

## 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

○ 鈴木悠暉<sup>†</sup>, 出口大輔<sup>‡</sup>, 川西康友<sup>†</sup>, 井手一郎<sup>†</sup>, 村瀬洋<sup>†</sup>

○ Yuki SUZUKI<sup>†</sup>, Daisuke DEGUCHI<sup>‡</sup>, Yasutomo KAWANISHI<sup>†</sup>,  
Ichiro IDE<sup>†</sup>, Hiroshi MURASE<sup>†</sup>

<sup>†</sup>: 名古屋大学 大学院情報科学研究科, [suzukiy@murase.m.is.nagoya-u.ac.jp](mailto:suzukiy@murase.m.is.nagoya-u.ac.jp),  
{[kawanishi](mailto:kawanishi@is.nagoya-u.ac.jp), [ide](mailto:ide@is.nagoya-u.ac.jp), [murase](mailto:murase@is.nagoya-u.ac.jp)}@is.nagoya-u.ac.jp

<sup>‡</sup>: 名古屋大学 情報連携統括本部, [ddeguchi@nagoya-u.jp](mailto:ddeguchi@nagoya-u.jp)

近年、車載カメラ画像を用いた安全支援技術に関する研究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、様々な手法が提案されている。一般に、走行環境によって見えが多様に変化し、それが原因で過検出が増加する。そのため、単一の検出器で全ての走行環境の多様な見えに対応することは困難である。ここで、異なる走行環境下においても似たような誤検出が現れるという性質がある。そこでこの性質を利用することにより、誤検出の低減を目指す。具体的には、誤検出の種類毎に学習画像をクラスタリングし、各クラスタに対して検出器を構築する。これらの検出器を組み合わせることで歩行者検出を行うことにより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、従来の手法による歩行者検出手法と比べ、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> 歩行者検出, 誤検出傾向, ITS, 適応型学習

### 1. はじめに

近年、車と歩行者の衝突を防ぐために、Advanced Driving Assistance Systems (ADAS) [1] の1つである衝突回避システムが熱心に研究されている。その中でも歩行者検出は重要であり、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。また、数多くあるセンサの中において、車載カメラは安価で解像度が比較的高いことから、車の周囲環境の情報を得るための重要なセンサとなっている。

車載カメラ画像を用いた歩行者検出において、機械学習の手法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]。これまでに、Dalalらの手法を拡張することにより、性能の改善を図る手法が様々提案されている。Felzenszwalbらは、歩行者のさまざまな見え（姿勢の変形など）に対処

