

CS-HOG：色の類似性に基づいた形状特徴量

後藤 雄飛^{†a)} 山内 悠嗣^{†b)} 藤吉 弘亘^{†c)}

CS-HOG: Color Similarity-Based HOG

Yuhi GOTO^{†a)}, Yuji YAMAUCHI^{†b)}, and Hironobu FUJIYOSHI^{†c)}

あらまし 従来の物体検出手法では、HOG 特徴量に代表されるように物体の形状を捉える局所特徴量が多く利用されている。近年では、色情報を利用した局所特徴量として Color Self-Similarity (CSS) 特徴量が提案されている。CSS 特徴量は色の類似度を算出し、同一物体らしさを観測することにより、物体検出に有効な特徴を表現することができる。また、CSS 特徴量は HOG 特徴量のような形状を捉える特徴量と併用することにより、検出性能が向上することが報告されている。そこで、本研究では色の類似性に基づいて物体形状を観測する CS-HOG 特徴量を提案する。CS-HOG 特徴量は、色の類似性を用いた同一物体らしさを利用することで、物体領域を推定して形状を捉えることができる。そのため、輝度のこう配に基づく HOG 特徴量と比較して、物体形状を明確に観測することが可能となり高精度な物体検出が実現できる。評価実験より、CS-HOG 特徴量は HOG 特徴量、CSS 特徴量と比較して約 22.5%、27.2% 識別精度が向上した。

キーワード 物体検出, 局所特徴量, HOG 特徴量, CSS 特徴量, Real AdaBoost

1. ま え が き

物体検出技術は、車載カメラによる Intelligent Transport System (ITS: 高度道路交通システム)、監視カメラによるセキュリティやマーケティングなど、多くの分野で実用化が期待されている。多くの物体検出手法は、学習サンプルから抽出した局所特徴量と統計的学習手法を用いて構築した識別器により実現している [1]~[7]。高精度な物体検出を実現するためには、物体の特徴を捉えるような局所特徴量を設計することが重要である。物体の特徴を表現する局所特徴量を設計する手法として、これまでに Histograms of Oriented Gradients (HOG) 特徴量 [8]、Edge Orientation Histograms (EOH) 特徴量 [9]、Edgelet 特徴量 [10] など、輝度のこう配に基づいて物体形状を捉える局所特徴量が多く提案されている。中でも、HOG 特徴量は局所領域においてこう配の方向ごとにこう配強度を累積した特徴量であり、局所領域で正規化処理

を行うため、照明変化や幾何学的変化に対して頑健な特徴量となる。

形状を捉えるこれらの特徴量は画像の輝度のこう配に基づいているため、色情報は利用されていない。カラー画像は、グレースケール画像の 3 倍の情報量をもつが、人における衣類や自動車におけるボディなど対象ごとに色が異なるため、物体検出に有効な特徴量を抽出することが困難とされていた。一方、物体追跡 [11], [12] や人物照合 [13], [14]、画像分類 [15], [16] や特定物体認識 [17], [18] などの分野では色情報を効果的な情報として多く利用されている。しかし、物体検出に色情報を用いた局所特徴量として Color Self-Similarity (CSS) 特徴量 [19] が提案され、色情報が物体検出に有効であることが確認されている。CSS 特徴量は、二つの局所領域の色ヒストグラムから色の類似性を捉え、同一物体らしさや非同一物体らしさを観測することで、物体検出に有効な特徴を表現をする。このように、高精度な物体検出を実現するためには、物体の特徴を効果的に表現する局所特徴量の設計が重要なこととなる。

そこで、本研究では物体の特徴を効果的に表現する局所特徴量として Color Similarity-based HOG (CS-HOG) 特徴量を提案する。CS-HOG 特徴量はカラー

[†] 中部大学大学院工学研究科, 春日井市
Dept. of Computer Science, Chubu University, 1200
Matsumoto, Kasugai-shi, 487-8501 Japan

a) E-mail: yuhi@vision.cs.chubu.ac.jp

b) E-mail: yuu@vision.cs.chubu.ac.jp

c) E-mail: hf@cs.chubu.ac.jp

画像を利用し、色の類似性に基づいて物体形状を捉える局所特徴量である。CS-HOG 特徴量は、色の類似性を用いた同一物体らしさを利用することで、物体領域を推定して形状を効果的に捉えることができる。これまでの輝度の明暗関係を利用したこう配の場合、検出対象ごとに物体と背景の明暗関係が反転するおそれがある。そのため、HOG 特徴量は明暗関係によるこう配の反転を防ぐ処理を行っている。これに対して、CS-HOG 特徴量は色の類似性を用いて物体領域を推定することで、物体領域と背景領域からこう配を取得して形状を求めることができる。そのため、検出対象ごとに統一したこう配を求めることを可能にし、HOG 特徴量と比べて効果的な特徴表現をする。また、CS-HOG 特徴量は色の類似性と物体形状の関係性に着目することで、対象物体に似た形状による誤検出を抑制できる。このように、CS-HOG 特徴量により色の類似性に基づいて適応的に物体形状を捉えることで、高精度な物体検出を実現する。

