

Relational Binarized HOG 特徴量と Real AdaBoost による バイナリ選択を用いた物体検出

松島 千佳[†] 山内 悠嗣[†] 山下隆義^{†,††} 藤吉 弘亘[†]

[†] 中部大学 〒487-8501 愛知県春日井市松本町 1200

^{††} オムロン株式会社 〒525-0035 滋賀県草津市西草津 2 丁目 2-1

E-mail: [†] matsu, yuu@vision.cs.chubu.ac.jp, hf@cs.chubu.ac.jp, ^{††} takayosi@omm.ncl.omron.co.jp

あらまし 本稿では, Relational Binarized HOG 特徴量 (RB-HOG) と Real AdaBoost によるバイナリ選択を用いた物体検出法を提案する. HOG 特徴量は, 人検出に有効な特徴量であるが, 局所領域に着目しているため高次元な特徴量である. そこで, 本稿では特徴量の情報量を削減するために, 2 つの局所領域から抽出した HOG 特徴量の大小関係によりバイナリパターン化する RB-HOG 特徴量を提案する. これにより, 局所領域間の関係性を捉えたバイナリパターンを作成することができるが, RB-HOG 特徴量は識別に不必要なバイナリが含まれる. そこで, Real AdaBoost を用いて学習する際に, “0” と “1” の 2 つのバイナリを許容する “*” を導入することにより, 識別に有効なバイナリを選択する. 評価実験の結果より, 提案手法はメモリ量を削減したにもかかわらず, 従来法である HOG 特徴量の検出精度と同程度以上であることを確認した.

キーワード 物体検出, 2 値化, Histograms of Oriented Gradients, Real AdaBoost

Object Detection Using Relational Binarized HOG Feature and Binary Selection by Real AdaBoost

Chika MATSUSHIMA[†], Yuji YAMAUCHI[†], Takayoshi YAMASHITA^{†,††}, and Hironobu FUJIYOSHI[†]

[†] Chubu University, 1200 Matsumoto, Kasugai, Aichi, 487-8501

^{††} OMRON Corporation, 2-2-1 Nishikusatsu, Kusatu, Shiga, 525-0035

E-mail: [†] matsu, yuu@vision.cs.chubu.ac.jp, hf@cs.chubu.ac.jp, ^{††} takayosi@omm.ncl.omron.co.jp

Abstract This paper proposes an object detection that uses a new type of feature called Relational Binarized HOG feature (RB-HOG), and binary selection with Real AdaBoost. HOG features are effective features for human detection, but their focus on local regions makes them high-dimension features. Therefore, to reduce the amount of feature dimension, this paper proposes a new type of feature called RB-HOG feature that creates binary patterns from the HOG features extracted from two local regions. This approach enables the created binary patterns to reflect the relationships between local regions. However, since RB-HOG features contain binary values not needed for classification, we have added a process to the Real AdaBoost learning algorithm that permits the two binary values of “0” and “1”, and selects valid binary values for classification. The results of our evaluation testing demonstrated that our method offered the same or better detection precision as the conventional method (HOG feature) despite managing to reduce the amount of memory.

Key words Object Detection, Binarization, Histograms of Oriented Gradients, Real AdaBoost

1. はじめに

近年, デジタルカメラや車載カメラの普及に伴い, 撮影画像の画質改善や運転者のアシストを実現するために人検出の実用化の要望が高まり, FPGA 等によるハードウェア化の研究が行われている [1] ~ [5]. ハードウェア

化する上で重要な点は, 高い精度を維持したまま高速かつ低メモリで動作することである.

近年, 提案されている人検出法は, 画像局所特徴量と統計的学習手法の組み合わせによる手法 [6] ~ [10] が多い. 特徴量としては, 局所領域の勾配に着目した特徴量が多数提案されている [11], [12]. これらの特徴量は, 物

体の形状を捉えることが可能であるが、局所領域毎に特徴量を求めるため非常に高次元、かつ実数値となる。これは実装メモリが少ない小規模なハードウェアでは大きな問題となる。この問題を解決するためには、特徴量の情報量を削減する必要がある。情報量の削減は2つの利点がある。1つ目はメモリ量の削減、2つ目は共通の性質を表現する特徴量を類型化できることである。

情報量を削減する手法には、特徴量の空間を圧縮して特徴量数を削減する手法と個々の特徴量の情報量自体を削減する2つの手法が考えられる。前者のアプローチは、特徴量数を削減するベクトル量子化[13]や特徴量の次元圧縮を行う主成分分析などがある。これらの手法は、元の情報量を維持しながら特徴量の次元数を削減することができる。しかし、人検出では膨大な数の検出ウィンドウを処理するため、計算コストが高いこれらの手法では非効率である。

後者のアプローチは、特徴量を低ビットに量子化する手法である。例えば、スカラー量子化は問題に応じてビット数を設定し、必要となる情報量を表現することが可能である。また、2値化も有効な方法の一つであり、最も少ないビット数で表現する他、ノイズに対する頑健性や扱いやすいという利点がある。その一手法として閾値処理による2値化が考えられ、簡単に処理することが可能であるため、計算コストが低いという利点がある。しかし、多数のサンプルに対して最適な閾値を一気に決定することは難しい。この他の2値化手法として大小関係を用いた2値化がある。Ojalaらにより提案されたLocal Binary Pattern(LBP)[14]を用いた手法[15],[16]や、これを発展させた手法[17]は、ある2つの値の大小関係から2値化するため、閾値が不要というメリットがある。また、閾値処理による2値化と大小関係を用いた2値化では、一つのバイナリが含む情報量が異なる。閾値処理では値の大きさのみを表現するのに対して、大小関係を用いた2値化は、さらに2つの値の関係性の情報を含むことができる。