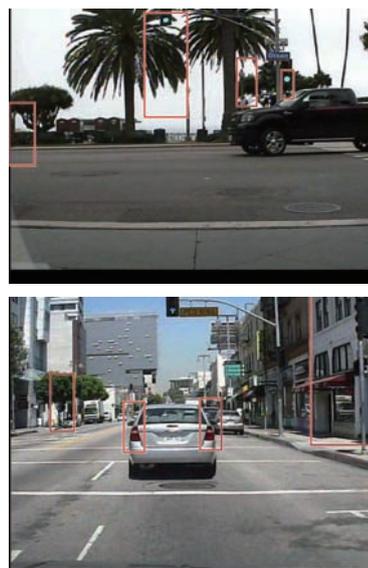


図 1 ACF+AdaBoost 検出器を実際の車の走行環境に適用した際に得られる、矩形で囲まれた誤検出例。

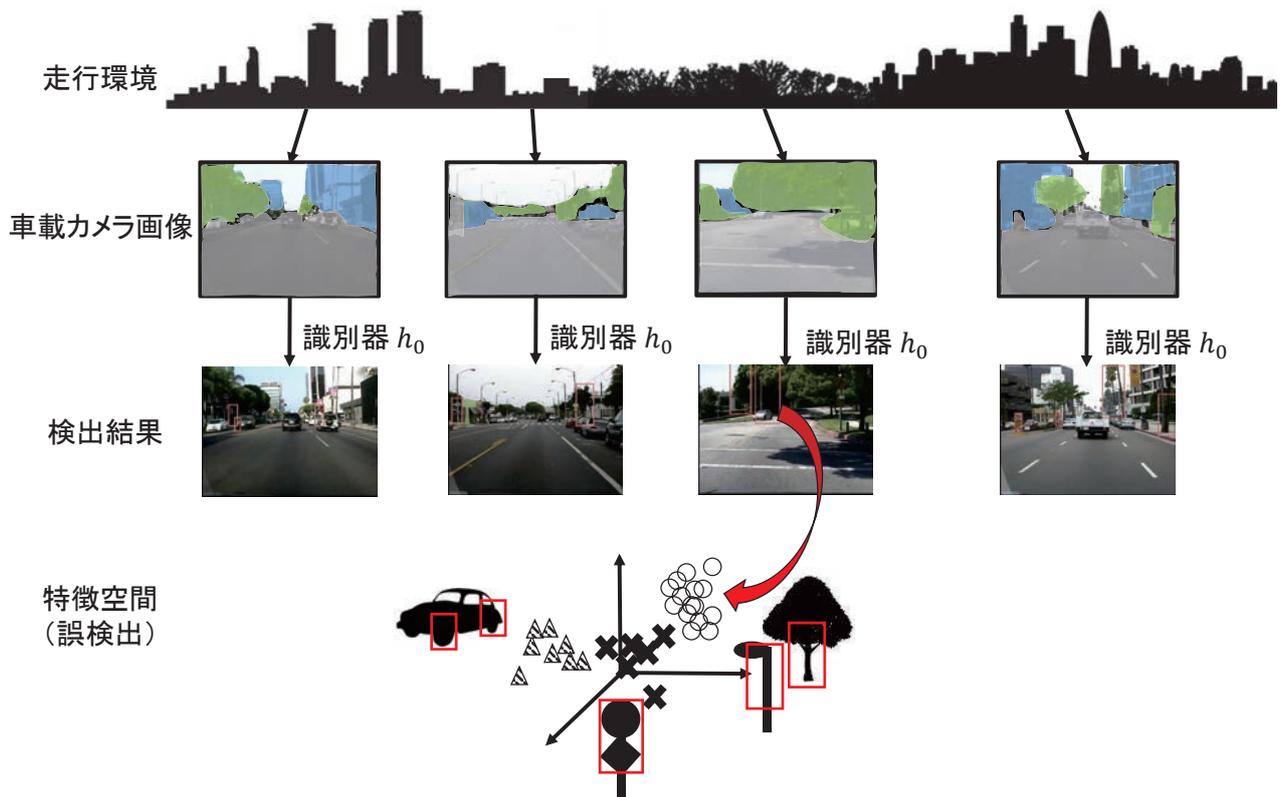


図 2 車が走行して得られる走行環境と車載カメラ画像の関係を表した概念図（本図では車載カメラ画像に含まれるシーンは人手で大まかに分類した）。各色は、緑：木が並ぶシーン、青：ビルなどの建物が並ぶシーン、灰色：交通に関する環境（道路や標識，車など）を表す。

するため、Latent SVM を用いたパーツベースのモデルによって Dalal らの手法を拡張している [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) 特徴を提案している [4]. これらの手法においては学習に用いるデータにマイニングを用いることが有効であり, Malisiewicz らはポジティブサンプルである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歩行者検出手法を提案している [5]. 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多くの学習データが必要である. しかしながら, 一般に学習データの増加にともなってクラス内のバリエーションが増加すると, 誤検出が増加して識別器の精度が低下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走行環境の見えが大きく変動する. そのため, 識別器が扱わなくてはならない走行環境は非常に多様であり,

単一の識別器のみでこの問題に対応することは困難である. 図 1 は, ACF+Adaboost 検出器を実際の車の走行環境に適用した際に得られた誤検出の例である. 図からわかるように, 誤検出として木やビルの窓, 標識, 車のランプなどの様々な大きさの矩形で囲まれた物体が得られている.

本発表では, このような誤検出を低減することで検出精度の向上を図る. 本稿の貢献は以下のとおりである.

1. 誤検出傾向に基づいてシーンを分類するという新しいコンセプトを提案する. この手法は識別器が苦手とするシーンを分類し, その傾向それぞれに対して識別器を対応させることで環境に対する頑健性を与える.
2. 誤検出傾向を用いて複数の識別器を構築する. 識別器が苦手とするシーン中の見えを網羅するために, 苦手な環境を学習した複数の識別器の検出結果を統合する. これにより多様な環境下での検出精度を改善する.