本論文の以降の章は次のように構成される。まず、**2.**において従来の局所特徴量である HOG 特徴量と CSS 特徴量の算出方法について説明し、二つの特徴量の特性について比較する。**3.**にて、本研究で提案する CS-HOG 特徴量の算出方法について説明し、HOG 特徴量との違いについて比較する。**4.**にて、CS-HOG 特徴量の有効性を確認するため評価実験を行う。最後に**5.**にて本研究についてまとめる。

2. 形状特徴と色の類似特徴

物体検出を実現するには、学習サンプルから抽出した局所特徴量を用いて学習した識別器を利用する手法が一般的である。物体の形状を捉える局所特徴量として Dalal らが HOG 特徴量を提案した。従来の物体検出手法は、HOG 特徴量に代表されるようにグレースケール画像から物体の形状を捉える局所特徴量が多く利用されている。これに対して、カラー画像の色情報を利用した局所特徴量として、Walk らが CSS 特徴量を提案した。CSS 特徴量は、色の類似性に着目することで物体検出に有効な特徴表現を可能にする。本章では、提案する CS-HOG 特徴量と関連の深い局所特徴量として、形状特徴である HOG 特徴量と、色の類似特徴である CSS 特徴量について述べる。

2.1 形状特徴：HOG 特徴量

Histograms of Oriented Gradients (HOG) 特徴量 [8] はグレースケールの入力画像から、セルと呼

ばれる局所領域ごとに作成したこう配方向ヒストグラムを特徴量とする。また、複数のセルで構成されるブロック領域ごとに特徴量を正規化することで、照明変化や幾何学的変化の影響を受けにくい特徴量となる。

算出手順は、はじめに入力画像をあらかじめ定義したサイズに正規化し、あるピクセルの輝度値 $L(x, y)$ のこう配強度 m とこう配方向 θ を式 (1)~(3) より算出する。

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

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

$$\begin{cases} L_x(x, y) = L(x+1, y) - L(x-1, y) \\ L_y(x, y) = L(x, y+1) - L(x, y-1) \end{cases} \quad (3)$$

算出したこう配方向 θ は $0^\circ \sim 360^\circ$ の値で算出されるが、 180° より大きくなる方向は 180° 引いて $0^\circ \sim 180^\circ$ とする。これにより、検出対象と背景領域の輝度の明暗関係に依存しないこう配方向を得ることができる。次に、算出したこう配強度 m とこう配方向 θ を用いて、式 (4) によりセル c ($M \times M$ ピクセル) におけるこう配方向ヒストグラム $\mathbf{V}_c = \{v_c(1), v_c(2), \dots, v_c(N^{hog})\}$ を作成する。

$$v_c(n) = \sum_x \sum_y m(x, y) \delta[f(\theta(x, y)), n] \quad (4)$$

ここで、 $n = \{1, 2, \dots, N^{hog}\}$ はヒストグラムのピンの番号、 $f(\theta)$ はこう配方向 θ を N^{hog} 方向に量子化した値、 $\delta[\cdot]$ はクロネッカーのデルタ関数を表しており、二つの要素が等しい場合は 1、それ以外は 0 を出力する関数である。このように、セル c ごとにヒストグラム化することにより、局所領域内の微小な幾何学的変化に対して頑健な特徴量となる。最後に、式 (5) を用いて各セル c で作成したこう配方向ヒストグラム \mathbf{V}_c を、複数のセルで構成されるブロック ($K \times K$ セル) 領域ごとに正規化する。

$$h_c(n) = \frac{v_c(n)}{\sqrt{\sum_{i=1}^{K \times K} \sum_{n=1}^{N^{hog}} v_i(n)^2 + \epsilon}} \quad (5)$$

ここで、 ϵ は分母が 0 になることを防ぐための定数 ($\epsilon = 1$) である。ブロック領域は、1 セルごとにセル領域が重なるように移動しながら正規化する。正規化後のこう配方向ヒストグラム \mathbf{H}_c は、

$H_c = \{h_c(1), h_c(2), \dots, h_c(K \times K \times N^{hog})\}$ となる。64 × 128 ピクセルの入力画像に対して、ブロック領域は $\{(64/M) - (K-1)\} \times \{(128/M) - (K-1)\}$ 個となる。この場合、HOG 特徴量の特徴ベクトルは、 $M = 8$, $K = 2$, $N^{hog} = 9$ のとき 105 個のブロック領域により正規化処理が行われるため、 $2 \times 2 \times 9 \times 105 = 3,780$ 次元となる。

2.2 色の類似特徴：CSS 特徴量

Color Self-Similarity (CSS) 特徴量 [19] は、二つのセル領域において作成した色ヒストグラムの類似度を特徴量としている。二つの局所領域における色の類似度を観測することで、検出対象の同一物体らしさや非同一物体らしさを捉える特徴量である。また、色の類似度を特徴量とすることで、検出対象の色に依存しにくいという利点がある。

