

Boosting に基づく特徴量の共起表現による人検出

山内 悠嗣[†] 藤吉 弘亘[†] 山下隆義^{†,††}

[†] 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
^{††} オムロン株式会社 技術本部 〒619-0283 京都府相楽郡木津町木津川台 9-1
 E-mail: [†]{yuu,hf}@vision.cs.chubu.ac.jp, ^{††}takayosi@ari.ncl.omron.co.jp

あらまし 本稿では, Boosting に基づく特徴量の共起表現法とこれによる人検出法を提案する. 既に特徴量間の共起を表現する手法として AdaBoost により 2 値に識別した符号を複数組み合わせる手法が提案され, 顔検出においてその有効性が確認されている. しかし, 入力特徴がどちらのクラスとも言い難い場合にも 2 値に識別して共起を表現するため, 間違えた符号を組み合わせる問題がある. そこで, 提案手法では Real AdaBoost の弱識別器の出力である連続値と評価値を, 演算子によって結合した共起特徴から Real AdaBoost を用いて人と人以外に判別する. 共起特徴は, 複数の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な特徴量であるため, 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る. 評価実験より, 従来法と比較して誤検出率 5.0% において約 6.8% 検出率を向上させることができた.

キーワード 人検出, 共起, Real AdaBoost, 多重解像度, Histograms of Oriented Gradients

Human Detection by Boosting-based Co-occurrence Feature Representation

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

[†] Dept. of Computer Science, Chubu Univ. 1200, Matsumoto, Kasugai, Aichi, 487-8501 Japan
^{††} OMRON Corporation. 9-1, Kizugawadai, Kizu-cho, Soraku-gum, Kyoto, Japan 619-0283
 E-mail: [†]{yuu,hf}@vision.cs.chubu.ac.jp, ^{††}takayosi@ari.ncl.omron.co.jp

Abstract This paper presents a novel method for representing co-occurrence of low-level features and a method for detecting human using the co-occurrence representation. Our approach uses output of weak classifier and evaluated value of Real AdaBoost to obtain the co-occurrence feature. Therefore, our method can apprehend more than the conventional method of co-occurrence representation. Capturing the co-occurrence feature of Histograms of Oriented Gradients(HOG) feature makes it possible to construct an effective human detector. Experimental results show that our method has about 6.8% higher performance than that of the conventional method.

Key words Human Detection, Co-occurrence, Real AdaBoost, Multi Resolution, Histograms of Oriented Gradients

1. はじめに

画像中から自動的に人を検出する技術は, 監視システムや ITS など多岐に渡る分野で実現が期待され, 近年最も盛んに行われている研究のひとつである. 近年の人検出法は, Viola と Jones が提案した顔検出に代表されるように low-level な特徴と統計的学習による Boosting の組み合わせによる手法 [1] が多い. 人検出に有効な low-level な特徴量として, 領域の累積エッジ強度の比を特徴量とする Edge Orientation Histograms (EOH) [2] や, 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 Histograms of Oriented Gradients (HOG) [3], エッジ同士をつなげた短い線, カーブを特徴量として表現する Edgelet Feature [4] などがあり, 高い検出精度を達成している.

さらに, 検出精度の向上として, アピアランスとは異なる特徴量の併用 [5] [6] [7] や特徴量自体の改良 [8], Boosting 手法の改良 [9] などが提案されている. また, 高精度な検出には特徴量をどのように捉えるかが重要である. Mita *et al.* は, 複数の Haar-like 特徴を用いて 2 値に識別し, これらの符号を組み合わせる Joint Haar-like 特徴 [11] を提案した. この手法は, Viola と Jones の顔検出法 [1] と比較して, 大きく検出性能を向上さ