本研究では後者の情報量を削減する手法の一つである大小関係を用いた2値化に着目する。特徴量の情報量を削減しつつ高精度な人検出を実現するために、2つの局所領域から抽出したHOG特徴量の大小関係を用いたRelational Binarized HOG特徴量(RB-HOG)を提案する。RB-HOG特徴量は、2つのHOG特徴量の大小関係を用いるため閾値が不要であり、局所領域間の関係性を表現することが可能である。しかし、RB-HOG特徴量は複数のバイナリを組合せた特徴量であるため、識別に不必要なバイナリが含まれる。そこで、学習する際に“0”と“1”の2つのバイナリを許容する“*”を導入する。これにより、Real AdaBoostを用いて識別に有効なバイナリを選択する。

2. HOG特徴量と2値化

本章では、本研究で用いるHOG特徴量とHOG特徴量の情報量を削減する2値化について述べる。

2.1 HOG特徴量

Dalalらによって提案されたHistograms of Oriented Gradients特徴量(HOG)[6]は、局所領域における輝度の勾配方向をヒストグラム化した特徴量であり、物体の形状を捉えることが可能である。この特徴量は、近接画素の勾配を局所領域ごとに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。

はじめに、入力画像をあらかじめ定義したパッチサイズにダウンサンプリングし、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 を用いて、式(4)よりセル領域 c ($p \times p$ ピクセル)における量子化勾配方向 θ' の勾配強度の和を算出する。

$$v_c^H(\theta') = \sum_x \sum_y m(x, y) \delta[\theta', \theta(x, y)] \quad (4)$$

δ はKroneckerのデルタ関数であり、勾配方向 $\theta(x, y)$ が量子化勾配方向 θ' と同じヒストグラムの要素である場合は1を返す。式(4)より算出した量子化勾配方向 θ' における勾配強度の和の集合を、 N 方向の勾配方向ヒストグラム $V_c^H = \{v_c^H(1), v_c^H(2), \dots, v_c^H(N)\}$ とする。最後に、式(5)を用いてブロック領域($q \times q$ セル)ごとに正規化し、特徴量を抽出する。

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

ブロック領域は、1セルごとに正規化対象のセル領域が重なるように移動しながら正規化する。正規化後の勾配方向ヒストグラム V_c^H は、 $V_c^H = \{v_c^H(1), v_c^H(2), \dots, v_c^H(B \times N)\}$ となる。ここで、 B はブロック領域に含まれるセル領域の数である。

2.2 Binarized HOG特徴量

抽出したHOG特徴量の情報量を削減するために、閾値処理によりHOG特徴量を2値化したBinarized HOG特徴量(B-HOG)を得る。この特徴量は、セル領域における N 方向分のバイナリを1つの特徴量(バイナリパ

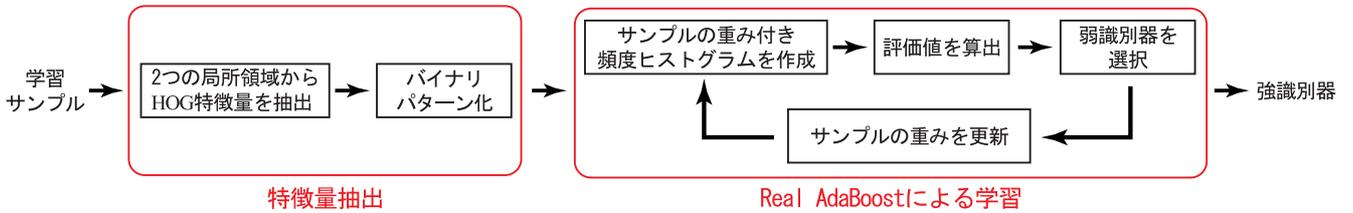


図 2 提案手法による学習の流れ

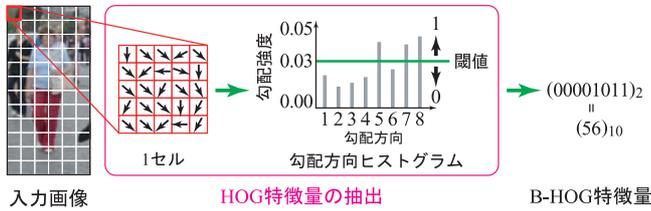


図 1 B-HOG 特微量の算出法

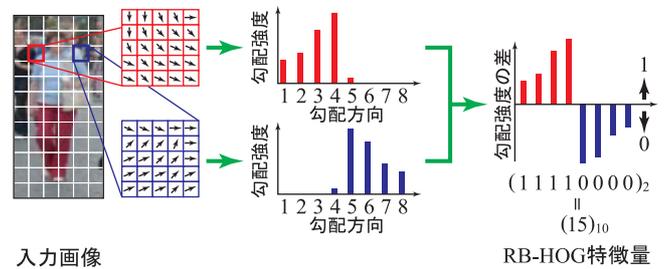


図 3 2つのセル領域の HOG 特微量を用いた 2 値化

ターン)として観測することにより、セル領域における勾配方向の関係性を捉える。

B-HOG 特微量は、セル領域における 8 方向の勾配方向ヒストグラム V_c^H を式 (6) に示すように閾値処理することにより、2 値化した特微量 $B_c^{BH} = \{b_c^{BH}(1), b_c^{BH}(2), \dots, b_c^{BH}(n), \dots, b_c^{BH}(8)\}$ となる。文献 [6] では、9 方向の量子化勾配方向を用いているが、本研究では特微量を 1 バイトで表現するために量子化勾配方向を 8 とする。

$$b_c^{BH}(n) = \begin{cases} 1 & \text{if } v_c^H(n) \geq th \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

