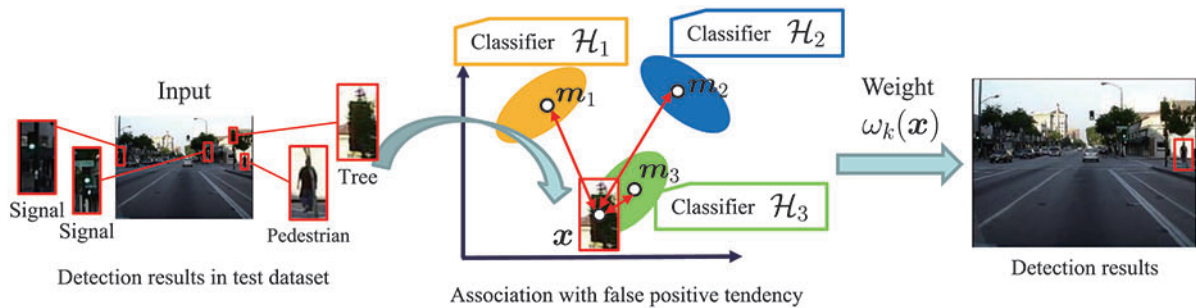


Fig. 7 Process flow of pedestrian detection by integration of scene adaptive classifiers

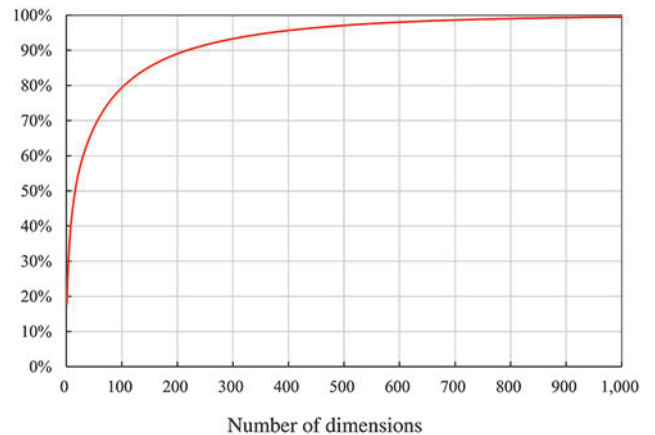
Fig. 8 Example of images in Caltech Pedestrian Detection Benchmark Dataset ¹⁰⁾

Fig. 9 Cumulative contribution of feature dimensions obtained by PCA

識別器の学習には Dollar らが提供する Matlab のプログラム ¹¹⁾ を用いた。さらに、3.1.2 におけるベースライン検出器 \mathcal{H}_0 の検出結果と誤検出傾向の対応を求める際は、PCA により特徴量を 5,120 次元から 300 次元に圧縮した。図 9 は PCA により得られた各次元の累積寄与率を示しており、本実験に用いた 300 次元での累積寄与率は約 93% であった。提案手法の有効性を確認するため、誤検出傾向に基づくシーン適応型歩行者検出に用いる学習画像クラスタの分割数を $K = 3, 5, 10$ に変化させて実験を行った。さらに、Dollar らの ACF+AdaBoost 検出器との比較も行った。これは、提案手法において $K = 1$ とした場合と等価であり、本論文で提案するシーン適応を行わない歩行者検出手法（提案手法におけるベースライン検出器 \mathcal{H}_0 と同じ）である。また、歩行者検出で選択されるシーン適応型識別器の数は $P = K$ とし、学習に用いるデータはすべての手法で同じものを用いた。

歩行者検出結果の評価指標として、Detection Error Trade-off (DET) 曲線を用いた。DET 曲線は、縦軸に Miss rate (1 - 適合率)、横軸に画像 1 枚あたりの誤検出数 (FPPI) を取るグラフである。この曲線は、左下に近づくほど誤検出が少なく、かつ検出精度が高いことを表す。

4.3 実験結果および考察

提案手法と比較手法の DET 曲線を図 10 に示す。提案手法はシーン適応を行わない比較手法と比べ、Miss rate が減少しており、誤検出傾向を用いた複数識別器の統合の有効性が確認できる。また、これらの手法を同一の Miss rate で比較した際の検出結果画像の例を図 11 に示す。

図 11 から分かるように、同じ Miss rate という条件で比較した場合、提案手法 ($K = 5$) における誤検出が少ないことが確認できる。このことは、図 10 に示した DET 曲線による精度評価からも確認することができる。特に図中の FPPI = 0.1 におい

