

Geometric Context を用いた特徴量間の共起による物体検出の高精度化

高木 雅成[†] 山内 悠嗣^{††} 三井 相和^{††} 藤吉 弘亘^{††}

[†] 株式会社デンソー 〒448-8661 愛知県刈谷市昭和町 1-1

^{††} 中部大学大学院工学研究科 〒487-8501 愛知県春日井市松本町 1200

E-mail: [†]MASANARI_TAKAGI@denso.co.jp, ^{††}{yuu,mitsui,hf}@vision.cs.chubu.ac.jp

あらまし 近年、人や車両などの物体検出には HOG 等のアピアランススペースの局所特徴量を用いた検出法が提案されている。しかし、アピアランススペースの特徴量だけでは、識別困難なサンプルが存在する。この問題に対して、コンテキストの利用として物体カテゴリ間の共起性に基づく手法が提案され、その有効性が確認されている。しかし、学習サンプル中では同時に存在する確率が低いカテゴリ間の物体に対して、物体間の共起性を表現するため識別精度に悪影響を及ぼす可能性がある。そこで、本稿では特徴量レベルで非対象カテゴリと対象カテゴリの特徴量間の共起性を表現することでコンテキストを用いた高精度な物体検出法を提案する。Geometric Context として得られる“空”や“地面”などの確信度を特徴量として利用し、Real AdaBoost の弱識別器の出力を演算子によって結合することにより共起性を表現する。従来の共起表現法では、和演算子と積演算子を用いた共起表現法のみであったが、本稿では検出対象とは異なるカテゴリの特徴量と対象カテゴリの特徴量間の共起性を表現するために差演算子を導入することにより、従来とは異なる共起表現を新たに加える。提案手法の有効性を確認するために車載カメラにおける人と車両の検出評価実験を行い、コンテキストを利用しない従来手法と比較して提案手法は高精度に検出可能であることを確認した。

キーワード Geometric Context, Histograms of Oriented Gradients, 物体検出, 共起確率特徴量, Boosting

Improving Object Detection Accuracy by Feature Co-occurrence Using Geometric Context

Masanari TAKAKI[†], Yuji YAMAUCHI^{††}, Tomokazu MITSUI^{††}, and Hironobu FUJIYOSHI^{††}

[†] DENSO CORPORATION, 1-1 Showa, Kariya, Aichi, 448-8661 Japan

^{††} Dept. of Computer Science, Chubu Univ. 1200, Matsumoto, Kasugai, Aichi, 487-8501 Japan

E-mail: [†]MASANARI_TAKAGI@denso.co.jp, ^{††}{yuu,mitsui,hf}@vision.cs.chubu.ac.jp

Abstract This paper presents a novel approach for improving accuracy of object detection by co-occurrence between local features using Geometric Context (GC). Our approach uses local features such as probability of geometric classes (ground plane, vertical things, sky) and output of weak classifier of Real AdaBoost to obtain the co-occurrence feature. Capturing the co-occurrence between Histograms of Oriented Gradients (HOG) and GC feature makes it possible to construct an effective object detector. By evaluating experiment of human and vehicle detection in in-vehicle camera, it shows that our method has higher performance than the conventional method of not using context information.

Key words Geometric Context, Histograms of Oriented Gradients, Object Detection, Co-occurrence Probability Feature, Boosting

1. はじめに

現在、交通事故発生件数は80万件、負傷者数は100万人を超えるなど、依然として憂慮すべき情勢である[1]。この問題に対処するため、ITSに関する研究が盛んに行われている。その中でもカメラ映像から画像認識技術により人や車両の検出を行い、運転者の視覚支援に役立て

る研究が多く取り組まれている。これらの手法は、輝度やエッジに着目したアピアランススペースの特徴量[2]~[4]であるため、検出対象の形状に似た物体を誤検出することや姿勢変動や隠れなどの影響により未検出する問題がある。この問題を解決する手法として、一般物体認識の分野ではコンテキストの一種である物体のカテゴリ間の共起を利用した認識手法[5],[6]が提案されている。岡部

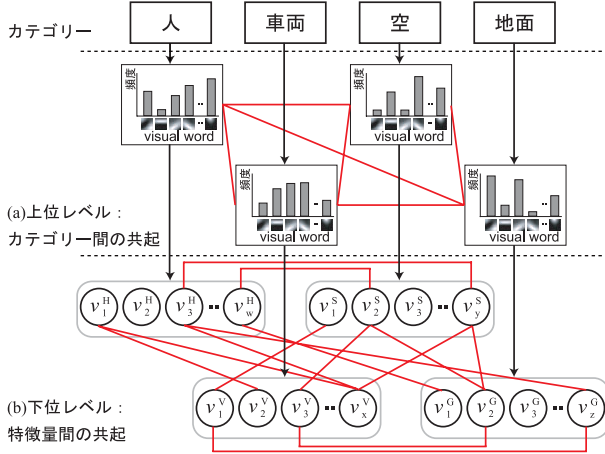


図 1 上位レベルと下位レベルのコンテキスト

らは、Bag of Feature(BoF)のアプローチにおいて、局所特徴の相対度数分布の線形結合に基づく尤度とカテゴリの共起に基づく事前確率を組み合わせた事後確率を最大化する MAP 推定により、各カテゴリの存在比率を推定し物体認識[5]を実現している。この手法は、入力画像中に“自動車”が存在した場合、隣を走っているのは“牛”の可能性は低く、“バイク”である可能性が高いことをカテゴリ間の共起として表現している。このようなカテゴリ情報の共起を表現することにより、アピランス情報のみでは識別が難しいサンプルに対しても正識別することができる。しかし、このような領域全体から表現される対象カテゴリ間 (図 1(a)) での共起は幾つかの問題がある。1 つ目は、学習サンプル中では同時に存在する確率が低いカテゴリの物体に対しても共起を表現するため、識別精度に悪影響を及ぼす可能性がある。2 つ目は、入力画像中に多くの物体が存在した場合、全ての物体のカテゴリの共起を表現すると、表現された共起関係が弱くなるため、識別に有効な情報が減少する問題である。