ここで、 th は閾値を表す。例えば、図 1 のような入力画像より得られる HOG 特微量を 2 値化した場合、 $B_c^{BH} = (00001011)_2$ となる。

2.3 B-HOG 特微量の利点と問題点

HOG 特微量と B-HOG 特微量には、特微量に含まれる情報量に違いがある。式 (5) により求めた HOG 特微量は、浮動小数点型 (8 バイト) で表現しなければならないが、B-HOG 特微量は符号なし整数型 (1 バイト) で表現できる。そのため、B-HOG 特微量は、HOG 特微量よりメモリ量を 1/8 に削減することができる。しかし、人検出を行う環境が異なる場合、その環境に合わせた最適な閾値 th を求めなければならないという問題がある。

3. 提案手法

本章では、提案する Relational Binarized HOG 特微量 (RB-HOG) と Real AdaBoost を用いた学習について述べる。提案手法による学習の流れを図 2 に示す。2.3 節に示した 2 値化の利点を維持したまま問題点を解決するために、まず、2 つの局所領域から抽出した HOG 特微量の大小関係を用いて 2 値化した RB-HOG 特微量を抽出する。次に、統計的学習手法である Real AdaBoost

を用いて学習する。RB-HOG 特微量は、局所領域間の関係性を捉えることができるが、識別に不必要なバイナリが含まれる。そこで、学習する際に“0”と“1”の 2 つのバイナリを許容する“*”を導入することにより、識別に有効なバイナリを Real AdaBoost により選択する。

3.1 Relational Binarized HOG 特微量

RB-HOG 特微量は、2 つの局所領域から得られる HOG 特微量の大小関係を用いることにより 2 値化する。そのため、RB-HOG 特微量を求める際には閾値が必要ない。また、B-HOG 特微量は、一つのバイナリで勾配の強さしか表現できないのに対して、RB-HOG 特微量はこの他に 2 つの特微量の関係性も表現することができる。提案手法である RB-HOG 特微量は HOG 特微量の大小関係により 2 値化するため、照明変動に対して一定の不変性を持つ。そのため、提案手法では、照明変動に対して頑健性を得る HOG 特微量の正規化処理を必要としない。正規化処理は、HOG 特微量を計算する上で最も高い計算コストを占めているため、提案手法は処理コストを大幅に削減することが可能である。

3.1.1 2つのセル領域の HOG 特微量を用いた 2 値化

RB-HOG 特微量は式 (7) に示すように、2 つのセル領域から得られる 8 方向の勾配方向ヒストグラム $V_{c_1}^H, V_{c_2}^H$ の大小関係を用いて比較し、2 値化した特微量 $B_{c_1 c_2}^{RBH} = \{b_{c_1 c_2}^{RBH}(1), b_{c_1 c_2}^{RBH}(2), \dots, b_{c_1 c_2}^{RBH}(n), \dots, b_{c_1 c_2}^{RBH}(8)\}$ である。

$$b_{c_1, c_2}^{RBH}(n) = \begin{cases} 1 & \text{if } v_{c_1}^H(n) \geq v_{c_2}^H(n) \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

図 3 に示すように、2 つのセル領域における特微量の大小関係を用いることにより、局所領域間の関係性を捉えるバイナリパターンを作成できる。このとき、2 値化ではなく多値化することにより、より多くの情報を表現す

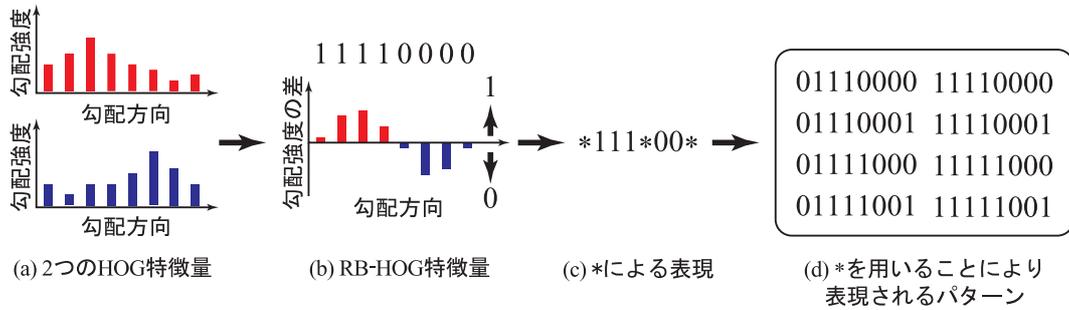


図5 “*”導入したバイナリパターンが表現するバイナリパターンの例

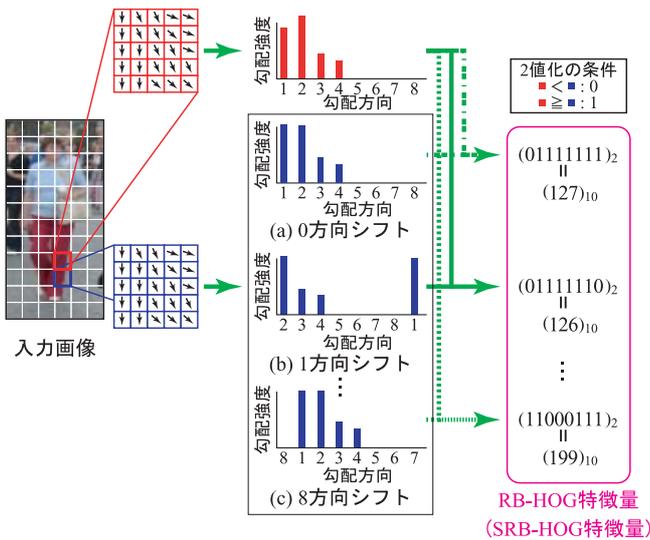


図4 勾配方向のシフトを導入

ることが可能となるが、最適な値の選定法や多値化する手法が必要となる。また、多値化することにより必要なメモリ量が増加するため、本稿では多値化ではなく2値化のみを対象とする。