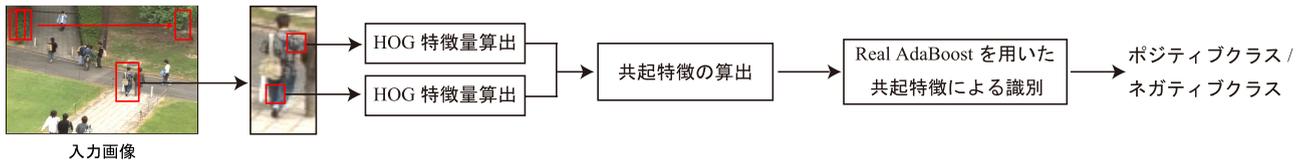


図2 提案手法の流れ

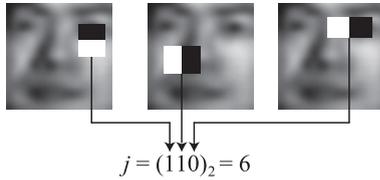


図1 Joint Haar-like 特徴 j

せた。しかし、弱識別器において選択された特徴量からはどちらのクラスとも言い難い場合にも2値に識別するため、間違えた符号を組み合わせる問題がある。

そこで、本研究では Real AdaBoost [10] の弱識別器の出力と評価値を演算子によって結合した新しい特徴量間の共起表現法と共起特徴による人検出法を提案する。提案する共起特徴は、複数の特徴量を演算子により多様な捉え方で観測することが可能であるため高精度な検出が期待できる。

本稿では、2.章において、従来の物体検出法と特徴量の共起表現法について説明した後、3.章にて low-level 特徴として HOG 特徴量を用いた共起特徴による人検出法について説明し、4.章では提案手法の有効性を示す。また、文献[7]にて、文献[11]の共起表現法を用いて異なる特徴量間の共起を表現したように、提案手法も同一の特徴量以外の共起表現が可能である。そこで、本研究では、異なる解像度間から得られる HOG 特徴量から共起の表現を行う。これについては、5.章にて説明する。

2. 従来の物体検出法

Viola と Jones の顔検出法 [1] が提案された以降、low-level な特徴量と統計的学習の組み合わせによる物体検出法が多く提案されている。提案されている人検出法としては、局所領域における勾配方向をヒストグラム化した Histograms of Oriented Gradients (HOG) と SVM を用いた人検出法 [3] がある。HOG 特徴量は照明変動に頑健であり、僅かな局所の変化ならば吸収できるため、高精度な人検出が可能である。近年では、検出性能を向上させるため、low-level 特徴を Boosting により組み合わせることで、識別に有効な mid-level な特徴量を生成する手法 [11][12] が提案されている。

Mita *et al.* は、AdaBoost によって選択された1個の弱識別器が、複数の Haar-like 特徴量を同時に観測することで共起を表現する Joint Haar-like 特徴 [11] を提案した。Joint Haar-like 特徴は、2値に識別した符号を組み合わせることにより表現され、組み合わせられた Joint Haar-like 特徴の同時確率に基づき識別を行う。図1に3つの符号を組み合わせる Joint Haar-like 特徴を算出する例を示す。この手法は、Viola と Jones の顔検出法と比較して、高精度かつ高速に識別することができる。し

かし、選択された特徴量からはどちらのクラスとも言い難い場合にも2値に識別するため、間違えた符号を組み合わせる場合、Joint Haar-like 特徴にも影響する。また、検出対象を人とした場合、人同士の画像中での重なりによる隠れが生じることが多い。そのため、隠れが生じている部分の特徴量を組み合わせることも考えられる。

そこで、我々は弱識別器の出力が連続値である Real AdaBoost を利用する。Real AdaBoost の弱識別器の出力は、大量の学習サンプルを用いることで統計的な信頼度を出力する。信頼度は、プラスの方向へ大きい場合は検出対象である信頼性が高く、マイナスの方向へ大きい場合は検出対象である信頼性が低いことを表す。我々は、弱識別器の出力を演算子によって組み合わせることで共起特徴を生成する。演算子には、和 (+) と積 (\times) を用いることにより、組み合わせによる表現のバリエーションを増やすことで、識別に有効な共起特徴を生成する。

3. 共起特徴による人検出

Real AdaBoost の複数の弱識別器から得られる出力と評価値を演算子により結合した共起特徴を提案する。図2に提案手法の流れを示す。本研究は、画像中から検出ウィンドウをラスタスキャンし、検出ウィンドウから得られる HOG 特徴量から共起特徴を生成し、この特徴により人 (ポジティブクラス) とそれ以外 (ネガティブクラス) の判別を行う。本章では、low-level 特徴として用いる HOG 特徴量について説明した後、提案する mid-level な特徴である共起特徴と Real AdaBoost による識別器構築について説明する。