算出手順は、入力画像をあらかじめ定義したサイズに正規化し、分割したセル c ($M \times M$ ピクセル) の色ヒストグラム $\mathbf{p}_c = \{p_c(H, 1), p_c(H, 2), \dots, p_c(H, N^{css}), \dots, p_c(V, N^{css})\}$ を作成する。式 (6) に示すように色ヒストグラム \mathbf{p}_c は、セルごとに HSV 表色系 $\{H, S, V\}$ で表される各ピクセルの色 $I_c(x, y, H)$, $I_c(x, y, S)$, $I_c(x, y, V)$ を累積することで作成する。

$$p_c(r, n) = \sum_{x=1}^M \sum_{y=1}^M \delta[f(I_c(x, y, r)), n] \quad (6)$$

ここで、 r は HSV 表色系 $\mathcal{D} = \{H, S, V\}$ の各要素、 n は量子化番号 ($n = 1, 2, \dots, N^{css}$)、 $f(I_c(x, y, r))$ は色 $I_c(x, y, r)$ を量子化した値を表している。二つのセル c_i , c_j から作成した色ヒストグラム \mathbf{p}_{c_i} , \mathbf{p}_{c_j} を用いて、セル間の色の類似度 $s(c_i, c_j)$ を式 (7) に示すユークリッド距離を利用して色の類似度を算出する。

$$s(c_i, c_j) = \sqrt{\sum_{r \in \mathcal{D}} \sum_{n=1}^{N^{css}} (p_{c_i}(r, n) - p_{c_j}(r, n))^2} \quad (7)$$

ヒストグラム間の距離の算出は、一般的にヒストグラムインタセクションや Bhattacharyya 係数が用いられる。本研究では文献 [19] と同様に、予備実験の結果、ユークリッド距離が最も有効に働いたため、類似度の算出法として採用する。64 × 128 ピクセルの入力画像に対して、セル c は $(64/M) \times (128/M)$ 個となる。この場合、CSS 特徴量の特徴ベクトルは、 $M = 8$ のとき 128 個のセルから全ての組合せについて計算するため、 ${}_{128}C_2 = 8,128$ 次元となる。

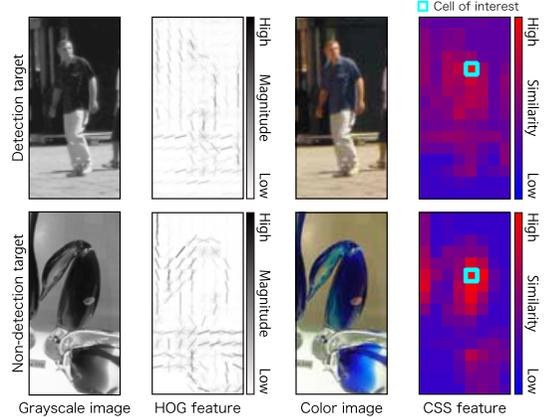


図 1 HOG 特徴量と CSS 特徴量の比較

Fig. 1 Comparison of HOG feature and CSS feature.

2.3 HOG 特徴量と CSS 特徴量の比較

HOG 特徴量と CSS 特徴量の特徴表現の特性について比較する。HOG 特徴量と CSS 特徴量を可視化して比較したものを図 1 に示す。HOG 特徴量は、局所領域内の輝度から得られるこの配を利用して物体の形状を表現する。しかし、図 1 の検出対象に示すように、画像のグレースケール画像はカラー画像と比較して、物体と背景の輝度が類似してしまう場合がある。このような画像から HOG 特徴量を算出すると、検出対象の形状を明確に捉えることができない。一方、局所領域間の色の類似性を表現する CSS 特徴量は、グレースケール画像より情報量の多いカラー画像を利用することで、物体と背景と色の違いを捉えることができている。このように、HOG 特徴量は物体形状を局所的に捉えることができる反面、輝度の情報に依存してしまうという問題がある。図 1 の非検出対象の例は、画像中に検出対象ではない物体が含まれる場合を示している。このとき、CSS 特徴量は物体の詳細な形状を表現できないため、検出対象に似た特徴を捉えてしまい誤検出となるおそれがある。このように、CSS 特徴量はカラー画像による色情報を利用することで、物体と背景の領域を捉えることができる反面、詳細な物体の特徴を表現できないという問題がある。そこで、本研究ではこのような HOG 特徴量と CSS 特徴量の特性を生かし、両者の長所を捉えることができる局所特徴量を設計する。次章より、本研究で設計する CS-HOG 特徴量について説明する。

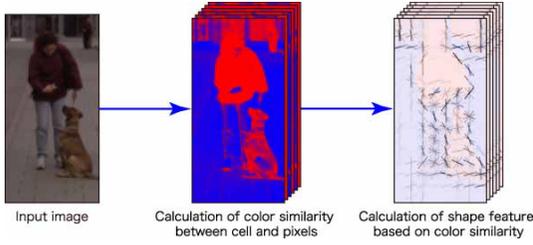


図 2 CS-HOG 特徴量の算出手順
Fig. 2 Calculation of CS-HOG feature.

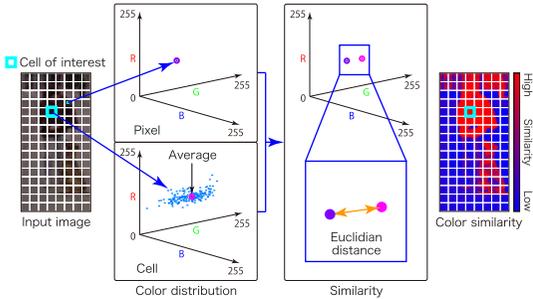


図 3 局所領域とピクセル間の色の類似度の算出
Fig. 3 Calculation of color similarity between cell and pixels.