図4に示すように、抽出した特徴量が類似している場合、比較する特徴量の大小関係が曖昧である。そのため、明確に2値で表現することが困難となる場合がある。

3.1.2 勾配方向のシフトを導入

RB-HOG 特徴量の問題点を解決するために、図4の(b)や(c)に示すように、一方のセル領域から抽出される8方向の勾配方向ヒストグラム $V_{c_2}^H$ の勾配方向を s 方向分シフトし ($s = 1, 2, \dots, 8$)、8通りの勾配方向ヒストグラム $V_{c_2s}^H$ を作成する。その後、式(7)と同様に、式(8)より大小関係を比較して2値化した8通りの特徴量 $B_{c_1c_2s}^{SRBH}$ を算出する。

$$b_{c_1c_2s}^{SRBH}(n) = \begin{cases} 1 & \text{if } v_{c_1}^H(n) \geq v_{c_2s}^H(n) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

勾配方向をシフトしたヒストグラムとの大小関係を算出することにより、抽出した特徴量が類似している場合でも、大小関係を明確に表現できる。本稿では、勾配方向をシフトしてRB-HOG 特徴量を抽出したものを Shifted Relational Binarized HOG 特徴量 (SRB-HOG) と呼ぶ。

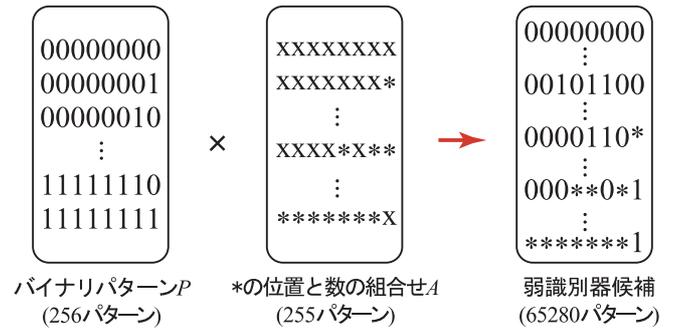


図6 一組のセル領域に対する弱識別器候補

3.2 “*”による識別に有効なバイナリ選択

RB-HOG 特徴量を抽出した後、統計的学習手法である Real AdaBoost [18] を用いて学習する。このとき、バイナリパターン化した特徴量に “*” を導入して識別に有効なバイナリを選択することにより、検出精度の向上が期待できる。

3.2.1 “*”の導入

提案する RB-HOG 特徴量は、図5(a)に示すような2つの HOG 特徴量を抽出した場合、図5(b)のようなバイナリパターンが得られる。しかし、比較する勾配強度の差が小さい場合、バイナリが容易に反転する可能性がある。このようなバイナリは信頼性が低いため、誤識別の原因となる。そこで、本稿では図5(c)に示すように、“0”と“1”の2つのバイナリを許容する “*” を導入したバイナリ選択法を提案する。これより、図5(d)のように、類似した複数のバイナリパターンを同時に表現することが可能となる。“*”の数とどのバイナリに適用するかは Real AdaBoost により選択する。そこで、各セル領域の組合せに対して図6に示すような “*” を導入した 65,280 パターンの弱識別器候補を作成する。提案手法の弱識別器候補の数 M は式(9)となる。

$$M = R \times P \times A \quad (9)$$

ここで、 R は局所領域 (セル) c の組合わせ数、 P はバイナリパターン数 ($2^8 = 256$)、 A は “*” の数 ($2^8 = 256$) である。これらの弱識別器候補を用いて Real AdaBoost により学習する。本稿では、全ての局所領域を組み合わせる弱識別器候補としている。弱識別器候補を限定する手法と

アルゴリズム 1: 学習アルゴリズム.

(a) 入力: I 個の学習サンプルにラベル付け $\{x_i, y_i\}_{i=1 \dots I}$,
 $y_i \in \{-1, 1\}$

(b) 初期化: サンプルの重みを初期化

$$D_1(i) = 1/I \quad (10)$$

(c) 学習:

For $t = 1, \dots, T$ // 学習回数

For $m = 1, \dots, M$ // “*” を導入したバイナリパターン

・ サンプルの重み付き頻度 W_+, W_- を算出

$$W_+ = \sum_{i: y_i = +1} D_t(i) \delta[F(x_i), m] \quad (11)$$

$$W_- = \sum_{i: y_i = -1} D_t(i) \delta[F(x_i), m] \quad (12)$$

・ 評価値 Z を算出

$$Z_m = |W_+ - W_-| \quad (13)$$

・ 評価値 Z が最大となるバイナリパターン m を
弱識別器 h_t として選択

$$h_t = \arg \max_{m \in M} Z_m \quad (14)$$

・ サンプルの重みを更新

$$D_{t+1}(i) = D_t(i) \exp[-y_i h_t(x_i)] \quad (15)$$

$$h(x_i) = \begin{cases} \frac{1}{2} \ln \frac{W_+ + \epsilon}{W_- + \epsilon} & \text{if } F(x_i) = m \\ \frac{1}{2} \ln \frac{(1 - W_+) + \epsilon}{(1 - W_-) + \epsilon} & \text{otherwise} \end{cases} \quad (16)$$

(d) 出力: 強識別器

$$H(x) = \text{sign} \left[\sum_{t=1}^T h_t(x) \right] \quad (17)$$

して、三田らのように Sequential forward selection(SFS)を用いる手法 [19] や、山内らのように Boosting により組み合わせる特徴量を限定する手法 [20] などが利用できる。また、バイナリパターン化した特徴量の全てのバイナリが識別に有効であるとは限らないため、“*” の数を制限することにより学習時間を短縮することができる。