以降, 2. で関連研究について述べ, 3. で誤検出傾向に基づくシーン適応の考え方について述べる. これに基づいたシーン適応の具体的な手法について 4. で述べ, 5. で提案手法の有効性を評価するための実験について述べる. 最後に 6. で本発表をまとめる.

## 2. 関連研究

図 2 は, 車が走行して得られる走行環境と車載カメラ画像の関係の概要について示したものである. 図に示すように, 車載カメラ画像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, 交通環境 (道路や標識, 車など), 空といった複数のシーンが含まれ, 一般にそれらの組み合わせで構築されている. 走行環境から得られる画像集合  $L$  からは, 複数のシーン  $Z = \{\zeta_1, \zeta_2, \dots, \zeta_c | \zeta_c \in L\}$  が得られる.

前節で走行環境の多様性の増加によって誤検出が増加するという問題について述べた. 実際にそれぞれのシーンから得られる誤検出 (木の幹や標識, 電柱など) は異なり, 木や道路, ビルなどで誤検出が得られる. これらの誤検出が増加する問題に対する妥当な解決方法の 1 つは, 見えの変化に適応可能な歩行者検出手法を用いることである.

Suzuo らは, 走行環境の見えに着目した走行環境の分類を行い, 走行環境毎に識別器を構築している [6]. Siva らは, 入力画像に適した (入力画像と見えが似ている) いくつかの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 [7]. 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, 環境が単一のシーンのみで構成される場合に高い精度が得られることが示されている.

一方, 走行環境中に様々なシーン (ビル群や木の集合など) が含まれる場合, 走行環境をどれか 1 つのシーンに分類できないという問題がある. このため, 一般に誤検出が増加する. そこでこの問題に対処するために画像の見えに基づく環境分類をシーンの分類に拡張することを考える.

車載カメラ画像に識別器を適用すると, 各シーンから特有の誤検出 (木や標識, 電柱など) が得られる. また, 誤検出に対応する特徴量を抽出する

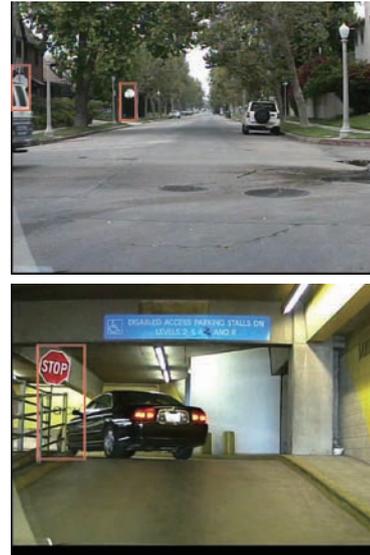


図 3 異なる走行環境において, 同一の誤検出 (標識) が現れる例.

と, 特徴空間上で似ている形状特徴が現れる. 似た特徴量が現れる誤検出をクラスタリングによりまとめることで, 誤検出の傾向に合わせてシーンをまとめることができる. 次節では, シーンに現れる誤検出の傾向に基づく識別器のシーン適応に関するアイデアを紹介する.

## 3. シーン適応型歩行者検出のアイデア

前節において走行環境の見えの変化に適応する手法を用いることにより, 誤検出の低減が可能であることを説明した. しかし, 1 枚の画像が複数のシーンから構成される場合, 画像全体をどれか 1 つのシーンに分類する手法は適切に動作しない. 図 3 は異なる環境 (車庫と木立) において, 同一の誤検出 (標識) が現れる例を示している. 木が多くみられる環境であったとしても, 木に関する誤検出以外の, 画像の大部分を占める見えとは異なる標識の誤検出が存在している. このことから, 誤検出は必ずしも画像全体の見えに依存せず, 検出に用いられる検出窓周辺のシーンによって誤検出の傾向が決まることがわかる.