3. 色の類似性に基づいた形状特徴： CS-HOG 特徴量

本研究では、色の類似性に基づいて形状を捉える Color Similarity-based HOG (CS-HOG) 特徴量を設計する。CS-HOG 特徴量は、色の類似度を利用した同一物体らしさを用いることにより、物体と背景の境界を示す形状を明確に表現することが可能である。HOG 特徴量と比較して物体の特徴を効果的に表現することができるため、高精度な物体検出が期待できる。CS-HOG 特徴量の算出の流れは、図 2 に示すように入力画像から局所領域とピクセル間の色の類似度を算出し、色の類似度を用いて形状特徴を算出する。以下より、CS-HOG 特徴量の算出方法について詳しく説明する。

3.1 局所領域とピクセル間の色の類似度の算出

CS-HOG 特徴量の算出手順は、まず、入力画像をあらかじめ定義したサイズに正規化し、局所領域（セル）と各ピクセルの色の類似度を算出する。図 3 に示すように、注目セル c' ($M \times M$ ピクセル) から RGB 表色系 $\{R, G, B\}$ で表される各ピクセルの色 $I_{c'}(x, y, R)$, $I_{c'}(x, y, G)$, $I_{c'}(x, y, B)$ により、注目セル c' 内の色

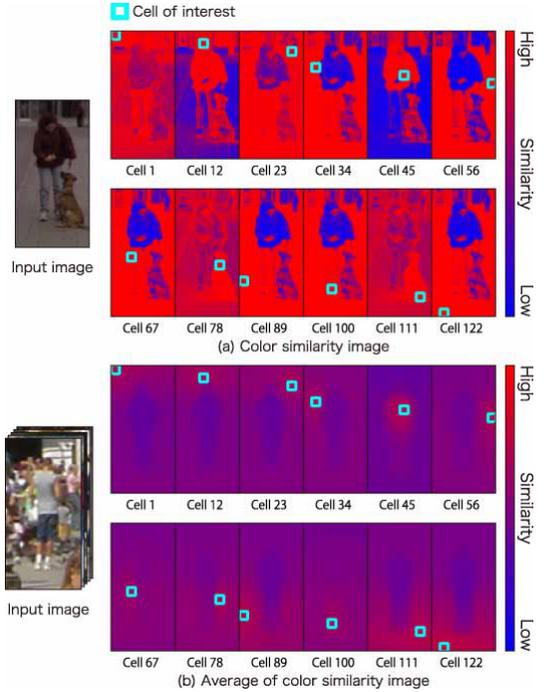


図 4 類似度画像の可視化
Fig. 4 Visualization of color similarity images.

の代表値 $I'(c', r)$ を式 (8) を用いて算出する。

$$I'(c', r) = \frac{1}{M \times M} \sum_x \sum_y^M I_{c'}(x, y, r) \quad (8)$$

ここで、 r は RGB 表色系 $\mathcal{D} = \{R, G, B\}$ の各要素を表している。注目セル c' の代表値 $I'(c', r)$ と各ピクセルの色 $I(x, y, r)$ を用いて、CSS 特徴量と同様に式 (9) のユークリッド距離を用いて類似度 $s'(x, y, c')$ を算出する。

$$s'(x, y, c') = \sqrt{\sum_{r \in \mathcal{D}} (I'(c', r) - I(x, y, r))^2} \quad (9)$$

類似度 $s'(x, y, c')$ は注目セル c' ごとに求められる。 64×128 ピクセルの入力画像に対して、注目セル c' は $(64/M) \times (128/M)$ 個となる。そのため、 $M = 8$ の場合は各ピクセルに対して 128 パターンの類似度が求められる（本研究では注目セル c' に対する類似度画像と呼ぶ）。図 4(a) に、1 枚の検出対象から作成した類似度画像の例を示す。図 4(a) から分かるように、CSS 特徴量と同様に注目セル c' が物体領域に選択されている場合（セル 45）、色の類似度を利用して同一

領域らしさを表現できている．また，注目セル c' が背景領域に選択されている場合（セル 89），物体領域は類似度が低くなることが確認できる．図 4 (b) は，学習用画像 2,416 枚の検出対象から作成した平均類似度画像を表している．図 4 (b) から分かるように，検出対象全体で同一物体領域は類似度が高く，非同一物体領域は類似度が低く表現される傾向がある．CS-HOG 特徴量は，このように作成した類似度画像の特徴を捉えて物体形状を捉える．

[HSV 表色系の利用]

文献 [19] では，色の類似度を算出する際に利用する表色系として HSV 表色系が最も有効に働くと報告している．CSS 特徴量は，各要素を量子化してヒストグラム化することで HSV 表色系による距離計算を行っている．しかし，CS-HOG 特徴量における局所領域とピクセル間の色の類似度の算出には，前述の算出方法では HSV 表色系を利用して類似度を求めることができない．HSV 表色系は H （色相）， S （彩度）， V （明度）で構成されており，RGB 表色系とは違い各要素の原点が異なる．そのため，HSV 表色系の各要素を利用してユークリッド距離を計算することができない．そこで，本研究では HSV 表色系を利用して距離計算を行うため，各要素 $\{H, S, V\}$ を式 (10) により座標系 $\{u, t, V\}$ に変換する．

$$\begin{cases} u = S \times \cos H \\ t = S \times \sin H \\ V = V \end{cases} \quad (10)$$