3.1 low-level 特徴：HOG 特徴量

Histograms of Oriented Gradients (HOG) 特徴量 [3] とは、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ベクトルである。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。まず最初に、各ピクセルの輝度 L から勾配強度 m と勾配方向 θ を次式より算出する。

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

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

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

次に、算出された勾配強度 m と勾配方向 θ を用いて、セル領域 (5×5 ピクセル) において、輝度の勾配方向ヒストグラムを作成する。ただし、算出された勾配方向は、 $0^\circ - 360^\circ$ となるが、 $0^\circ - 180^\circ$ に変換する。これにより、人の服装に対して

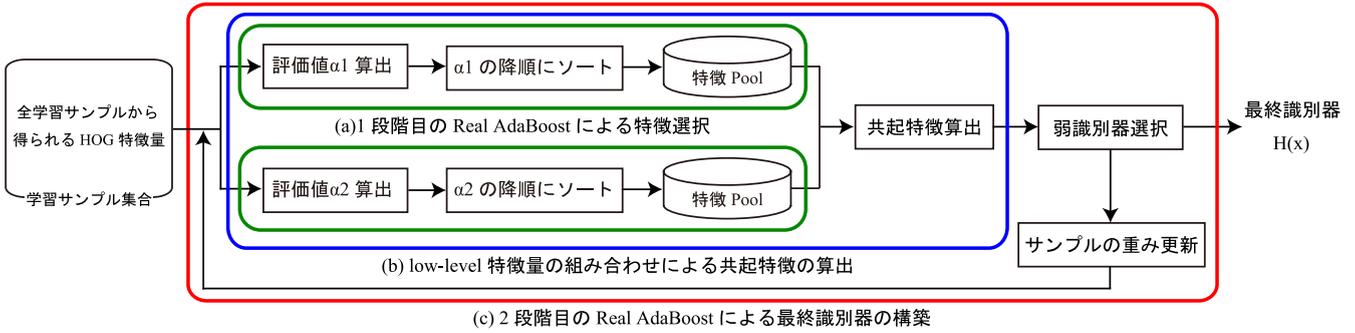


図 4 最終識別器の構築の流れ

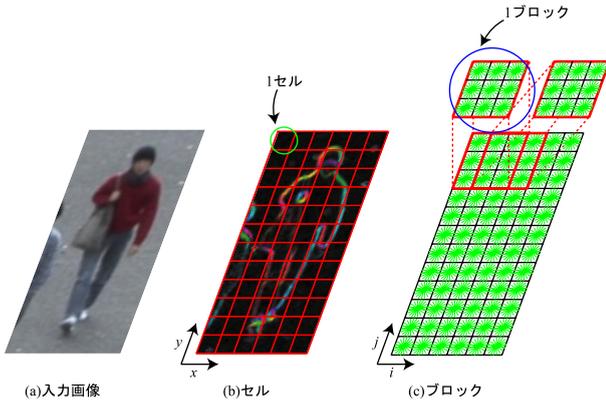


図 3 HOG 特徴量算出に用いる領域の構造

影響されない勾配方向が得られる．今回は勾配方向を 20° ずつに分割することにより 9 方向の勾配方向ヒストグラムを得る．最後に、次式によりブロック領域 (3×3 セル) 毎に特徴量を正規化して抽出する．

$$v = \frac{v}{\sqrt{\sum_{i=0}^k v(i)^2 + \varepsilon}} \quad (\varepsilon = 1) \quad (4)$$

ここで、 v は HOG 特徴量、 k はブロック内の HOG 特徴量の数、 ε は分母が 0 の場合に計算不能になることを防ぐ係数である．ブロックは 1 セルずつ移動させながら HOG 特徴量の抽出を行う．検出ウィンドウを 30×60 ピクセルに正規化した場合、3,240 個の特徴量が得られる．

3.2 共起特徴と Real AdaBoost による識別器構築

提案する共起特徴を算出し、人とそれ以外の判別する最終識別器の構築を行う．共起特徴を用いた学習の流れを図 4 に示す．ここでは、学習部を 3 つに分けて説明する．1 つ目は、1 段階目の Real AdaBoost による特徴選択 (図 4(a))、2 つ目は low-level 特徴の組み合わせによる共起特徴の算出 (図 4(b))、3 つ目は 2 段階目の Real AdaBoost による最終識別器の構築 (図 4(c)) である．