そこで、本稿では従来用いられている HOG 等のアピランス情報に加え、Geometric Context [7] によって得られる検出対象カテゴリと非検出対象カテゴリの確信度を特徴量として扱い、局所領域から得られる特徴量間の共起を表現することにより物体検出の高精度化を目的とする。図 1(b) に示すように、Geometric Context より得られる“空”や“地面”などの確信度を特徴量として利用し、Real AdaBoost [8], [9] の弱識別器の出力を演算子によって結合することにより共起性を表現する。そのため、検出対象物体のアピランスとその周辺情報であるジオメトリの両方を考慮した検出が可能となる。また、従来の共起表現法 [10] では、和演算子と積演算子を用いた共起表現のみであった。本稿では、検出対象とは異なるカテゴリの特徴量と対象カテゴリの特徴量間の共起性を表現するために差演算子を導入する。これにより、多様な共起を表現することができるため、高精度な検出が期待できる。

2. HOG 特徴量と Geometric Context

提案手法では局所特徴量として、アピランス情報である HOG 特徴量と、Geometric Context によって得られる検出対象カテゴリと非検出対象カテゴリの確信度を用いる。本章では、提案手法で用いる HOG 特徴量と Geometric Context について述べる。

2.1 HOG 特徴量 [3]

Histograms of Oriented Gradients(HOG) 特徴量とは、局所領域における勾配方向をヒストグラム化した特徴量である。近接画素の勾配を局所領域によってヒストグラム化するため、照明の影響を受けにくく、局所的な幾何学変化に頑健という特徴がある。まず、各ピクセルの輝度 I から勾配強度 m と勾配方向 θ を次式より算出する。

$$m(x, y) = \sqrt{I_x(x, y)^2 + I_y(x, y)^2} \quad (1)$$

$$\theta(x, y) = \tan^{-1} \frac{I_y(x, y)}{I_x(x, y)} \quad (2)$$

$$\begin{cases} I_x(x, y) = I(x+1, y) - I(x-1, y) \\ I_y(x, y) = I(x, y+1) - I(x, y-1) \end{cases} \quad (3)$$

次に、算出した勾配強度 m と勾配方向 θ より、セル領域 $c(p \times p$ ピクセル) における N 方向の勾配方向ヒストグラム $V_c^{HOG} = \{v_c^{HOG}(1), v_c^{HOG}(2), v_c^{HOG}(3), \dots, v_c^{HOG}(N)\}$ を作成する。

$$V_c^{HOG} = \sum_x \sum_y m(x, y) \delta[\theta', \theta(x, y)] \quad (4)$$

ここで、 δ は Kronecker のデルタ関数を表し、勾配方向 $\theta(x, y)$ が量子化した方向 θ' に含まれる場合に 1 を返す。最後に、セル c の勾配方向ヒストグラムの勾配強度をブロック領域 ($q \times q$ セル) において正規化する。

$$v_c^{HOG}(n) = \frac{v_c^{HOG}(n)}{\sqrt{\left(\sum_{k=1}^{q \times q \times N} v_c^{HOG}(k)^2 \right) + \varepsilon}} \quad (\varepsilon = 1) \quad (5)$$

ブロック領域は、1 セル毎ずらしながら正規化する。そのため、セル c に含まれる勾配方向ヒストグラムの勾配強度は、何度もブロック領域によって正規化され、複数の特徴量が抽出される。あるセル c の勾配方向ヒストグラムの勾配方向が $1 \sim N$ とすると、正規化後の勾配方向ヒストグラム V_c^{HOG} は、 $V_c^{HOG} = \{v_c^{HOG}(1), v_c^{HOG}(2), v_c^{HOG}(3), \dots, v_c^{HOG}(B \times N)\}$ となる。ここで、 B はセル c がブロック領域に含まれる数である。

2.2 Geometric Context [7]

Hoiem 等により提案された Geometric Context(GC) は、1 枚の入力画像に対して三次元の面構造を推定し、

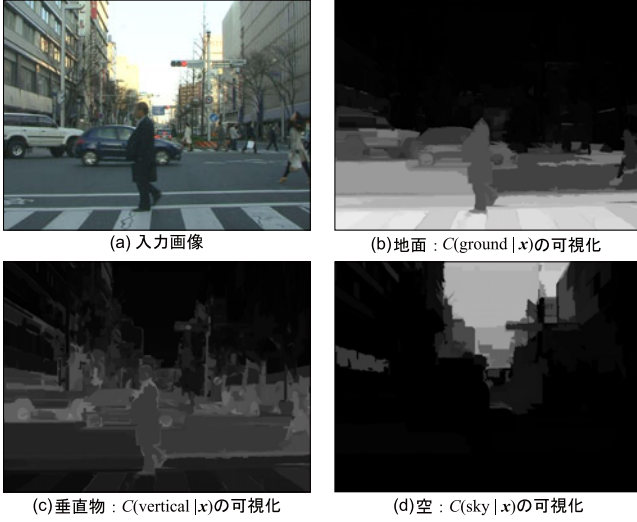


図2 Geometric Context

ラベリングする手法である。この手法は、色 (16 特徴量) やテクスチャ (15 特徴量)、位置 (8 特徴量)、形状 (4 特徴量)、三次元構造 (35 特徴量) の 73 特徴量を入力とし、ロジスティック回帰を利用した AdaBoost [11] によりモデルを学習している。識別時には、学習時に得られたモデルを用いて、未知入力画像に対して“地面 (ground)”と“垂直物 (vertical)”, “空 (sky)”に分類する。