これにより，RGB 表色系と同様に原点が統一の座標系として，HSV 表色系を表現することができる．CS-HOG 特徴量は，このように HSV 表色系を座標系 $\{u, r, V\}$ に変換してから距離計算を行い色の類似度を算出する．

3.2 色の類似度を用いた形状特徴の算出

CS-HOG 特徴量は図 5 に示すように，局所領域とピクセル間の色の類似度により作成した類似度画像を用いて形状特徴を観測する．形状特徴は，注目セル c' ごとに作成された類似度画像を利用して，HOG 特徴量同様に各ピクセルのこう配をセル c ごとにヒストグラム化したものを特徴量とする．算出手順は，はじめに注目セル c' より算出した類似度 $s'(x, y, c')$ を利用して，各ピクセルのこう配強度 m' とこう配方向 θ' を式 (11)~(13) より算出する．

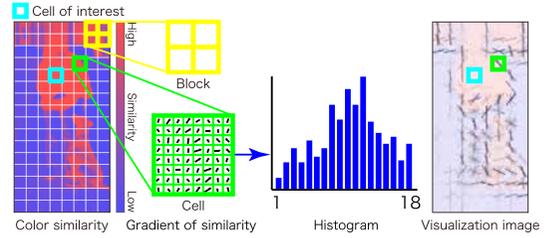


図 5 CS-HOG 特徴量

Fig. 5 Color Similarity-based HOG feature.

$$m'(x, y, c') = \sqrt{s'_x(x, y, c')^2 + s'_y(x, y, c')^2} \quad (11)$$

$$\theta'(x, y, c') = \tan^{-1} \frac{s'_x(x, y, c')}{s'_y(x, y, c')} \quad (12)$$

$$\begin{cases} s'_x(x, y, c') = s'(x+1, y, c') - s'(x-1, y, c') \\ s'_y(x, y, c') = s'(x, y+1, c') - s'(x, y-1, c') \end{cases} \quad (13)$$

HOG 特徴量は，輝度値の明暗関係によるこう配方向 θ' の反転を防ぐため， 180° より大きくなるこう配方向 θ を 180° 引いた方向で求めている．提案する CS-HOG 特徴量は，注目セル c' との同一物体らしさを色の類似度を利用して表現しているため， $0^\circ \sim 360^\circ$ 全てのこう配方向 θ' を用いて特徴表現をすることができる．次に，注目セル c' の類似度画像より算出したこう配強度 m' とこう配方向 θ' を用いて，式 (14) によりセル c ($M \times M$ ピクセル) におけるこう配方向ヒストグラム $\mathbf{V}'_c(c') = \{v'_c(1, c'), v'_c(2, c'), \dots, v'_c(N^{cshog}, c')\}$ を作成する．

$$v'_c(n, c') = \sum_x \sum_y m'(x, y, c') \delta[f(\theta'(x, y, c')), n] \quad (14)$$

ここで， $n = \{1, 2, \dots, N^{cshog}\}$ はヒストグラムのピンの番号， $f(\theta')$ はこう配方向 θ' を N^{cshog} 方向に量子化した値を表している．CS-HOG 特徴量では，こう配方向ヒストグラム $\mathbf{V}'_c(c')$ は，注目セル c' ごとに作成される．よって， 64×128 ピクセルの入力画像に対して $M = 8$ の場合，一つのセル c に対して 128 パターンのこう配方向ヒストグラム $\mathbf{V}'_c(c')$ が作成される．最後に，式 (15) を用いて注目セル c' ごとに，HOG 特徴量と同様にしてこう配方向ヒストグラム $\mathbf{V}'_c(c')$ を，複数のセルで構成されるブロック ($K \times K$ セル) 領域ごとに正規化する．

$$h'_c(n, c') = \frac{v'_c(n, c')}{\sqrt{\sum_{i=1}^{K \times K} \sum_{n=1}^{N^{cshog}} v'_i(n, c')^2 + \epsilon}} \quad (15)$$

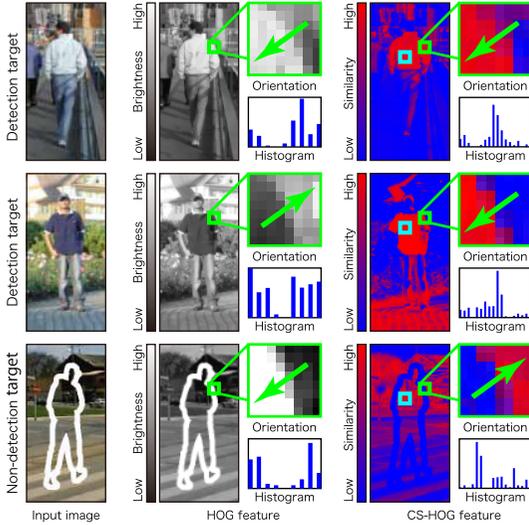


図 6 HOG 特徴量と CS-HOG 特徴量の比較

Fig. 6 Comparison of HOG feature and CS-HOG feature.