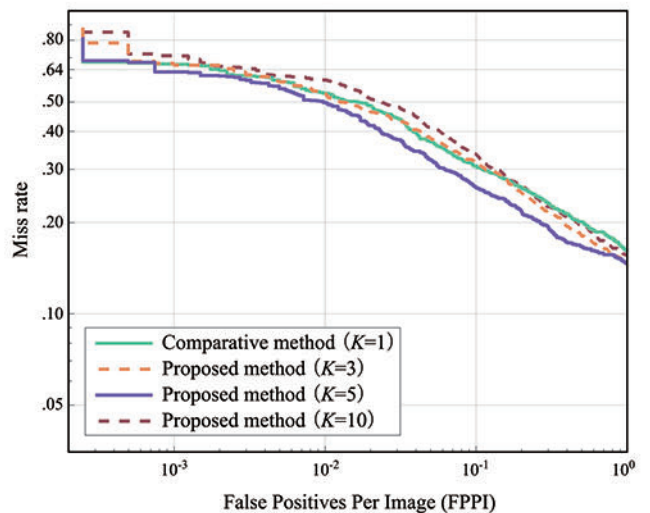
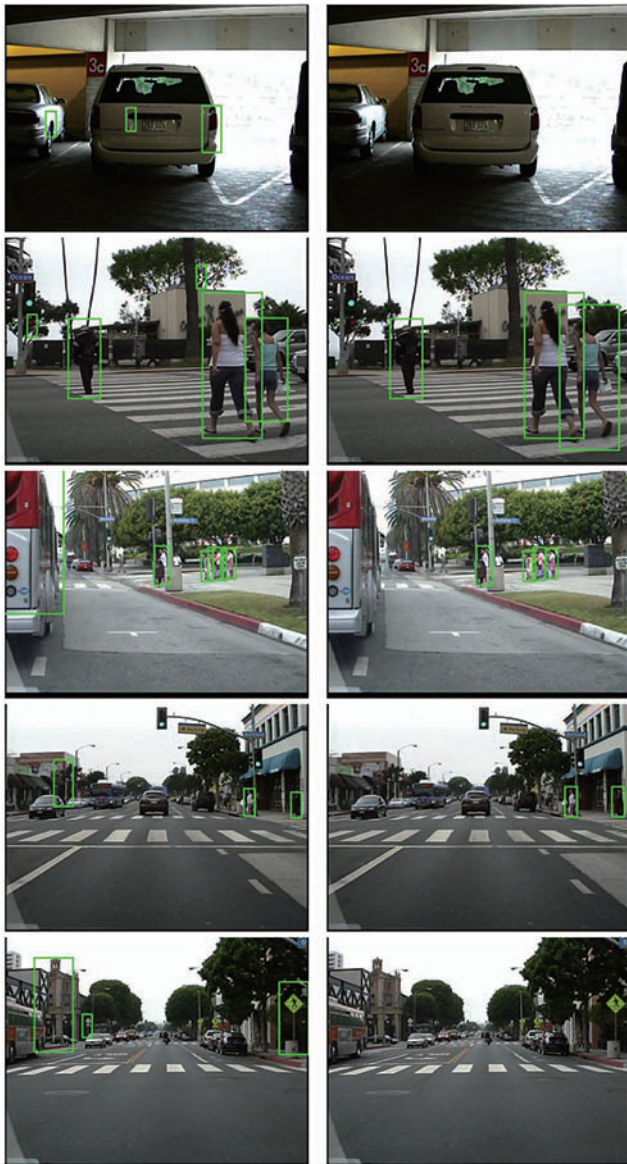


Fig. 10 Performance comparison of the proposed method and the comparative method by DET curves

て、比較手法の Miss rate と比べて 4% 低減可能であった。これは、誤検出傾向に基づくシーン適応型の歩行者検出によって誤検出の抑制が効果的に働いたことと考えられる。

図 10 は、提案手法における誤検出傾向の分割数を変化させた場合の性能も示している。 $K = 5$ とすることで比較手法よりも高い性能が得られるものの、 $K = 3$ では比較手法と比べて検出精度の向上が見られず、 $K = 10$ では比較手法よりも検出精度の低下が見られた。 $K = 10$ において検出精度が低下した主な原因を調査したところ、 $K = 5$ では各誤検出傾向と関



(a) Comparative method ($K = 1$) (b) Proposed method ($K = 5$)

Fig. 11 Results of the proposed method and the comparative method

連付けられる学習画像枚数が平均で約 1,200 枚であったのに対し、 $K = 10$ の場合は約 600 枚と約半分になっていることを確認した。このように、提案手法では K の数が多くなるにしたがって各誤検出傾向に関連付けられる画像の枚数が減少する。そのため、十分な学習用画像を用いて識別器を構築できなかったと考えられる。一方、 $K = 3$ では誤検出傾向の分割が不十分であったことが主な原因としてあげられる。これらのことから、提案手法が有効に機能するような分割数 K に関する検討が必要であり、今後は K を柔軟に決定するためのクラスタリング手法の検討も必要である。

次に、3.2 におけるベースライン検出器 (\mathcal{H}_0) と組み合わせる識別器の数 P を変化させた場合の性能を評価した。表 1 は、 $K = 5$ において P を変化させた場合の Miss rate (FPPI = 0.1) を示している。表より、 P の値によらず提案手法は比較手法よりも Miss rate を低減できていることが確認できる。このことから、組み合わせる識別器の数を変化させることで提案手法の