学習は、まず入力画像を Felzenszwalb 等が提案した手法により super pixel [12] と呼ばれる領域に分割する。次に、分割された領域から色やテクスチャ、位置、形状、三次元構造の特徴量を抽出し、弱識別器に (6) 式を用いた AdaBoost により、幾何学的に均一な領域となるように最適化を行う。

$$f_m(x_1, x_2) = \sum_i \log \frac{P(y_1 = y_2, |x_{1i} - x_{2i}|)}{P(y_1 \neq y_2, |x_{1i} - x_{2i}|)} \quad (6)$$

ここで、 x_1, x_2 は隣接する super pixel の持つすべての特徴量、 y_1, y_2 は正解ラベル、 n_f は特徴数である。

最後に、次式を用いてセグメンテーションされた領域毎に“地面”、“垂直物”、“空”の3つのラベルの確信度 C を求める。

$$C(y_i = e|x) = \sum_j^{n_h} P(y_j = e|x, h_{ji})P(h_{ji}|x) \quad (7)$$

ここで、 y はクラスのラベル、 e は推定ラベル、 x は画像データ、 n_h は領域分割された数、 h_{ji} は AdaBoost により最適化された領域、 j は最適化された領域に含まれる super pixel の番号である。右辺の確率尤度関数はそれぞれ、第一項が“領域が対象ラベルである確信度”、第二項が“領域の均一性”を表現している。

入力画像 (図 2(a)) から GC による三次元の面構造を推定した結果を図 2(b), (c), (d) に示す。図 2(b) は地面、図 2(c) は垂直物、図 2(d) は空を表し、画像中の輝度が高いほどラベルの確信度が高いことを表している。

本研究では、HOG 特徴量の算出に用いたセル領域 c 毎に、GC によって得られた“地面”や“垂直物”、“空”の確信度を局所特徴量として用いる。そのため、セル領域内の確信度の平均値を求め、GC から得られる局所特徴量 $v_c^{ground}, v_c^{vertical}, v_c^{sky}$ とする。

3. 提案手法

本章では、HOG と GC を用いた特徴量間の共起による物体検出について述べる。図 3 に人検出における提案手法の流れを示す。まず、入力画像からアピアランス情報である HOG 特徴量と、GC によりジオメトリ情報である“地面”、“垂直物”、“空”の確信度を算出する。これらの局所特徴量を用いて共起確率特徴 (CPF: Co-occurrence Probability Feature) [10] を算出する。従来の共起表現法では和 (+), および積 ($\times$) の演算子のみを用いていた。本稿では、これらの演算子以外にも差 ($-$) 演算子を新たに導入することで、より多様な共起を表現する。

3.1 検出対象クラスの確率

まず、共起確率特徴量を算出するために、局所特徴量を入力とした Real AdaBoost の弱識別器の出力を算出する。弱識別器の出力は、大量の学習サンプルを用いることで統計的な信頼度を得ることができる。そのため、弱識別器の出力は検出対象クラスの確率を表す。

はじめに、局所特徴量から確率密度関数 W_+, W_- を作成する。確率密度関数 $W_{\pm}$ は、1 次元のヒストグラムにより表現され、学習サンプルの重み D_t に基づいて次式より作成する。

$$W_+^j = \sum_{i: j \in J \wedge y_i = +1} D_t(i) \quad (8)$$

$$W_-^j = \sum_{i: j \in J \wedge y_i = -1} D_t(i) \quad (9)$$

ここで、 t は学習回数、 i は学習サンプルの番号、 j は 1 次元ヒストグラムの要素番号、 y_i はクラスのラベル $y \in \{1, -1\}$ を表す。学習サンプル i から特徴量を算出し、特徴量の値に対応する 1 次元ヒストグラムの要素番号 j へ学習サンプルの重み $D_t(i)$ を加算することで、確率密度関数 $W_{\pm}$ を作成することができる。作成した確率密度関数 $W_{\pm}$ は、クラス毎の確率密度関数の総和が 1 となるように正規化する。作成したポジティブクラスとネガティブクラスの確率密度関数 $W_{\pm}$ を用いて、弱識別器の出力 $h(v)$ を次式より算出する。

$$h(v) = \frac{1}{2} \ln \frac{W_+^j + \epsilon}{W_-^j + \epsilon} \quad (10)$$

ここで、 v は特徴量、 ϵ は分母が 0 の場合に計算不能になることを防ぐための係数である。

さて、局所特徴量の共起を考えた際、二通りの場合が考えられる。一つ目は、異なる局所領域における特徴量間の共起表現、二つ目は、同一局所領域における特徴量間の共起表現である。