ここで、 ϵ は分母が 0 になることを防ぐための定数 ($\epsilon = 1$) である。正規化後のこの配方向ヒストグラム $H'_c(c')$ は、 $H'_c(c') = \{h'_c(1, c'), h'_c(2, c'), \dots, h'_c(K \times K \times N^{cshog}, c')\}$ となる。CS-HOG 特徴量は、この処理 (式 (11)~(15)) により、全ての注目セル c' から作成した類似度画像からこの配方向ヒストグラム $V'_c(c')$ を取得し、各量子化番号の値を特徴量とする。CS-HOG 特徴量は、 64×128 ピクセルの入力画像に対して、 $M = 8$, $K = 2$, $N^{cshog} = 18$ のとき、128 パターンの注目セル c' より作成した類似度画像と 105 個のブロック領域により正規化処理が行われる。そのため、 $128 \times 2 \times 2 \times 18 \times 105 = 967,680$ 次元となる。

3.3 HOG 特徴量と CS-HOG 特徴量の比較

HOG 特徴量と CS-HOG 特徴量の特徴表現の特性について比較する。HOG 特徴量と CS-HOG 特徴量を可視化して比較したものを図 6 に示す。図 6 の二つの検出対象から分かるように、HOG 特徴量は物体領域と背景領域の輝度の明暗関係によるこの配方向を防ぐため、この配方向を $0^\circ \sim 180^\circ$ に制限している。この配方向ヒストグラム (図中のヒストグラム) から確認できるように、この配方向を制限させることで検出対象で統一した物体形状を捉えることができている。一方、CS-HOG 特徴量は同一物体らしさを色の類似度を利用して表現しているため、この配方向を制限する必要がなく全てのこの配方向を表現できる。そ

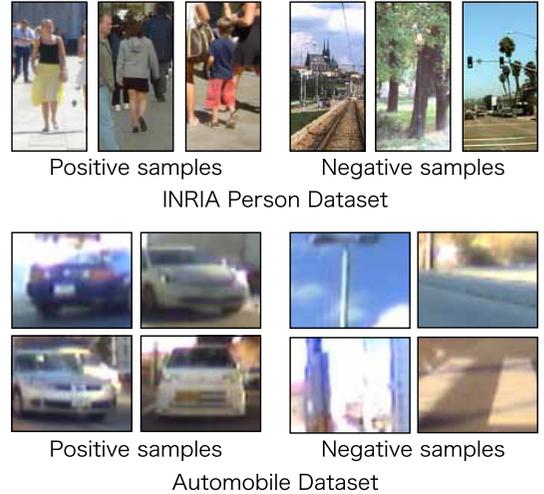


図 7 データセットの例
Fig. 7 Examples of datasets.

のため、HOG 特徴量と比較して特徴表現が多様となり、物体形状を効果的に捉えることができると考えられる。図 6 の非検出対象は、検出対象に類似した形状を人為的に背景へ加えた画像を示している。この場合、HOG 特徴量は局所領域の形状のみを観測するため、ヒストグラムからも分かるように、検出対象と似た形状を捉えてしまうことで誤検出となるおそれがある。しかし、CS-HOG 特徴量は注目セルとして選択した局所領域を利用して、同一物体らしさから物体領域を捉えることで、このような影響を抑制できる効果がある。このように、CS-HOG 特徴量は色の類似性に基づいて形状特徴を求めることで、多様な特徴表現をして効果的に物体形状を捉えることができることから、高精度な物体検出が期待できる。

4. 評価実験

CS-HOG 特徴量による特徴表現の有効性を確認するため評価実験を行う。

4.1 データセット

評価実験には、あらかじめ物体の領域を切り出したデータセットを用いる。データセットには、人検出のベンチマークとして広く利用されている INRIA Person Dataset [8] と車両のデータセットを用いる。図 7 に各データセットの検出対象 (ポジティブサンプル)、非検出対象 (ネガティブサンプル) の例を示す。INRIA Person Dataset は、学習サンプルにはポジティブサンプル 2,416 枚、ネガティブサンプル 13,161 枚を使用

し、評価サンプルはポジティブサンプル 1,126 枚、ネガティブサンプル 15,282 枚を使用する。また、学習サンプル、評価サンプルは 64×128 ピクセルの大きさに正規化して使用する。車両のデータセットは、学習サンプルにはポジティブサンプル 2,466 枚、ネガティブサンプル 5,035 枚を使用し、評価サンプルはポジティブサンプル 1,900 枚、ネガティブサンプル 5,153 枚を使用する。また、学習サンプル、評価サンプルは 80×64 ピクセルの大きさに正規化して使用する。

4.2 実験概要

CS-HOG 特徴量の有効性を示すために、以下の局所特徴量を用いて識別精度を比較する。

- HOG 特徴量 (HOG)
- CSS 特徴量 (CSS)
- CS-HOG 特徴量:RGB 表色系 (CS-HOG:RGB)
- CS-HOG 特徴量:HSV 表色系 (CS-HOG:HSV)