誤検出の傾向を利用してシーンに適応した識別器を構築することで, 異なる走行環境下に現れる誤検出に対応できると考えられる. ここで, 1 枚の車載カメラ画像  $I_n$  に識別器  $h_0$  を適用して誤検出  $FP_n$  が得られた場合を考える. 誤検出  $FP_n$  は

学習段階

-----> 誤検出傾向の利用

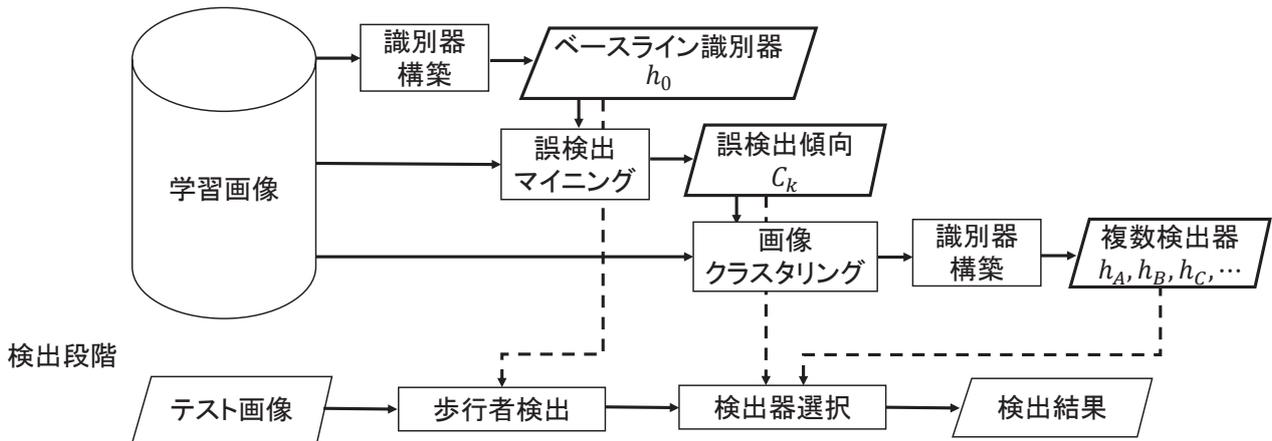


図 4 提案手法の処理手順

画像中のいずれかのシーン  $\zeta_c$  中に現れる。また、誤検出が現れるシーン  $\zeta_c$  と誤検出  $FP_n$  の関係は、誤検出が現れるシーン  $\zeta_c$  を含む画像  $I_{\zeta_c}$  から抽出した特徴量  $x_i$  を識別器  $h_0$  に入力することで得られる。これらの誤検出は、識別器  $h_0$  が苦手とするシーンを表していると考えられる。そのため、このような誤検出は識別器  $h_0$  が苦手とするシーンを理解するための手がかりとなる。この識別器  $h_0$  が誤検出を生じさせる要因には、学習画像から取り出したネガティブサンプルが影響している。このことから、誤検出が現れるシーンを含む学習サンプルをクラスタリングし、このクラスタ毎に識別器  $h_0$  が苦手なシーンに適応した新たな識別器  $h_{FP}$  を構築する。

#### 4. 誤検出マイニングに基づくシーン適応型歩行者検出

本稿では、誤検出を減らすために、誤検出マイニングに基づいて複数の識別器を構築する手法を提案する。図 4 は、提案手法の処理手順を示したものである。提案手法は、学習段階と検出段階の 2 つの段階からなる。初めに、提案手法の学習の流れについて説明する。シーン適応型検出器の構築手法は以下の 4 つの処理からなる。

- 処理 1. 従来手法によるベース歩行者識別器を構築する。
- 処理 2. 類似している誤検出をクラスタリングすることで誤検出マイニングを行う。

- 処理 3. 誤検出傾向に基づく学習画像のクラスタリングを行う。
- 処理 4. 学習画像クラスタリングにより得られるクラスタ毎に、複数の識別器を構築する。

最後に、以上の処理によって得られた複数の識別器の識別スコアを重み付けして和をとることで歩行者の検出スコアとする。以降の節において、各処理について詳しく説明する。

##### 4.1 ベース歩行者識別器の構築

この節では、画像から得られる誤検出マイニングに用いるベースとなる検出器  $h_0$  を構築する。この検出器は、全ての学習画像  $L$  を用いて学習する。識別器  $h_0$  の構築に用いる学習サンプルは、事前に歩行者領域の外接矩形が付与されている画像から、歩行者画像と非歩行者画像を切り出したものである。非歩行者画像は、すべての走行環境の背景から歩行者領域の外接矩形との重なりが 10% 以下となるようにランダムに選択する。