3.2.2 学習手法

提案する学習手法は、3.2.1 にて述べた多数の弱識別器候補の中から RealAdaBoost により識別に有効な弱識別器を選択する。学習の流れについて説明する。

はじめに、前処理としてアルゴリズム 1(a) においてラベルの付いた I 枚の学習サンプル $(x_1, y_1), \dots, (x_I, y_I)$ を用意する。 x_i は画像であり、 y_i はクラスラベルである。検出対象のクラスは $y_i = +1$ 、非検出対象のクラスは $y_i = -1$ とする。また、各サンプルの重みを $D_t(i)$ とし、式 (10) により初期化する。

次に、Real AdaBoost により強識別器を学習をする。まず、学習サンプルの重み付き頻度 W_+, W_- を算出する。 $+$ は検出対象のクラス、 $-$ は非検出対象のクラス、 m は “*” を導入したバイナリパターンを表す。 W_+, W_- は、 m が各クラスの学習サンプルから抽出したバイナリ

パターンと一致する頻度であり、式 (11), (12) より算出する。ここで、 $F(x)$ は入力画像 x からバイナリパターンを観測する関数である。 δ は Kronecker のデルタ関数であり、入力する 2 つのバイナリパターンが一致する場合は 1 を返す。

学習サンプルの重み付き頻度 W_+, W_- を算出後、式 (13) より分離度を表す評価値 Z を算出する。評価値 Z は大きいほど、ポジティブサンプルとネガティブサンプルの頻度の差が大きいことを意味する。そのため、式 (14) より全バイナリパターン M から、評価値 Z が最大となるバイナリパターン m を、 t ラウンド目において最も識別に有効な弱識別器 h_t として選択する。

弱識別器を選択後、誤識別した学習サンプルが次のラウンドで正しく識別されるために、式 (15) より学習サンプルの重みを更新する。このとき、選択された弱識別器のサンプルの重み付き頻度 W_+, W_- を用いて、式 (14) より弱識別器の出力 $h(x)$ を算出する。ここで、 ϵ は分母が 0 になることを防ぐ係数 ($\epsilon = 1/I$) である。

最後に、サンプルの重み付き頻度の算出から学習サンプルの重み更新までの処理を一定の学習回数、または一定の識別率が得られるまで繰り返し、式 (17) に示す強識別器 $H(x)$ を得る。

4. 評価実験

提案手法の有効性を評価するために、2 つの評価実験をする。1 つ目は RB-HOG 特徴量の有効性を確認するために、HOG 特徴量や B-HOG 特徴量と精度を比較する。2 つ目は “*” を導入することの有効性を確認するために、導入前と導入後の精度を比較する。

4.1 データセット

実験では、姿勢の変化による形状の変化が大きい人と、剛体に近いとされる車両のデータセットを用いて実験する。人のデータセットは INRIA person dataset [6]、車両のデータセットは文献 [21] で用いられているデータセットを用いる。各データセットの一部を図 7 に示す。INRIA person dataset の学習用サンプルはポジティブサンプル 2,415 枚、ネガティブサンプル 12,180 枚であり、評価用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 453 枚である。評価用のネガティブサンプルは、人が存在しない 1 枚の背景画像であり、検出ウィンドウをラストスキャンして評価に用いる。車両のデータセットの学習用サンプルはポジティブサンプル 710 枚、ネガティブサンプル 8,800 枚であり、評価用サンプルはポジティブサンプル 1,230 枚、ネガティブサンプル 3,880 枚である。

4.2 評価実験の概要

評価実験では、下記に示す方法を比較する。

- HOG 特徴量 (HOG)