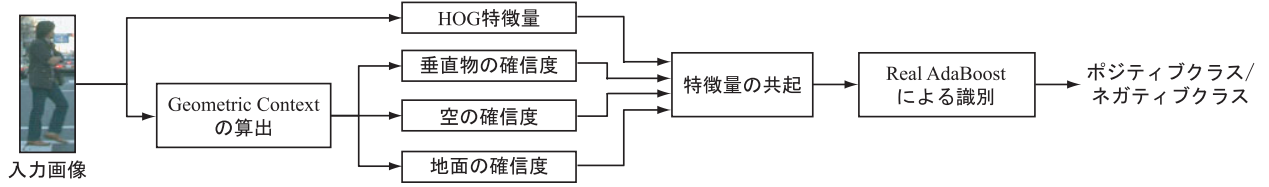


図 3 提案手法による識別の流れ

3.2 GC を用いた異なる局所領域における特徴量間の共起表現 (和・積演算子による共起)

異なる局所領域における特徴量の共起を表現するために、和 (+)、および積 (×) の演算子を用いた共起確率特徴量を生成する。ここまでで得られた Real AdaBoost の弱識別器の出力を用いて次式により共起確率特徴量を生成する。

$$CPF_1 = \begin{cases} h_1(v_{c_1}^{HOG}(n_1)) \\ h_1(v_{c_1}^{ground}) \\ h_1(v_{c_1}^{vertical}) \\ h_1(v_{c_1}^{sky}) \end{cases} + \begin{cases} h_2(v_{c_2}^{HOG}(n_2)) \\ h_2(v_{c_2}^{ground}) \\ h_2(v_{c_2}^{vertical}) \\ h_2(v_{c_2}^{sky}) \end{cases} \quad (11)$$

$$CPF_2 = \begin{cases} h_1(v_{c_1}^{HOG}(n_1)) \\ h_1(v_{c_1}^{ground}) \\ h_1(v_{c_1}^{vertical}) \\ h_1(v_{c_1}^{sky}) \end{cases} \times \begin{cases} h_2(v_{c_2}^{HOG}(n_2)) \\ h_2(v_{c_2}^{ground}) \\ h_2(v_{c_2}^{vertical}) \\ h_2(v_{c_2}^{sky}) \end{cases} \quad (12)$$

ここで、 $h_1(v_{c_1}^{HOG}(n_1))$, $h_2(v_{c_2}^{HOG}(n_2))$ は HOG 特徴量 $v_{c_1}^{HOG}(n_1)$, $v_{c_2}^{HOG}(n_2)$ を入力した際の弱識別器の出力、 $h_1(v_{c_1}^{ground})$, $h_2(v_{c_2}^{ground})$ は GC から得られる“地面”の確信度 $v_{c_1}^{ground}$ を入力した際の弱識別器の出力、 $h_1(v_{c_1}^{vertical})$, $h_2(v_{c_2}^{vertical})$ は GC から得られる“垂直物”の確信度 $v_{c_1}^{vertical}$ を入力した際の弱識別器の出力、 $h_1(v_{c_1}^{sky})$, $h_2(v_{c_2}^{sky})$ は GC から得られる“空”の確信度 $v_{c_1}^{sky}$ を入力した際の弱識別器の出力を表す。

和 (+) 演算子は特徴量間の弱い関係性を表現する共起、積 (×) 演算子は特徴量間の強い関係性を表現する共起を表している。図 4 に異なる局所領域における特徴量間の共起の例を示す。図 4 では検出対象が人である場合、人の肩付近における HOG 特徴量と、人の足元付近における“地面”の確信度の共起を表現している。このような共起を表現することで、対象カテゴリの局所特徴量と非対象カテゴリの局所特徴量間の共起をコンテキストとして表現することができる。

3.3 GC を用いた同一局所領域内における特徴量間の共起表現 (差演算子による共起)

和 (+)、および積 (×) の演算子を用いた共起確率特徴量では、特徴量から検出対象の確率を算出し、共起を表現した。この共起表現法では、GC より得られる非検出対象カテゴリの確信度を検出対象の確率として算出している。本研究では、HOG 特徴量の検出対象の過剰な確率を抑制するために、ジオメトリ情報を有効活用する差

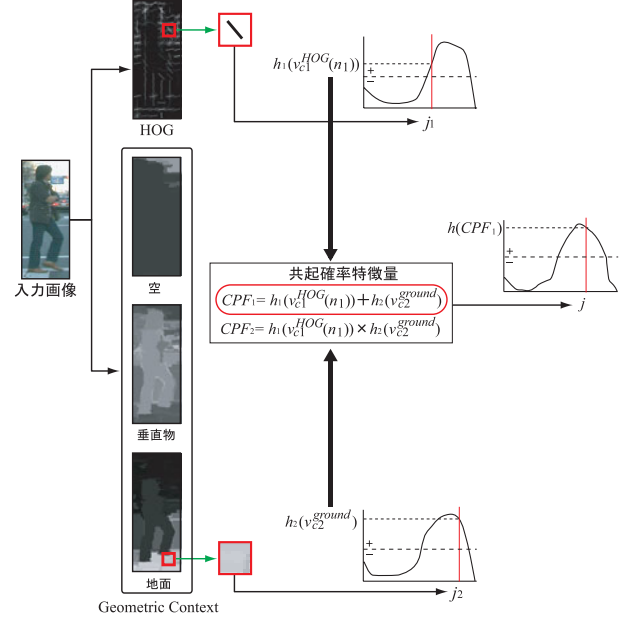


図 4 異なる局所領域における共起表現の例

演算子を用いた共起表現を提案する。これにより、検出対象とは異なるカテゴリの確信度の共起を表現することが可能となるため、HOG 特徴量のみでは信頼性の低い特徴量であっても、異なるカテゴリの確信度を考慮することができる。次式より差 (-) 演算子を用いた共起確率特徴量を算出する。

$$CPF_3 = h_1(v_{c_1}^{HOG}(n_1)) - \begin{cases} v_{c_1}^{ground} \\ v_{c_1}^{vertical} \\ v_{c_1}^{sky} \end{cases} \quad (13)$$

図 5 に差演算子を用いた共起確率特徴量の生成例を示す。HOG 特徴量だけでは車両らしさが高くなるような学習サンプルに対して、その局所領域における“地面”らしさが非常に高い場合は、その局所領域における HOG 特徴量を用いた識別が抑制される。