##### 4.2 誤検出マイニング

この節では、4.1 で構築した検出器を用いて木や標識のような誤検出を収集し、シーン中に現れる誤検出傾向を求めるために誤検出のクラスタリングを行う。

初めに学習画像から誤検出を得るために、1 枚の学習画像  $I_n$  にベース識別器  $h_0$  を適用して歩行者検出を行う。ここで、ベース識別器  $h_0$  の構築に用いたネガティブサンプルの切り出しと同様、検

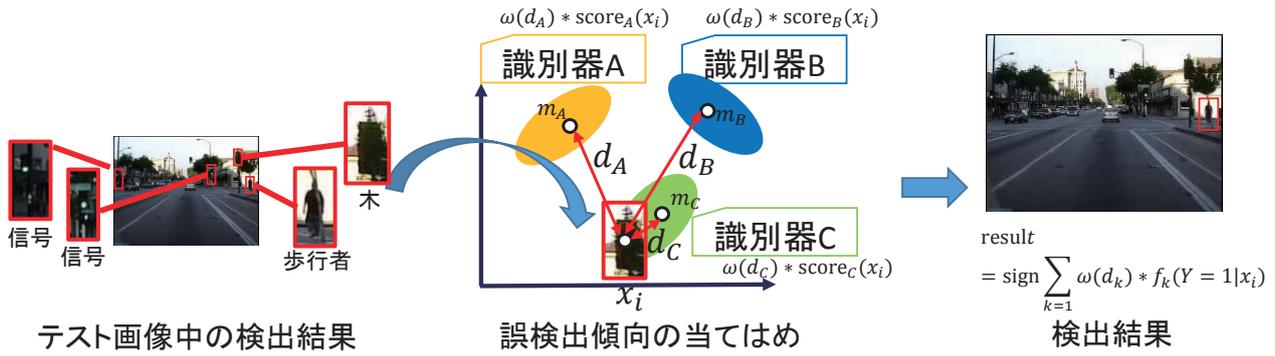


図 5 シーン適応型識別器を組み合わせた歩行者検出の概要

出結果の中から歩行者画像と 10%以上重なりがないもののみを学習画像  $I_n$  の  $i$  番目の誤検出  $\text{FP}_{n,i}$  とみなし、これらを収集する。

次に、画像集合  $L$  から収集した各誤検出  $\text{FP}_i$  からベース検出器の構築に用いた手法によって特徴量  $x_i$  (HOG 特徴量や ACF 特徴量など) を抽出する。そして、全ての特徴量を 1 つの特徴空間に射影する。ここで、誤検出が  $K$  種類に分類できると仮定し、特徴空間上で  $k$ -means クラスタリング手法を適用することで誤検出のクラスタを取得する。誤検出傾向はそれぞれの誤検出クラスタ中心  $m_k$  とその集合  $M$  によって

$$M = \{m_k\}, k = 1, 2, \dots, K \quad (1)$$

と表現される。ここで用いる特徴量  $x_i$  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ぬ可能性がある。そこで提案手法では、主成分分析 (PCA) 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  $M$  を作成する。

#### 4.3 誤検出傾向に基づく学習画像のクラスタリング

提案手法では、誤検出傾向を用いて学習画像  $L$  の中から誤検出を収集できた画像  $I_{\text{FP}_n}$  をクラスタリングすることで、誤検出傾向に基づく学習画像を用いてシーン適応型検出器  $h_{m_k}$  を構築する。

初めに、4.2 において各学習画像  $I_{\text{FP}}$  から収集した誤検出  $\text{FP}_{n,i}$  と各誤検出クラスタ  $m_k$  を結びつける。誤検出  $\text{FP}_{n,i}$  から得られる特徴量  $x_i$  対