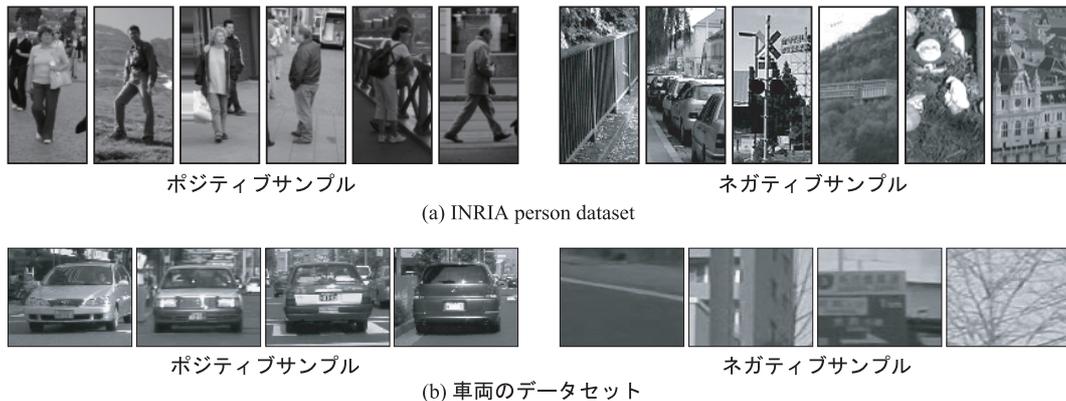


図 7 学習に用いるデータセットの一部

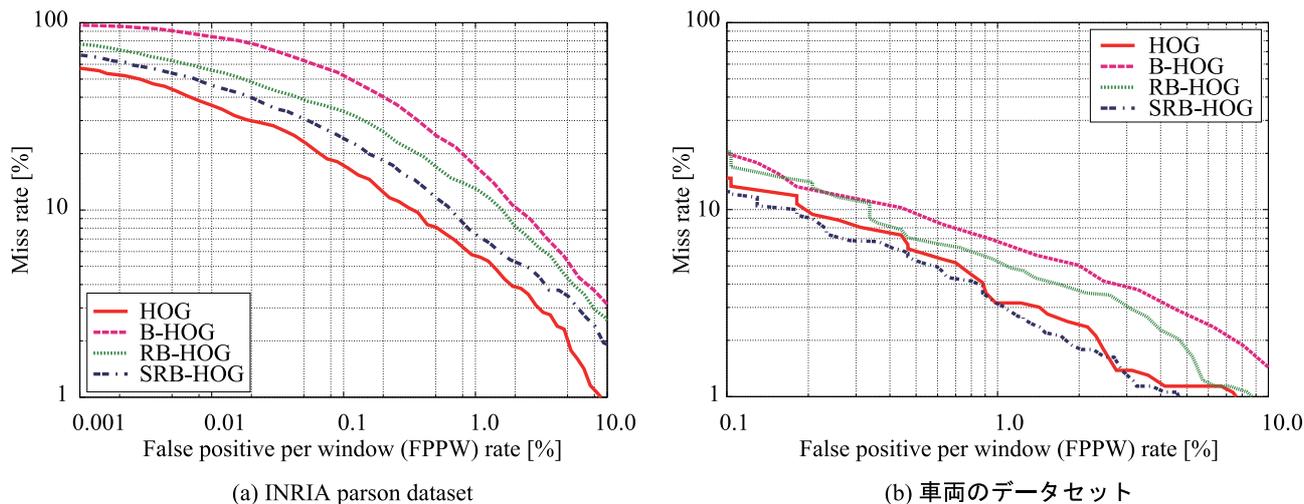


図 8 RB-HOG 特徴量の有効性の実験結果

表 1 実験に使用するデータベースのパラメータ

データセット	画像サイズ [pix.]	セル [pix.]	ブロック [セル]	勾配 方向数
INRIA person dataset [6]	64 × 128	8	2	8
車両データセット [21]	72 × 54	9	2	8

- Binarized HOG 特徴量 (B-HOG)
- Relational Binarized HOG 特徴量 (RB-HOG)
- RB-HOG 特徴量 + 勾配方向のシフト有 (SRB-HOG)

各データセットのパラメータを表 1 に示す。評価には、Detection Error Tradeoff (DET) カーブを用いる。DET カーブは、横軸に False Positive Per Window (FPPW)、縦軸に miss rate を表し、左下の原点に近いほど検出精度が高いことを示す。また、B-HOG 特徴量を算出する際に用いる閾値 th は、予備実験にて決定し、INRIA person dataset の場合は $th = 0.09$ 、車両データセットの場合は $th = 0.10$ を用いる。

4.3 実験 1: RB-HOG 特徴量の有効性

実験 1 では、提案手法である SRB-HOG 特徴量の有効性を確認する。各データセットの実験結果の DET カーブを図 8 に示す。

まず、B-HOG 特徴量と RB-HOG 特徴量を比較する。図 8(a) より FPPW が 1.0% のときの検出率を比較すると、RB-HOG 特徴量の検出率が B-HOG 特徴量より約 4.6% 向上した。図 8(b) より FPPW が 1.0% のときの検出率を比較すると、RB-HOG 特徴量の検出率が B-HOG 特徴量より約 1.7% 向上した。

次に、RB-HOG 特徴量と SRB-HOG 特徴量を比較する。図 8(a) より FPPW が 1.0% のときの検出率を比較すると、SRB-HOG 特徴量の検出率が RB-HOG 特徴量より約 1.9% 向上した。図 8(b) より FPPW が 1.0% のときの検出率を比較すると、SRB-HOG 特徴量の検出率が RB-HOG 特徴量より約 2.1% 向上した。また、HOG 特徴量と比較するとほぼ同程度の検出精度であることがわかる。

4.4 実験 2: “*” によるバイナリ選択の有効性

実験 2 では、Real AdaBoost で識別器を学習する際に、“*” を導入して識別に有効なバイナリを選択することの有効性を確認する。

4.4.1 評価実験の結果

本実験では、実験 1 の結果より提案手法の中で最も検出精度がよい SPB-HOG 特徴量に着目し、学習時に “*”