3.4 共起確率特徴量を用いた Real AdaBoost による最終識別器の構築

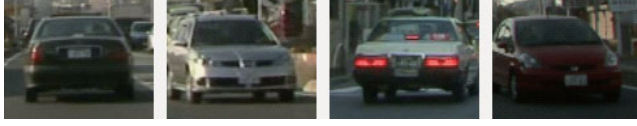
ここまでで述べた共起確率特徴量を用いて、検出対象と非検出対象の判別を行う最終識別器を学習する。学習は、全ての特徴量の組み合わせを行った場合、組み合わせ数が多く、膨大な学習時間を必要とする。そこで、文献 [10] と同様に 2 段階の Real AdaBoost により最終識別器を構築する。1 段階目の Real AdaBoost では、全て



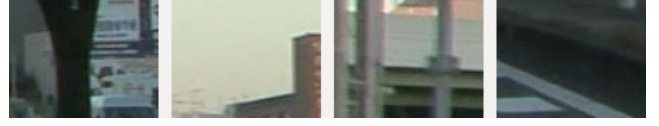
(a)人検出用ポジティブサンプル



(b)人検出用ネガティブサンプル



(c)車両検出用ポジティブサンプル



(d)車両検出用ネガティブサンプル

図 6 学習用サンプル例

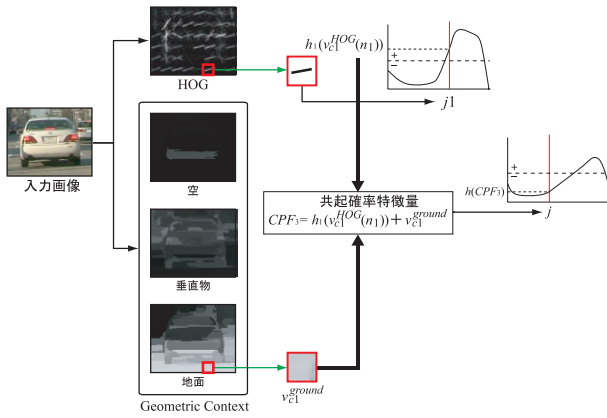


図 5 同一局所領域内における共起表現の例

の局所特徴量から識別に有効な特徴量の選択を行い、特徴 pool を作成する。そして、作成された特徴 pool の特徴量のみを用いて共起確率特徴量を生成する。

2 段階目の Real AdaBoost では、生成された共起確率特徴量を用いて、式 (8), (9) と同様の方法により確率密度関数 $W_{\pm}$ を作成し、次式より評価値 α を求める。

$$\alpha = \sum_j \sqrt{W_+^j W_-^j} \quad (14)$$

生成した共起確率特徴量の中から最も評価値 α が小さい共起確率特徴量を弱識別器として選択する。最終的な識別器は、次式で表される弱識別器の線形和となる。

$$H(CPF) = \text{sign} \left[\sum_{t=1}^T h_t(CPF) \right] \quad (15)$$

ここで CPF は、 $CPF \in \{CPF_1, CPF_2, CPF_3\}$ とする。

4. 評価実験

本章では、提案手法の有効性を示すために、検出対象を人、および車両とし、実験により従来法との比較を行う。

4.1 学習・評価データベース

画像データベースは、車両のルームミラー付近に取り

表 1 データベース

検出対象	学習用		評価用	
	ポジティブ	ネガティブ	ポジティブ	ネガティブ
人	1,215	10,416	1,836	2,108
車両	710	8,800	1,230	3,880

付けたカメラにより、実環境下で 13 時～19 時の間に撮影した映像をもとに作成する。撮影した映像の中から、人、および車両の全体が含まれるように切り出すことで作成する。切り出した画像は、背景や照明、検出対象の向き、検出対象同士による画像中の重なりによる部分的な隠れなどが含まれている。表 1 に人、および車両の学習と評価に用いるサンプル数を示す。人および車両でポジティブクラスとネガティブクラスのサンプル数の比率が異なるが、Real AdaBoost の重みの初期化時に考慮することで対応することができる。また、図 6 に学習に用いたサンプルの一部を示す。

4.2 評価方法

表 1 のデータベースを用いて、学習および評価を行う。評価実験では、下記に示す HOG ベースの従来法と HOG と GC の共起を用いた提案手法の比較を行う。

- 従来法 1: HOG+Real AdaBoost [3]
- 従来法 2: HOG 特徴量の共起表現 [10]
- 提案手法 1: GC を用いた異なる局所領域における特徴量間の共起表現 (和と積演算子による共起)
- 提案手法 2: GC を用いた同一局所領域における特徴量間の共起表現 (差演算子による共起)
- 提案手法 3: GC を用いた異なる及び同一局所領域における特徴量間の共起表現 (和と積, 差演算子による共起)

実験結果の比較には、Receiver Operating Characteristic (ROC) カーブを用いる。ROC カーブとは、横軸に誤検出率 (FP)、縦軸に検出率 (TP) を表したものである。識別器の閾値を変化させることによって、誤検出率に対する検出率の比較を行うことが可能である。グラフ左上 (FP:0%, TP:100%) に近いほど検出性能が良いことを表す。

4.3 GCを用いた異なる局所領域における和と積演算子による実験結果

GCを用いた異なる局所領域における特徴量間の共起表現(和と積演算子による共起)の有効性を確認するために、従来法1、従来法2、提案手法1の比較を行う。図7にROCカーブによる実験結果を示す。図7に示すように、人および車両の両検出対象において提案手法1がグラフの左上にあることから、最も検出性能が高いことがわかる。検出対象が人の場合では、誤検出率1.0%の場合、提案手法1は従来法1と比較して約5.0%、従来法2と比較して約1.2%検出率が向上した。検出対象が車両の場合では、誤検出率0.1%の場合、提案手法は従来法1と比較して約4.6%、従来法2と比較して約1.6%検出率が向上した。この結果から、GCから得られるジオメトリの情報は人および車両検出において有効であることがわかる。