3.2.1 1 段階目の Real AdaBoost による特徴選択

共起特徴は、複数の HOG 特徴量を組み合わせで算出する．ここでは、2 個の特徴量の組み合わせについて説明するが、提案手法では M 個の特徴量の組み合わせについても容易に拡張することが可能である．得られた 3,240 個の HOG 特徴量の全

組み合わせから、Real AdaBoost により最もエラーの少ない 2 個の特徴量の組み合わせを探索した場合、膨大な学習時間が必要となる．本研究では、効率的な学習を行うために 3,240 個の HOG 特徴量から、Real AdaBoost のアルゴリズムを用いて特徴選択を行い、識別に有効な N 個の HOG 特徴量のみを組み合わせを行う．これにより、検出精度を維持しながら、学習時間の削減することができる．

まず、HOG 特徴量から確率密度分布 W_+ 、 W_- を作成する．確率密度分布 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．

$$W_+^j = \sum_{i: j \in J \wedge y_i = +1} D_t(i) \quad (5)$$

$$W_-^j = \sum_{i: j \in J \wedge y_i = -1} D_t(i) \quad (6)$$

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 BIN の番号、 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 BIN の番号 j へ学習サンプルの重み $D_t(i)$ を加算することで、確率密度分布 W_{\pm} を作成することができる．1 次元ヒストグラムの BIN 数は、学習サンプルの数に応じて適切な値にする必要がある．本研究では、実験より 1 次元ヒストグラムの BIN の数を 64 とした．作成した確率密度分布 W_{\pm} は、クラス毎の確率密度分布の総和が 1 となるように正規化する．

次に確率密度分布 W_{\pm} を用いて、分布の分離度を表す評価値 α を次式より算出する．

$$\alpha = 1 - \sum_j \sqrt{W_+^j W_-^j} \quad (7)$$

この評価値 α が大きいほどポジティブクラスとネガティブクラスの分布が分離していることを意味する．最後に、全ての HOG 特徴量から得られる確率密度分布の分離度を表す評価値 α を用いて、大きい順にソートを行い、 α の大きい方から N 個を特徴 Pool とする．本研究では、特徴 Pool を 200 個とした．

3.2.2 low-level 特徴の組み合わせによる共起特徴の算出

提案する共起特徴は、Real AdaBoost の複数の弱識別器から得られる出力と評価値を演算子によって結合した特徴量である．そのため、複数の特徴量を演算子により多様な捉え方で観測することが可能であるため、単一の特徴のみでは識別困難な

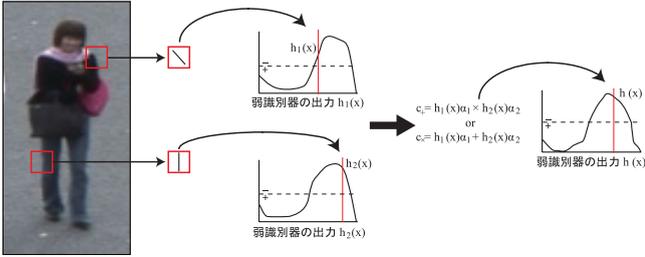


図5 HOG特徴量による共起特徴

パターンに対しても高精度な検出が可能となる。

共起特徴算出の方法について説明する。まず、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分布 W_+ , W_- を用いて、弱識別器の出力 $h(\mathbf{x})$ を算出する。学習サンプルより得られる特徴量 \mathbf{x} の値から、1次元ヒストグラムのBINの番号 j を算出し、それに対応した確率密度分布 W_{\pm} から弱識別器の出力 $h(\mathbf{x})$ を次式より算出する。

$$h(\mathbf{x}) = \frac{1}{2} \ln \frac{W_+ + \epsilon}{W_- + \epsilon} \quad (8)$$

ここで、 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。本研究では、実験より $\epsilon = 0.0000001$ とした。学習サンプルより得られる特徴量 \mathbf{x} の値から、1次元ヒストグラムのBINの番号 j より確率密度分布 W_{\pm} の値が得られる。