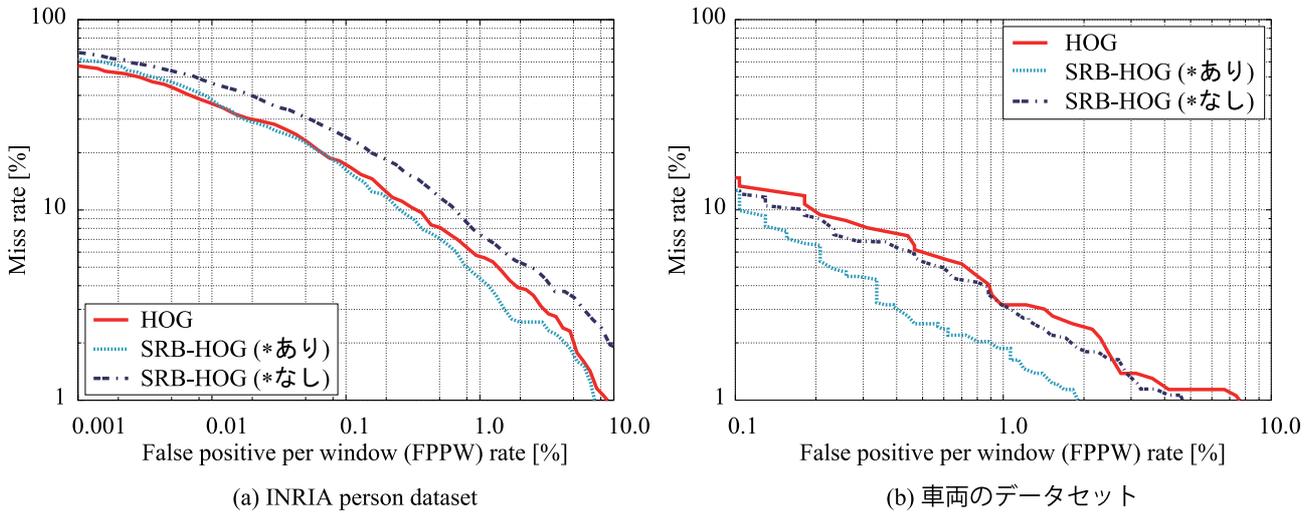


図9 “*”によるバイナリ選択の有効性の実験結果

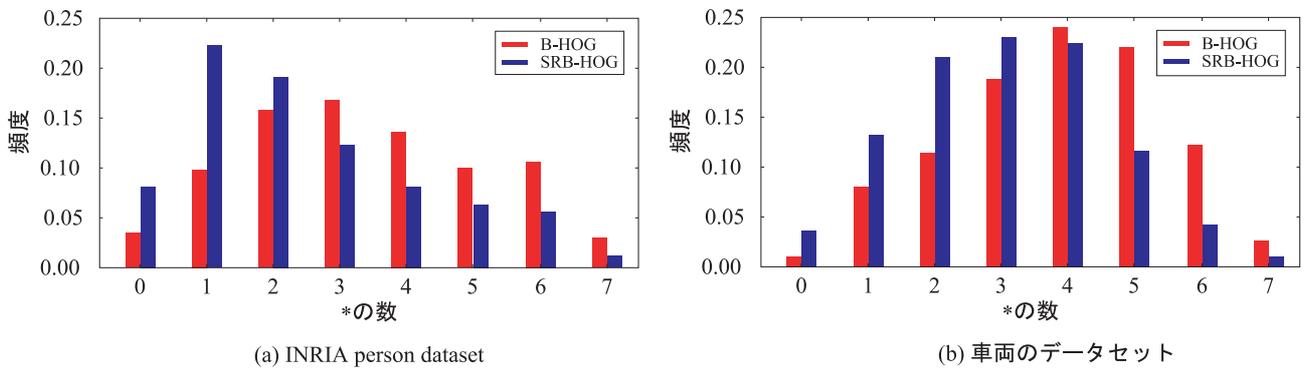


図10 選択された“*”の数の割合

を導入した手法と導入しない手法と比較する．各データセットの実験結果のDETカーブを図9に示す．

図9より，学習時に“*”を導入することにより精度が向上し，HOG特徴量と同程度以上の検出精度であることがわかる．図9(a)の人のデータセットでは，FPPWが1.0%における検出率を比較すると，“*”を導入したSRB-HOG特徴量は，“*”を導入しない手法と比較すると約3.0%向上し，HOG特徴量より約1.5%向上した．また，図9(b)の車両のデータセットでは，FPPWが1.0%における検出率を比較すると，“*”を導入したSRB-HOG特徴量は，“*”を導入しない手法とHOG特徴量より約1.2%向上した．これは，“*”を導入して反転しやすいバイナリを観測しないことにより，識別に有効なバイナリが選択できているからであると考えられる．

4.4.2 識別に有効なバイナリの考察

識別に有効なバイナリを考察するために，各データセットにおいて選択された“*”の数の割合を比較する．比較結果を図10に示す．図10より各特徴量の最も選択頻度が高い“*”の数は，人の場合，B-HOG特徴量が3個，SRB-HOG特徴量が1個である．また，車両の場合，B-HOG特徴量は4個，SRB-HOG特徴量は3個である．この結果より，検出対象が人及び車両の両方において，B-HOG特徴量よりもSRB-HOG特徴量の方が選択され

表2 メモリ量の比較

特徴量	HOG	B-HOG (BH)	RB-HOG (RBH)	SRB-HOG (SRBH)
メモリ量 [KB]	3.91	0.50	0.49	0.98

る“*”の数が少ないことがわかる．“*”は，“0”と“1”の2つのバイナリを許容するため，閾値との差や大小関係が曖昧な反転しやすいバイナリが“*”として表現される．SRB-HOG特徴量は，2つの局所領域から抽出したHOG特徴量の勾配方向をシフトしながら大小関係を算出する．そのため，大小関係が明確なバイナリが生成され，B-HOG特徴量よりも“*”の数が減少したといえる．

4.5 メモリ量の比較と処理時間

まず，特徴量の情報量を削減することにより，どの程度のメモリ量が削減できるかを確認するため，HOG特徴量と提案手法に必要なメモリ量を比較する．弱識別器数 T を500個としたときの，1検出ウィンドウ当たりのメモリ量を表2に示す．メモリ量 M の計算には次式を用いる．

$$M^{HOG} = T \times \text{Type}_d \quad (18)$$

$$M^{BH} = T \times \text{Type}_{uc} + th \quad (19)$$

$$M^{RBH} = T \times \text{Type}_{uc} \quad (20)$$

$$M^{SRBH} = T \times (\text{Type}_{uc} + S) \quad (21)$$

ここで、Type はデータの型を表し、 Type_d は浮動小数点型 (8 バイト)、 Type_{uc} は符号なし整数型 (1 バイト) である。また、 th は閾値であり浮動小数点型 (8 バイト)、 S はシフトする勾配方向数であり符号なし整数型 (1 バイト) である。