応する誤検出傾向  $m_k$  は、

$$m_k = \arg \min_k d_{ik} \quad (2)$$

$$d_{ik} = \text{dist}(m_k, x_{ik}) \quad (3)$$

により求められ、 $\text{dist}(\cdot)$  は PCA によって写像された特徴空間における入力特徴量  $x_i$  と誤検出のクラスタ中心  $m_k$  の間の距離  $d_{ik}$  を定義する関数である。本稿では Euclid 距離を用いる。そして、この距離が最も近いクラスタ  $m_k$  に誤検出  $\text{FP}_{n,i}$  を投票する。もし投票数がしきい値以上であれば、画像  $I_{\text{FP}_n}$  としきい値以上の全ての誤検出傾向  $m_k$  を結び付ける。この投票は以下の式で表される。

$$i_{m_x} \leftarrow \frac{\text{bin}(I_{\text{FP}_n}, m_x)}{\max_{k=1, 2, \dots, K} \text{bin}(I_{\text{FP}_n}, m_k)}, \quad (4)$$

$\text{bin}(\cdot)$  は各誤検出傾向における投票数を返す関数である。ここで、最大投票数の 80% をしきい値とする。これにより、1 枚の画像が 1 つ以上の誤検出傾向と対応付けられる。

#### 4.4 シーン適応型歩行者識別器の構築

本節では、誤検出傾向  $M$  に基づいて歩行者識別器  $H_m = \{h_{m_1}, h_{m_2}, \dots, h_{m_K}\}$  を学習する手順について述べる。ネガティブサンプルには、前節でクラスタリングした学習画像  $I(m_k)$  から切り出した画像を用いる。ポジティブサンプルは 4.1 のベース識別器  $h_0$  の学習に利用した歩行者画像と同一である。各誤検出傾向  $m_k$  は誤検出のクラス数  $K$  と等しいため、誤検出傾向毎に構築される識別器  $H_m$  の数も等しくなる。シーン  $\zeta_c$  の誤検出傾向  $m_c$  といずれかの検出器の誤検出傾向  $m_k$  が一

致した場合、これらのシーン適応型歩行者検出器  $h_{m_k}$  は精度よく歩行者を検出できると期待される。

#### 4.5 歩行者検出

提案手法では、誤検出傾向に基づいて学習した複数の検出器  $h_{m_k}$  を組み合わせて歩行者検出を行う。ベース検出器  $h_0$  の検出結果において誤検出傾向  $m_k$  が一致した検出器  $h_{m_k}$  を用いて歩行者の識別を行うことで、誤検出の低減を図る。図5はシーン適応型歩行者検出手法の概要を示している。

初めに、テスト画像に対してベース検出器  $h_0$  を用いて歩行者検出を行い、検出結果から抽出した特徴量  $x_i$  を適用して誤検出傾向  $m_k$  を求める。このとき、歩行者を含めて検出結果にいずれかのシーン適応型歩行者識別器  $h_{m_k}$  を適用することになる。本稿では、各識別器は歩行者を検出できるものと仮定して、検出結果の識別を行う。次に、4.2の誤検出のクラスタリングと同様、誤検出クラスタの中心  $m_k$  と検出結果の特徴量間で Euclid 距離  $d_{ik}$  を (3) 式より求める。誤検出のクラスタ中心  $m_k$  からの距離が近い識別器  $h_{m_k}$  を用いて歩行者の識別を行う。シーン適応型識別器  $H_M$  を組み合わせた歩行者の識別では、入力特徴量  $x_i$  とクラスタ中心  $m_k$  からの距離に応じて識別器の選択と識別器に関する重みづけを行うことで歩行者を検出する。具体的には、誤検出傾向に適応した複数検出器の中から  $p$  個の検出器を選択し、検出結果から抽出した特徴量  $x_i$  とクラスタ中心  $m_k$  からの距離に応じて識別器の重み  $\omega(x_i, m_k)$  を算出する。重み  $\omega$  の算出について以下の式に示す。

$$\omega(x_i, m_k) = \frac{\exp(-\frac{d_{jk}-\mu}{\sigma})}{\sum_{p=1}^P \exp(-\frac{d_{jp}-\mu}{\sigma})}, \quad (5)$$

$$\mu = \frac{1}{p} \sum_{p=1}^P d_{ip},$$

$$\sigma = \frac{1}{p} \sum_{p=1}^P (d_{ip} - \mu)^2$$

ここで  $\mu$  は選択された  $p$  個の検出器の特徴量間距離の総和、 $\sigma$  は距離の分散である。なお、選択されなかった識別器の重みは0である。次に、選択した識別器を用いて歩行者の識別を行い、識別スコア