算出した Real AdaBoost の弱識別器の出力 $h(\mathbf{x})$ と評価値 α を用いて共起特徴を算出する。検出ウィンドウから得られた HOG 特徴量を \mathbf{x} , 弱識別器の出力を $h_1(\mathbf{x}_1)$, $h_2(\mathbf{x}_2)$, 評価値を α_1 , α_2 とした場合、下記の2パターンの組み合わせより共起特徴 $C1$, $C2$ を次式より表現する。

$$C1 = h_1(\mathbf{x})\alpha_1 + h_2(\mathbf{x})\alpha_2 \quad (9)$$

$$C2 = h_1(\mathbf{x})\alpha_1 \times h_2(\mathbf{x})\alpha_2 \quad (10)$$

まず、共起特徴 $C1$ は、評価値により弱識別器の出力を重み付けした値の和であるため、2つの特徴量を総合的に捉える特徴量となる。そのため、片方の特徴量が隠れやノイズ等の外乱の影響を受けた場合でも、もう1つの特徴量がクラスをよく表現していればそのクラスを反映した特徴量となる。次に、共起特徴 $C2$ は、弱識別器の出力に評価値を重み付けした値の積をとるため、両方の重みつき出力が高いときは高い共起特徴の値、それ以外は小さい共起特徴の値となる。検出時には、検出ウィンドウから得られる HOG 特徴 \mathbf{x} から弱識別器の出力を $h_1(\mathbf{x}_1)$, $h_2(\mathbf{x}_2)$ を求め、学習の際に算出した評価値を α_1 , α_2 から共起特徴 $C1$ もしくは $C2$ を算出する。2個の HOG 特徴量から共起特徴を生成する概念図を図5に示す。特徴 Pool が 200 個の場合、弱識別器の候補は $200 \text{ 個} \times 200 \text{ 個} \times 2 \text{ 演算子} = 80,000$ 個となる。これらの弱識別器候補は、Real AdaBoost を用いた識別器構築の特徴選択時に自動的に選択される。

3.2.3 2段階目の Real AdaBoost による最終識別器の構築
生成した共起特徴を用いて2段階目の Real AdaBoost により最終識別器の構築を行う。図6に共起特徴を用いた Real AdaBoost のアルゴリズムを示す。

3.2.1にて説明した共起特徴算出のために作成する特徴 Pool は、更新される学習サンプルの重みに従い学習ラウンド毎に作成し直す。これにより、前の学習ラウンドで誤識別した学習サンプルに対して重みが高くなるため、このサンプルを正識別するような特徴 Pool が各ラウンド毎に作成される。

Real AdaBoost のアルゴリズム

1. 前処理
 N 個の学習サンプル x_1, x_2, \dots, x_N
クラスに対応したラベル $y \{+1, -1\}$ の用意
2. 学習サンプルの重み D_t を初期化
 $D_{1n} = 1/N$
3. For $t = 1, 2, \dots, T$ T 回の学習回数
For $m = 1, 2, \dots, M$ M 個の弱識別器
 - 3.1 共起特徴の算出
 - 3.2 弱識別器 h_t の確率密度分布 W_t の作成
 $W_+^j = \sum_{i: j \in J_i^y = +1} D_t(i)$
 $W_-^j = \sum_{i: j \in J_i^y = -1} D_t(i)$
 - 3.3 弱識別器 $h(x)$
 $h(x) = \frac{1}{2} \ln \frac{W_+^j + \epsilon}{W_-^j + \epsilon}$
 - 3.4 弱識別器選択のための評価値 Z_m の算出
 $Z_m = 2 \sum_j \sqrt{W_+^j W_-^j}$
4. 最も Z_m が小さい弱識別器を選択
 $h_t = \arg \min Z_{t,m}$
5. 学習サンプルの重み $D_t(i)$ の更新
 $D_{t+1}(i) = D_t(i) \exp[-y_i h_t(x_i)]$
6. 学習サンプルの重み $D_t(i)$ の正規化
7. 最終識別器
 $H(x) = \text{sign}(\sum_{t=1}^T h_t(x))$

図6 Real AdaBoost のアルゴリズム

4. 評価実験

本章では、提案する共起特徴の有効性を示すために、歩行者検出実験により従来法との比較を行う。

4.1 データベース

文献[7]で用いられているデータベースを利用して実験を行う。このデータベースには、屋外で撮影された映像から、人の全身を切り出すことで作成されている。背景や照明、人の向き、人同士による画像中の重なりによる部分的な隠れ、切り出す際の位置ずれなどを含んでいる。学習用のポジティブサンプルは 2,053 枚、ネガティブサンプルは 6,253 枚、評価用のポジティブサンプルは 1,023 枚、ネガティブサンプルは 1,233 枚用意されている。図7に学習に用いたサンプルの一部を示す。