図8に、従来法1では非検出対象と判定されたが、提案手法1では検出対象と判定された評価サンプル例を示す。図中に示す値は、従来法と提案手法の強識別器の出力値を示す。これらの評価サンプルは、夕刻以降に撮影された画像であるため、アピアランスの情報であるHOG特徴量のみでは、人の輪郭をはっきり捉えることができない。この問題に対して、提案手法ではGCから得られるジオメトリ情報も用いているため、アピアランスのみでは識別が難しいサンプルに対しても正しい識別が可能となった。

4.4 GCを用いた同一局所領域における差演算子による実験結果

GCを用いた同一局所領域における特徴量間の共起表現(差演算子による共起)の有効性を確認するために、従来法1と提案手法2の比較を行う。図9にROCカーブによる実験結果を示す。図9に示すように、差(-)の演算子による共起表現を導入することで、従来手法1よりグラフ全体が左上にシフトしているため、検出性能が向上していることがわかる。また、従来手法では極端に識別時の信頼度が高く誤検出していたデータ例を図10に示す。図中に示す値は従来法と提案手法の強識別器の出力値を示す。提案手法は、空や地面等のGCを用いた差の演算子の共起を導入することにより、強識別器の出力値が低くなっている。これは、学習時に差の演算子により識別に有効ではないHOG特徴量が選択されることを避けることができ、誤検出を抑制する方向に作用したためといえる。

4.5 和と積、差演算子による共起確率特徴による実験結果

和と積、差演算子による共起確率特徴による実験結果を表2に示す。検出対象が人の場合では、誤検出率1.0%の場合、提案手法3はHOGのみを用いた従来法1と比

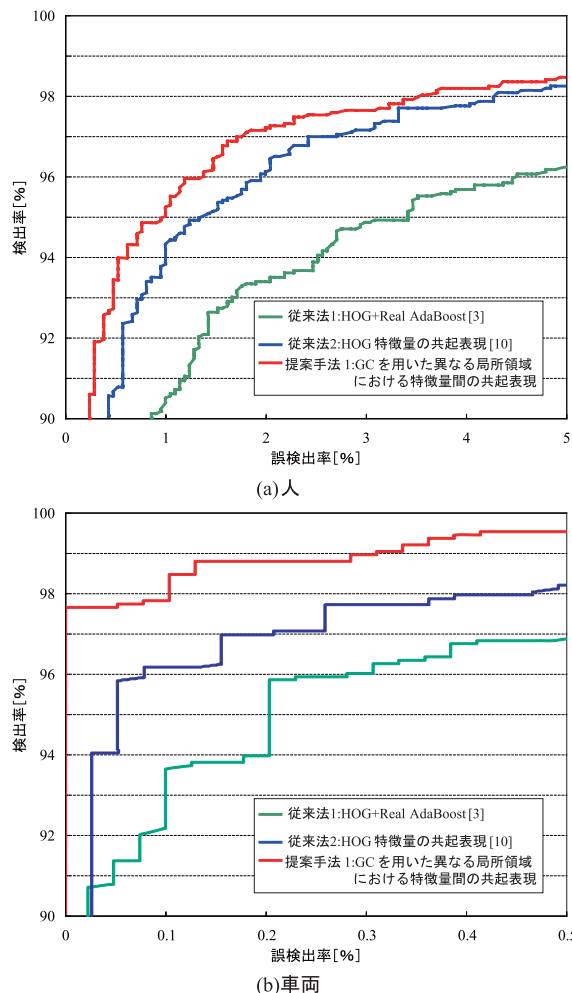
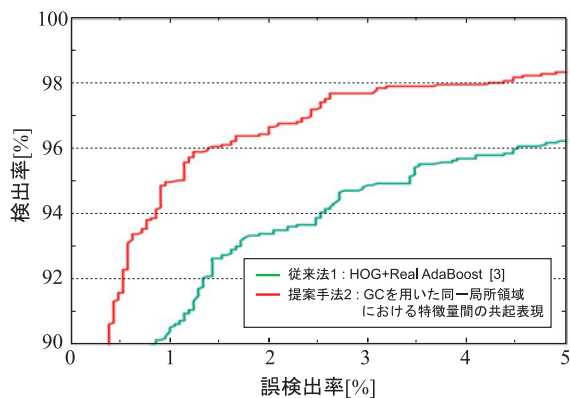


図7 和・積演算子による共起確率特徴による実験結果

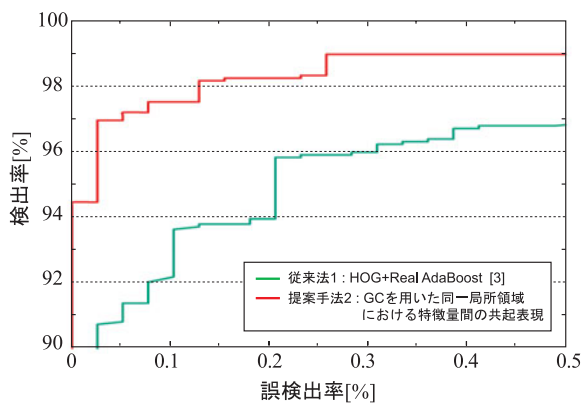


図8 従来法では未検出となる例

較して約6.8%検出率が向上した。検出対象が車両の場合では、誤検出率0.1%の場合、提案手法3は従来法1と比較して約6.4%検出率が向上した。さらに提案手法3は、GCを用いた和と積の演算子による提案手法1と、