CS-HOG 特徴量は、色の類似度に RGB 表色系を利用したものと HSV 表色系を利用したものをを用いる。各特徴量のパラメータは、セルサイズ $M = 8$ ピクセル、ブロックサイズ $K = 2$ セル、量子化数は $N^{hog} = 9$, $N^{css} = 3$, $N^{cshog} = 18$ とした。CS-HOG 特徴量と比較手法は AdaBoost [20] の改良である Real AdaBoost [21] を利用して識別器を学習する。また、Real AdaBoost により学習する弱識別器は 500 個で統一して評価実験を行う。評価には Detection Error Trade-off (DET) カーブを用いる。DET カーブは横軸に誤検出率、縦軸に未検出率を表しており、原点に近いほど高精度であることを表す。

4.3 実験環境

実験環境として、OS:Windows Server(R) 2008 Enterprise x64, CPU:Intel(R) Xeon(R) CPU X7542 @ 2.67 GHz, RAM:256 GByte のスペックのコンピュータを用いる。CS-HOG 特徴量は HOG 特徴量及び CSS 特徴量と比較して膨大な特徴次元数であるため、学習に約 1 週間程度の時間を要する。

4.4 実験結果

各データセットによる実験結果の DET カーブを図 8(a) に示す。DET カーブより、INRIA Person Dataset では誤検出率 0.05% のとき CS-HOG:RGB は HOG, CSS と比較して約 5.7%, 8.1% 識別精度が向上している。CS-HOG:HSV は HOG, CSS と比較して約 22.5%, 27.2% 識別精度が向上しており、CS-HOG:RGB と比較して約 16.8% 識別精度が向上している。これにより、人検出において CS-HOG 特

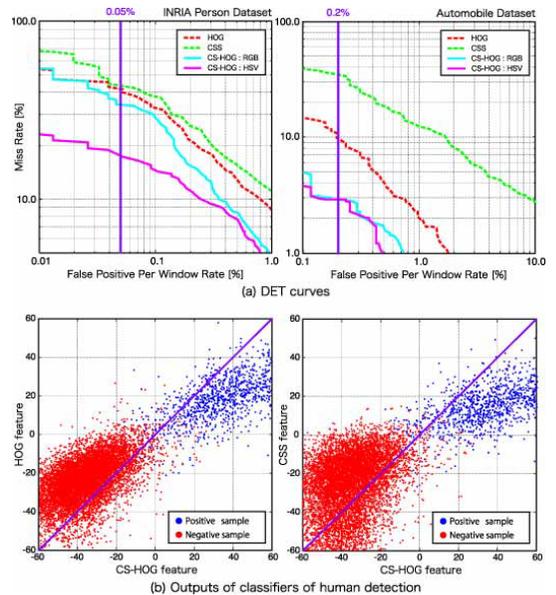


図 8 実験結果

Fig. 8 Experiment results.

徴量は HSV 表色系を利用して類似度計算をすることで、有効な特徴表現を可能にしたことが確認できた。また、車両のデータセットでは誤検出率 0.2% のとき CS-HOG:RGB 及び CS-HOG:HSV は HOG, CSS と比較して約 6.5%, 23.8% 識別精度が向上している。車両検出においては CS-HOG 特徴量に使用する表色系は、RGB と HSV どちらを用いても識別精度に大きな差は見られなかった。

人検出結果を用いて CS-HOG 特徴量と HOG 特徴量、CSS 特徴量で識別器の出力値にどのような変化があるかを確認する。図 8 (b) に評価サンプルにおける識別器の出力値の分布を示す。図 8 (b) は、横軸に CS-HOG 特徴量を用いた識別器の出力値、縦軸に HOG 特徴量または CSS 特徴量を用いた識別器の出力値を示し、対角線上にサンプルが分布すると同一の出力値であることを表す。識別器の出力値の分布から分かるように、CS-HOG 特徴量は HOG 特徴量及び CSS 特徴量と比較して、ポジティブサンプルの約 86.7%, 88.1% が出力値が高い値となっている。また、ネガティブサンプルの約 67.2%, 86.7% が出力値が低い値となっている。これにより、CS-HOG 特徴量は HOG 特徴量、CSS 特徴量と比較して、各サンプルがより正解クラスらしい出力となることが確認できた。よって、図 9 に示す人検出例からも分かるように、



図9 人検出例

Fig.9 Examples for human detection.

CS-HOG 特徴量は色の類似性に基づいて形状を観測し、物体の特徴を効果的に捉えることで高精度な物体検出を実現した。評価には Real AdaBoost の学習アルゴリズムにより選択された 500 次元の特徴量のみを使用するため、CS-HOG 特徴量を用いた識別器の処理時間は HOG 特徴量、CSS 特徴量と比較して大幅に増加しない。今回の評価実験に使用する INRIA Person Dataset の場合、評価サンプルに対する平均処理時間は HOG 特徴量が約 2ms なのに対して、CS-HOG 特徴量は約 7ms で識別が可能である。