Table 1 Comparison of Miss rate on FPPI = 0.1 when changing P

Method	Miss rate
Comparative method ($K = 1$)	0.308
Proposed method ($K = 5, P = 1$)	0.312
Proposed method ($K = 5, P = 2$)	0.305
Proposed method ($K = 5, P = 3$)	0.269
Proposed method ($K = 5, P = 4$)	0.266
Proposed method ($K = 5, P = 5$)	0.260

性能を改善できる可能性を確認した。今後、最適な K と P の組み合わせの決定方法に関する検討が必要である。

提案手法では得られた誤検出の矩形とその周囲の特徴には何らかの関連性があると考え、同じような誤検出を含む画像群と誤検出傾向について重複を許して対応付けを行った。そして、得られた画像群を用いて各誤検出傾向の識別器を構築した。これにより、十分な数の誤検出を収集できない場合であっても一定数の学習標本を集めることを可能とした。その反面、直接誤検出傾向に関係ない標本も学習に用いられてしまい、学習時のノイズとなる可能性が存在する。今後、大規模なデータセットを用いて大量の誤検出を収集した際の性能の変化について詳細な検討が必要であろう。

実験では、PCA により特徴量を 300 次元に圧縮した上でベースライン検出器 \mathcal{H}_0 の検出結果と誤検出傾向の対応を求めている。しかし、圧縮する次元の数を変化させることでこれらの対応関係は変化する可能性がある。今後、PCA により圧縮する次元数を変化させた際の提案手法の性能について検討が必要だと考えられる。

5. 結 言

本論文では、誤検出マイニングに基づくシーン適応型の歩行者検出手法を提案した。車載カメラを用いた歩行者検出の場合、場所や日時によって走行環境の見えが大きく変動するため、検出対象である歩行者とその背景の組み合わせのバリエーションは膨大である。そのため、単一の識別器のみでは誤検出が多く現れるという問題がある。そこで本論文では、誤検出をクラスタリングすることでその傾向を求め、誤検出傾向毎に識別器を構築するシーン適応型の歩行者検出手法により誤検出の低減を図った。提案手法の有効性を評価するために、Caltech Pedestrian Detection Benchmark¹⁰⁾ を用いて性能評価を行った。その結果、提案手法は Dollar らの ACF+AdaBoost 検出器をそのまま用いる場合と比べ、検出精度を向上できることを確認した。今後は、正事例のバリエーションと誤検出傾向の関係の調査、DPM³⁾ や Deep Learning といった他の識別器と組み合わせた場合の提案手法の効果、大規模なデータセットを用いて大量の誤検出を収集した際の提案手法の性能評価、検出結果と誤検出傾向の対応関係を求める際の次元圧縮手法の検証などが挙げられる。

謝 辞

本研究の一部は、科学研究費補助金、JST 名古屋大学 COI 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。

参 考 文 献

- 1) D. Geronimo, M. A. Lopez, D. A. Sappa, and T. Graf: Survey of pedestrian detection for advanced driver assistance systems, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**, 7, (2010) 1239.
- 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*, **1**, (2005) 886.
- 3) P. Felzenszwalb, D. McAllester, and D. Ramanan: A discriminatively trained, multiscale, deformable part model, *Proceedings of 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (2008) 1.
- 4) P. Dollar, R. Appel, S. Belongie, and P. Perona: Fast feature pyramid for object detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **36**, 8, (2014) 1532.
- 5) Y. Tian, P. Luo, X. Wang, and X. Tang: Deep learning strong parts for pedestrian detection, *Proceedings of 2015 IEEE International Conference on Computer Vision*, (2015) 1904
- 6) T. Malisiewicz, A. Gupta, and A. A. Efros: Ensemble of exemplar-SVMs for object detection and beyond, *Proceedings of 2012 IEEE International Conference on Computer Vision*, (2011) 89.
- 7) D. Suzuo, D. Deguchi, I. Ide, H. Murase, H. Ishida, and Y. Kojima: Environment adaptive pedestrian detection using in-vehicle camera and GPS, *Proceedings of International Conference on Computer Vision Theory and Applications 2014*, (2014) 354.
- 8) P. Siva, R. Chris, X. Tao, and A. Lourdes: Looking beyond the image: Unsupervised learning for object saliency and detection, *Proceedings of 2013 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (2013) 3238.
- 9) 鈴木 悠暉, 出口 大輔, 川西 康友, 井手 一郎, 村瀬 洋: 誤検出マイニングに基づくシーン適応型歩行者検出の検討, 動的画像処理実用化ワークショップ (DIA2016) 講演論文集, OS2-1, (2016).
- 10) P. Dollar, C. Wojek, B. Schiele, and P. Perona: Pedestrian detection: A benchmark, *Proceedings of 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (2009) 304.
- 11) P. Dollar: Piotr's Image & Video Matlab Toolbox, <https://github.com/pdollar/toolbox> accessed 2016.1.8.