表 2 に各特徴量のメモリ量を示す。まず、B-HOG 特徴量と RB-HOG 特徴量を比較する。B-HOG 特徴量を抽出するためには閾値が必要であるが、RB-HOG 特徴量のメモリ量とほぼ同程度である。次に、RB-HOG 特徴量と SRB-HOG 特徴量を比較する。SRB-HOG 特徴量は、明確な大小関係を得るために勾配方向をシフトするため、RB-HOG 特徴量と比較すると約 2 倍のメモリ量が必要となる。最後に、HOG 特徴量と SRB-HOG 特徴量を比較する。SRB-HOG 特徴量は、大小関係を用いて HOG 特徴量の情報量を削減することにより、メモリ量を約 75.0%削減することができる。

次に、1 検出ウィンドウ (64×128 ピクセル) における特徴量の算出と識別に必要な処理時間について述べる。HOG 特徴量の処理時間は 5.39×10^{-7} [ms]、SRB-HOG 特徴量の処理時間は 2.70×10^{-7} [ms] である。提案手法は HOG 特徴量と比べ、約 50.0%の処理時間を削減することができた。これは、SRB-HOG 特徴量を算出する際に処理コストが高い HOG 特徴量の正規化処理をする必要がないからである。

5. おわりに

本稿では、2 つのセル領域から抽出した HOG 特等量の大小関係により 2 値化した Relational Binarized HOG 特徴量を用いた人検出を提案した。本稿の貢献は 2 点ある。1 つ目は、2 つのセル領域から抽出した HOG 特等量の大小関係により 2 値化した Relational Binarized HOG 特徴量である。この特徴量は、1 つバイナリで勾配強度の大きさと局所領域間の関係性を同時に捉えることができる。これにより、閾値が不要となり、1 つのバイナリが含む情報量が閾値処理よりも多くなる。

2 つ目は、“*” を導入することにより識別に有効なバイナリを Real AdaBoost で選択することである。勾配強度の差が小さい場合はバイナリが反転する可能性があり、識別精度の低下を引き起こす要因となる。そのため、“0” と “1” の 2 つのバイナリを許容する “*” を導入することにより、バイナリの反転を抑制し、特徴量の情報量を削減しても高精度な検出を実現した。

今後は、検出精度向上のために学習方法について検討する予定である。

文 献

[1] V. Nair, P. O. Laprise, and J. J. Clark, “An FPGA-based people detection system”, EURASIP Journal

on Applied Signal Processing, pp.1047–1061, 2005.

[2] A. Ess, B. Leibe, K. Schindler, and L. V. Gool, “A mobile vision system for robust multi-person tracking”, CVPR, pp.1–8, 2008.

[3] K. Khattab, J. Dubois, and J. Miteran, “Cascade boosting-based object detection from high-level description to hardware implementation”, EURASIP Journal on Embedded Systems, 2009.

[4] 門田亮二, 中村行宏, 宮本龍介, “実時間歩行者認識に向けた HOG 特徴抽出のハードウェア実装”, 電子情報通信学会技術研究報告, pp.43–48, 2009.

[5] 山中悠歩, 山崎俊彦, 相澤清晴, “FPGA による histogram of oriented gradients アルゴリズムの高速化”, 情報処理学会創立 50 周年記念全国大会, 2010.

[6] N. Dalal, and B. Triggs, “Histograms of oriented gradients for human detection”, CVPR, pp.886–893, 2005.

[7] B. Wu, and R. Nevatia, “Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors”, ICCV, pp.90–97, 2005.

[8] B. Leibe, E. Seemann, and B. Schiele, “Pedestrian detection in crowded scenes”, CVPR, pp.878–885, 2005.

[9] O. Tuzel, F. Porikli, and P. Meer, “Human detection via classification on riemannian manifolds”, CVPR, pp.1–8, 2007.

[10] 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, 2008.

[11] C. Hou, H. Ai, and S. Lao, “Multiview pedestrian detection based on vector boosting”, ACCV, pp.210–219, 2007.

[12] A. Bosch, A. Zisserman, and X. Munoz, “Representing shape with a spatial pyramid kernel”, Proceedings of ACM international conference on Image and video retrieval, pp.401–408, 2007.

[13] Y. Linde, A. Buzo, and R. Gray, “An algorithm for vector quantizer design”, IEEE Transactions on Communications, vol.28, pp.84–95, 1980.

[14] T. Ojala, M. P. Ainen, and D. Harwood, “A comparative study of texture measures with classification based on featured distributions”, Pattern Recognition, vol.29, pp.51–59, 1996.

[15] X. Wang, T. X. Han, and S. Yan, “An hog-lbp human detector with partial occlusion handling”, ICCV, pp.1–8, 2009.

[16] Y. Mu, S. Yan, Y. Liu, T. Huang, and B. Zhou, “Discriminative local binary patterns for human detection in personal album”, CVPR, pp.1–8, 2008.

[17] A. Hadid, M. Pietikainen, and T. Ahonen, “A discriminative feature space for detecting and recognizing faces”, CVPR, pp.797–804, 2004.

[18] R. E. Schapire, and Y. Singer, “Improved boosting algorithms using confidence-rated predictions”, Machine Learning, no.37, pp.297–336, 1999.

[19] 三田雄志, 金子敏充, 堀修, “顔検出に適した共起に基づく joint haar-like 特徴”, 電子情報通信学会論文誌, vol.J89-D, no.8, pp.1791–1801, 2006.

[20] 山内悠嗣, 山下隆義, 藤吉弘巨, “Boosting に基づく特徴量の共起表現による人検出”, 電子情報通信学会論文誌, vol.J92-D, no.8, pp.1125–1134, 2009.

[21] 高木雅成, 山内悠嗣, 三井相和, 藤吉弘巨, “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”, 画像の認識・理解シンポジウム (MIRU), pp.643–650, 2009.