5. む す び

本研究では色の類似性に基づいて形状を観測する CS-HOG 特徴量を提案した。CS-HOG 特徴量は、カラー画像の色の類似性に基づいて形状を算出するこ

とで、HOG 特徴量に比べて効果的に物体形状を捉えることを可能にした。人及び車両検出における評価実験により、CS-HOG 特徴量は従来の局所特徴量と比較して、物体の特徴を効果的に表現し高精度な検出を実現した。また、人検出において CS-HOG 特徴量を用いることにより、従来の局所特徴量と比較してより正解クラスらしい出力を算出することを可能にした。今後は、物体検出以外の画像認識の問題に対して、CS-HOG 特徴量による色の類似性に基づいた形状特徴の有効性を調査する予定である。

文 献

- [1] P. Viola and M. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol.57, pp.137-154, 2004.
- [2] N. Dalal, B. Triggs, and C. Schmid, "Human detection using oriented histograms of flow and appearance," *European Conference on Computer Vision*,

- 2006.
- [3] Y.D. Mu, S.C. Yan, Y. Liu, T. Hunag, and B.F. Zhou, "Discriminative local binary patterns for human detection in personal album," *Comput. Vis. Pattern Recognit.*, pp.1-8, 2008.
- [4] W.R. Schwartz, A. Kembhavi D. Harwood, and L.S. Davis, "Human detection using partial least squares analysis," *International Conference on Computer Vision*, 2009.
- [5] T. Watanabe, S. Ito, and K. Yokoi, "Co-occurrence histograms of oriented gradients for pedestrian detection," *Advances in Image and Video Technology*, pp.37-47, 2009.
- [6] X. Wang, T.X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," *International Conference on Computer Vision*, pp.32-39, 2009.
- [7] C. Wojek, S. Walk, S. Roth, and B. Schiele, "Monocular 3D scene understanding with explicit occlusion reasoning," *Comput. Vis. Pattern Recognit.*, pp.1933-2000, 2011.
- [8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *Comput. Vis. Pattern Recognit.*, vol.1, pp.886-893, 2005.
- [9] K. Levi and Y. Weiss, "Learning object detection from a small number of examples: The importance of good features," *Comput. Vis. Pattern Recognit.*, vol.2, pp.53-60, 2004.
- [10] B. Wu and R. Nevatia, "Detection of multiple, partially occluded humans in a single image by Bayesian combination of edgelet part detectors," *International Conference on Computer Vision*, vol.1, pp.90-97, 2005.
- [11] M. Isard and A. Blake, "Condensation conditional density propagation for visual tracking," *Int. J. Comput. Vis.*, vol.29, no.1, pp.5-28, 1998.
- [12] D. Comanisiu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," *Comput. Vis. Pattern Recognit.*, pp.142-149, 2000.
- [13] X. Wang, G. Doretto, T. Sebastian, J. Rittscher, and P. Tu, "Shape and appearance context modeling," *International Conference on Computer Vision*, pp.1-8, 2007.
- [14] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," *Comput. Vis. Pattern Recognit.*, pp.2360-2367, 2010.
- [15] A. Torralba, R. Fergus, and W.T. Freeman, "80 Million tiny images: A large data set for nonparametric object and scene recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.30, no.11, pp.1958-1970, 2008.
- [16] J. Shotton, M. Johnson, and R. Cipolla, "Semantic texton forests for image categorization and segmentation," *Comput. Vis. Pattern Recognit.*, pp.1-8, 2008.
- [17] V. Lepetit and P. Fua, "Keypoint recognition using randomized trees," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.28, no.9, pp.1465-1479, 2006.
- [18] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua, "Fast key-point recognition using random ferns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.32, no.3, pp.448-461, 2010.
- [19] S. Walk and N. Majer, "New features and insights for pedestrian detection," *Comput. Vis. Pattern Recognit.*, pp.1030-1037, 2010.
- [20] Y. Freund and R.E. Schapire, "Experiments with a new boosting algorithm," *International Conference on Machine Learning*, pp.148-156, 1996.
- [21] R.E. Schapire and Y. Singer, "Improved boosting algorithms using confidence-rated predictions," *Mach. Learn.*, vol.37, no.3, pp.297-336, 1999.

(平成 24 年 7 月 30 日受付, 25 年 2 月 5 日再受付)



後藤 雄飛 (学生員)

2011 中部大・工・情報工学卒。同年より同大学院修士課程に在籍。コンピュータビジョン、パターン認識に関する研究に従事。



山内 悠嗣 (正員)

2012 中部大学大学院博士後期課程了。2012 より同大学院博士研究員。2010 独立行政法人日本学術振興会特別研究員 DC。コンピュータビジョン、パターン認識の研究に従事。



藤吉 弘亘 (正員)

1997 中部大学大学院博士後期課程了。1997~2000 米カーネギーメロン大学ロボット工学研究所 Postdoctoral Fellow。2000 中部大学講師, 2004 同大准教授を経て 2010 より同大教授。2005~2006 米カーネギーメロン大学ロボット工学研究所客員研究員計算機視覚, 動画像処理, パターン認識・理解の研究に従事。2005 年度ロボカップ研究賞。2009 年度情報処理学会論文誌コンピュータビジョンとイメージメディア優秀論文賞, 2009 年度山下記念研究賞。情報処理学会, 電気学会, IEEE 各会員。