(a)人



(b)車両

図9 差演算子による共起確率特徴による実験結果

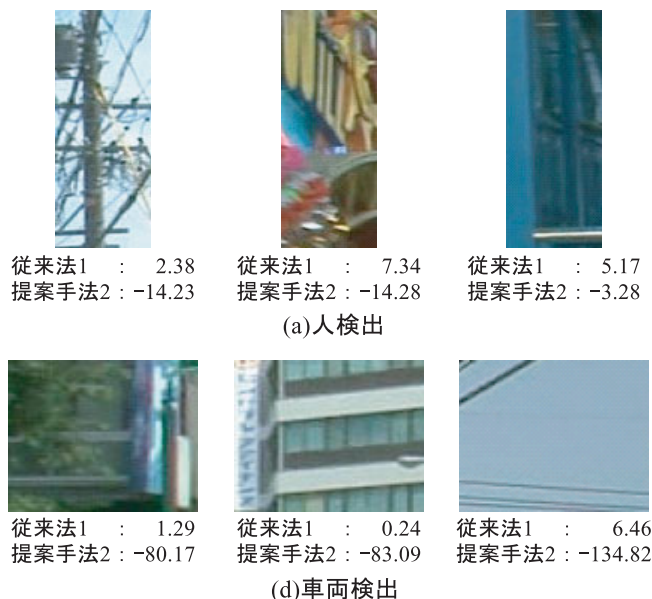


図10 従来法では誤検出となる例

差演算子のみの提案手法2より検出率が高い。以上より、和と積、差の全ての演算子を用いることで多様な共起表現が可能となり、検出性能を向上させることができた。

4.6 考察

4.6.1 人検出、および車両検出に有効な特徴

図11にHOG特徴量、およびGCから得られる“地

表2 和と積、差演算子による検出率[%]
(人：誤検出率1.0%、車両：誤検出率0.1%)

	従来法1	提案手法1 (和、積)	提案手法2 (差)	提案手法3 (和、積、差)
人	90.5	95.5	95.0	97.3
車両	92.1	97.3	97.5	98.5

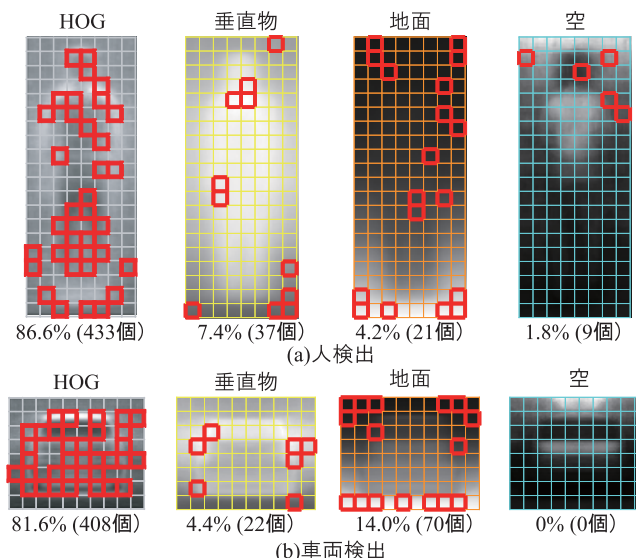


図11 学習により選択された特徴量の割合と個数

面”や“垂直物”，“空”の中から選択された特徴量とその領域を正方形(赤色)で示す。画像の輝度値が高いほど、HOGの場合は勾配が高く、ジオメトリ情報は各カテゴリ(垂直物、地面、空)の確信度が高いことを意味している。HOG特徴量に関しては、検出対象が人、車両であってもシルエット全体を捉えるような傾向があり、選択された特徴量の中で最も多く、約80%をHOG特徴量が占めている。GCの“地面”や“垂直物”，“空”を表す特徴量に関しては、車両の場合、人に比べて“地面”らしさが多く選択されていることがわかる。人の場合、ガードレールなどにより足元に隠れが発生することが多い。車両の場合は、車両の下端に隠れが発生することは人に比べて頻度が少ない。そのため、検出対象を人にした場合は、車両に比べて“地面”らしさが選択される状況が少なく、代わりに僅かではあるが“空”や“垂直物”が選択されている。

次に、“地面”に着目すると、確信度の低い領域も選択されている。これは、人検出においては人の足元付近が“地面”であることだけでなく、人の上部付近は“地面”でないことを表現していることがわかる。車両検出においても人と同様に、車両の下部が“地面”であることだけでなく、車両の上部付近は“地面”でなく、“車両”を表現していることがわかる。つまり、ある局所領域における“地面”らしさ、“空”らしさが高いものが必ずしも“人”や“車両”らしさを表現するだけでなく、その周辺の“地面”らしさが低いからこそ“人”や“車両”らしいといえる場合もある。

4.6.2 未検出, および誤検出について

図12にGCを用いた異なる及び同一局所領域における特徴量の共起表現による誤識別(未検出, および誤検出)を示す. 検出対象を人とした場合の未検出は, 過剰な隠れ, およびHOG特徴量やGCから得られる“垂直物”を表す特徴量が算出されにくい夜間のデータで発生する傾向がある. 検出対象を車両とした場合の未検出も同様に, HOG特徴量やGCから得られる“地面”を表す特徴量が算出されにくい夜間のデータで発生する傾向がある. 検出対象を人とした場合の誤検出は, 人を表すような縦エッジが抽出されやすく, “垂直物”を表す特徴量が抽出されやすいデータで発生する傾向がある. 検出対象を車両とした場合の誤検出は, 車両を表すような横エッジが抽出されやすく, “地面”を表す特徴量が抽出されやすいデータで発生する傾向がある. これらの原因のひとつは, 人検出の場合, 特徴量の選択状況において, “垂直物”らしさを選択する割合が高いために発生する問題である. 例えば, 人画像において“垂直物”らしさが抽出されにくいと未検出に, 背景画像において“垂直物”らしさが抽出されやすいと誤検出になると考えられる. また, 車両検出の場合では, 特徴量の選択状況において, “地面”らしさを選択する割合が高い. そのため, 車両画像において“地面”らしさが抽出されにくいと未検出に, 背景画像において“地面”らしさが抽出されやすいと誤検出になると考えられる. これらの誤識別への対応は今後の課題としたい.