$$f_k(Y = 1|x_n) = f(Y = 1|x_n, m_k) \quad (6)$$

を算出する。本稿では、識別スコア  $f_k(Y = 1|x_n)$  は Adaboost における全ての決定木の各葉ノードがもつ信頼度の総和とする。この識別器毎に算出した識別スコア  $f_k(Y = 1|x_n)$  と識別器の重み

$$\omega_k(x_n) = \omega(k|x_n) \quad (7)$$

の積を計算し、重み付きスコアの総和を取ること、歩行者の最終検出スコア  $f(Y = 1|x_n)$  とする。

$$f(Y = 1|x_n) = \sum_{k=1}^K f_k(Y = 1|x_n) \omega_k(x_n) \quad (8)$$

もし最終識別スコア  $f(Y = 1|x_n)$  がしきい値以下の場合、検出結果を棄却する。このとき歩行者として認識された検出結果において、歩行者周辺に多くの検出矩形が現れることがある。これらの検出矩形に対して、Non-Maximal Suppression (NMS) [8] による検出矩形の選択を行い、統合した検出矩形を提案手法の最終的な歩行者検出結果とする。

## 5. 評価実験

提案手法の有効性を評価するために、評価実験を行った。Dollar らによる Aggregated Channel Feature (ACF) と Adaboost を組み合わせた検出手法 [4] と提案手法を比較した。以降の節では、データセットの詳細な説明と検出器を構築した際のパラメータについて述べる

### 5.1 実験データ

本稿では、評価に Caltech Ppedestrian Detection Benchmark [9] を用いた。このデータセットには、単眼カメラで撮影したカラー画像が含まれている。実験において、32,077 枚の学習画像と 4,024 枚のテスト画像を用いた。学習画像は歩行者領域の外接矩形が付与されており、全ての歩行者画像と無作為に切り出した非歩行者画像を用いた。非歩行者画像は、歩行者画像と 10%以上の重なりがない領域のみを学習に用いた。

### 5.2 識別器の構築パラメータ

実験における歩行者の最小検出サイズは  $48 \times 96$  pixel としたため、ACF 特徴量の次元数は、5,120 次元である。識別器の構築には、Dollar らが提供

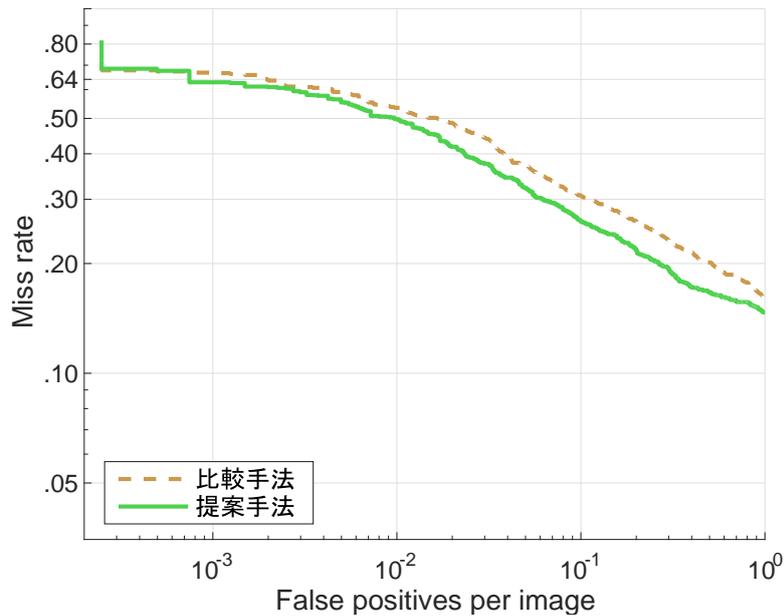


図 6 Caltech データセット [9] を用いた歩行者検出結果 (FROC 曲線)

する Matlab のプログラム [10] を用いた。誤検出傾向に基づくシーン適応型歩行者識別器に用いた学習画像クラスタの分割数は 5 とした。

### 5.3 実験結果と考察

表 1 FPPI=0.1 のときの各手法の Miss Rate

| 手法             | MissRate |
|----------------|----------|
| 比較手法           | 0.308    |
| 提案手法 ( $p=1$ ) | 0.312    |
| 提案手法 ( $p=2$ ) | 0.305    |
| 提案手法 ( $p=3$ ) | 0.269    |
| 提案手法 ( $p=4$ ) | 0.266    |
| 提案手法 ( $p=5$ ) | 0.260    |