5. おわりに

本稿では, GC用いた特徴量間の共起による物体検出能力の高精度化について述べ, その有効性を示した. 提案手法は, HOG特徴量とGCにより得られる“地面”, “垂直物”, “空”らしさを表す特徴量を入力とし, 検出対象とは異なるカテゴリの特徴量と対象カテゴリの特徴量間の共起特徴から Real AdaBoost を用いて検出対象とそれ以外に判別する. “地面”, “垂直物”, “空”らしさを表す特徴を演算子により多様な捉え方で観測することが可能になり, 従来の共起を表現する方法よりも高精度な検出が可能となった. 今後は, 共起確率特徴量に用いる演算子に商($\div$)を加え, 多クラスの識別を拡張する予定である.

文 献

- [1] 平成20年度警察白書: <http://www.npa.go.jp/>
- [2] P. Viola and M. Jones: “Robust real-time face detection”, Int. Journal of Computer Vision, Vol. 57, pp. 137-154 (2004).
- [3] N. Dalal and B. Triggs: “Histograms of oriented gradients for human detection”, IEEE Computer Vision and Pattern Recognition, Vol. 1, pp. 886-893 (2005).
- [4] P. Negri, X. Clady and L. Prevost: “Benchmarking Haar and Histograms of Oriented Gradients features applied to vehicle detection”, Int. Conference on Informatics in Control (2007).

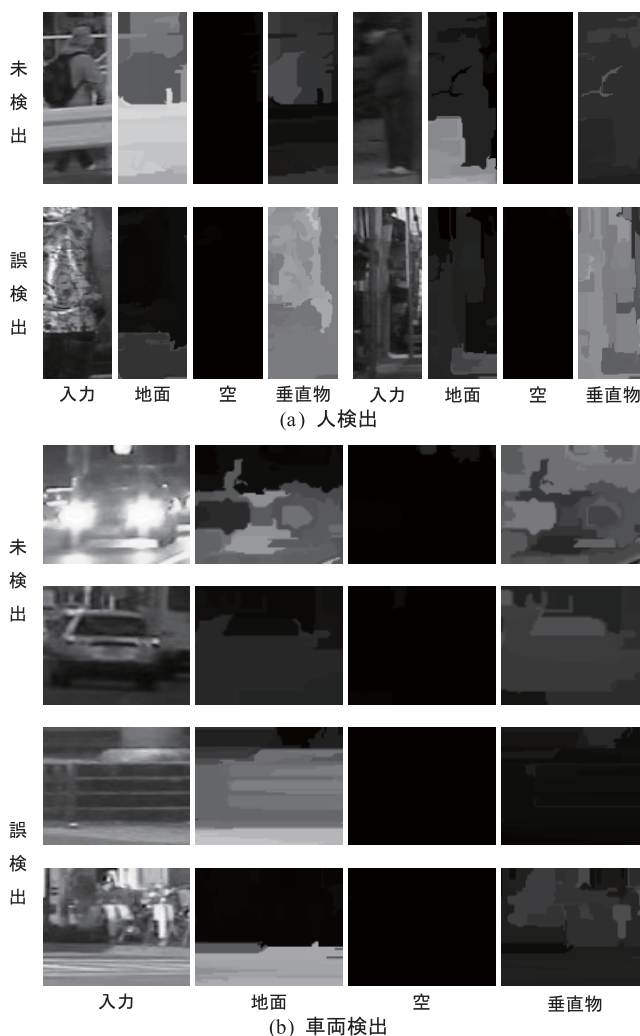


図12 提案手法により誤識別したサンプル例

- [5] 岡部孝弘, 近藤雄飛, 木谷クリス真実, 佐藤洋一: “カテゴリの共起を考慮した物体認識”, 画像の認識・理解シンポジウム (MIRU2008), pp. 217-222 (2008)
- [6] A. Rabinovich, A. Vedaldi, C. Galleguillos, E. Wiewiora and S. Belongie: “Objects in Context”, IEEE Int. Conference on Computer Vision, pp. 1-8 (2007)
- [7] D. Hoiem, A. A. Efros and M. Hebert: “Geometric context from a single image”, IEEE Int. Conference on Computer Vision, Vol. 1, pp. 654-661 (2005)
- [8] R. E. Schapire and Y. Singer: “Improved boosting algorithms using confidence-rated predictions”, Machine Learning, Vol. 37, pp. 297-336 (1999)
- [9] B. Wu and R. Nevatia: “Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors”, IEEE Computer Vision and Pattern Recognition, pp. 90-97 (2005)
- [10] 山内悠嗣, 藤吉弘亘, 山下隆義: “Boostingに基づく特徴量の共起表現による人検出”, 画像の認識・理解シンポジウム (MIRU2008), pp. 207-214 (2008)
- [11] M. Collins, R. E Schapire and Y. Singer: “Logistic Regression, AdaBoost and Bregman Distances”, Machine Learning, Vol. 48, No. 1-3, (2002)
- [12] P. Felzenszwalb and D. Huttenlocher: “Efficient Graph-Based Image Segmentation”, Int. Journal of Computer Vision, Vol. 59, No. 2, pp. 167-181 (2004).