提案手法の有効性を示すために、従来手法 [4] による歩行者検出手法と比較した。従来手法では、学習画像を全て用いて歩行者識別器の学習を行った。図 6 は、5 個の識別器を選択した場合の提案手法と従来手法の FROC 曲線による実験結果を示している。表 1 は、FPPI=0.1 のときの比較手法と  $p$  個の識別器を選択した提案手法の各 Miss Rate を示している。FROC 曲線から分かるとおり、提案手法は従来手法より Miss Rate が低く、精度が高いことがわかる。これは、提案手法により誤検出

が低減できたことを示している。特に、シーン適応型識別器を 3 つ以上選択し、識別スコアの重み付けによる歩行者識別を行うことで、FPPI が 0.1 における Miss Rate を 4 % 低減することができた。その結果、提案手法によって同一の検出率における誤検出を低減でき、検出精度の向上を確認した。

一方、選択した識別器の数が 2 つ以下であった場合、精度の向上が見られなかった。その理由として、すべての歩行者画像を用いてシーン適応型歩行者検出器を構築したため、歩行者に対するシーン適応型歩行者検出器の識別スコアが低下し、効果的な歩行者検出に失敗したと考えられる。

今後の展望として、誤検出傾向に基づいて学習に用いる歩行者画像の分類を行うことで、歩行者画像にシーン適応型識別器を適用した際の識別スコアが向上するか調査する予定である。また、他の DPM や Deep Learning に提案手法を適用したいと考えている。

## 6. むすび

本稿では、誤検出マイニングに基づくシーン適応型識別器を用いる歩行者検出手法を提案した。

走行環境が多様なために単一の識別器で対応できずに生じる誤検出に対して、誤検出の特性に基づくシーン適応を図ることで誤検出の低減を図る。

提案手法では、誤検出を収集してクラスタリングを行うことで誤検出傾向を求め、誤検出傾向ごとに分かれた学習画像クラスタごとに識別器の構築を行った。そして、構築した識別器を組み合わせて歩行者検出を行った。提案手法の有効性を評価するために、Caltech pedestrian detection benchmark[9]を用いて、評価実験を行った。その結果、提案手法は従来手法を用いた歩行者検出手法と比べ、検出精度を向上できた。今後はポジティブサンプルに対する誤検出傾向の影響の調査とDPMやDeep Learningなど他の学習手法に対して提案手法を適用する予定である。

## 謝辞

本研究の一部は、科学研究費補助金による。

## 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 Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 7, pp.1239–1258, July 2010.
- [2] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," In *Proc. 2005 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, vol. 1, pp.886–893, June 2005.
- [3] P. Felzenszwalb, D. McAllester, and D. Ramanan, "A discriminatively trained, multi-scale, deformable part model," In *Proc. 2008 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.1–8, June 2008.
- [4] P. Dollar, R. Appel, S. Belongie, and P. Perona, "Fast feature pyramid for object detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol.36, no.8, pp.1532–1545, 2014.
- [5] T. Malisiewicz, A. Gupta, and A. A. Efros, "Ensemble of exemplar-SVMs for object detection and beyond", In *Proc. 2012 IEEE Int. Conf. on Computer Vision*, pp.89–96. Nov. 2011.
- [6] D. Suzuo, D. Deguchi, I. Ide, H. Murase, H. Ishida, and Y. Kojima, "Environment adaptive pedestrian detection using in-vehicle camera and GPS," In *Proc. Int. Conf. on Computer Vision Theory and Applications 2014*, 354–361, 2014.
- [7] P. Siva, R. Chris, X. Tao, and A. Lourdes, "Looking beyond the image: Unsupervised learning for object saliency and detection," In *Proc. 2013 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.3238–3245, 2013.
- [8] J. Gil, and W. Michael, "Computing 2-D min, median, and max filters," *Proc. 1993 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.504–507, 1993.
- [9] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: A benchmark," *Proc. 2009 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pp.304–311, June 2009.
- [10] Dollar らによる Matlab のツールボックス <http://vision.ucsd.edu/~pdollar/toolbox/doc/> (2016/1/8 閲覧)