4.2 実験概要

文献[7]の評価用データベースを用いて、提案手法と HOG+Real AdaBoost, 特徴量として HOG 特徴量を用いた従

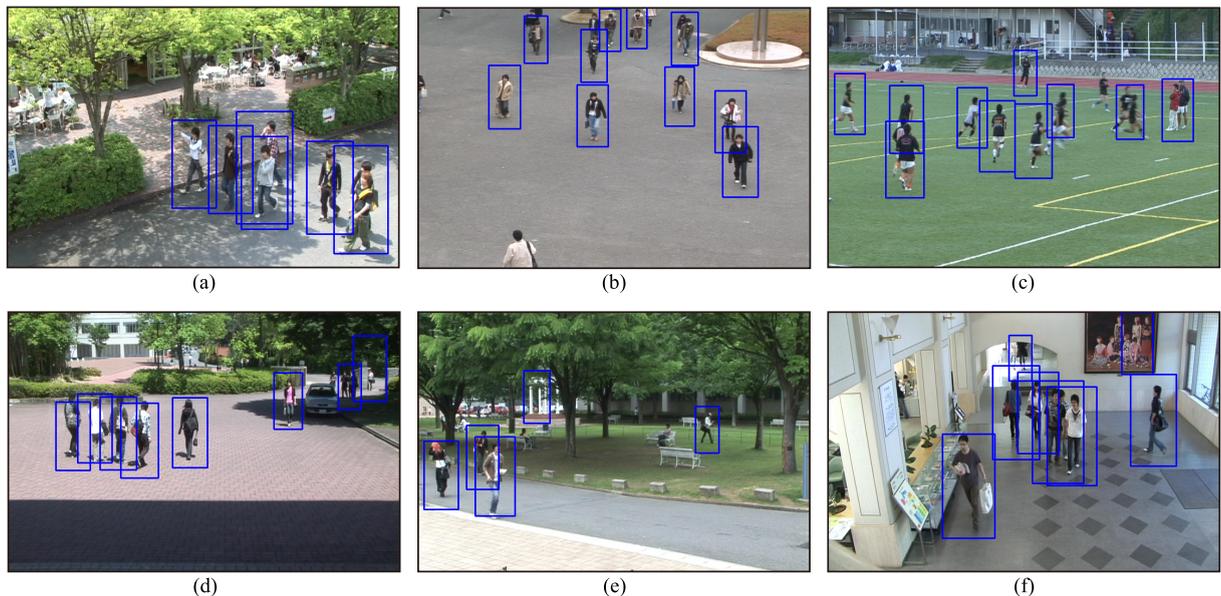


図9 提案手法による人検出例



図7 学習用サンプルの一部

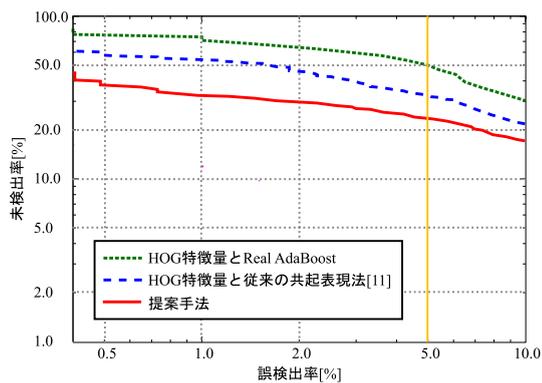


図8 DET カーブ

来の共起表現法 [11] の比較を行う。

比較には Detection Error Tradeoff(DET) [13] によって評価を行う。DET とは、横軸に誤識別率、縦軸に未識別率を両対数グラフによって表したものである。識別器のしきい値を変化させることによって、誤識別率に対する未識別率の比較を行うことが可能である。原点に近いほど検出性能が良いことを表す。

4.3 実験

図8に実験結果を示す。DET カーブでは、原点に近いほど検出性能が高いことを意味するため、提案手法が他の手法と比べ、全体的に検出性能が高いことがわかる。誤検出率 5.0% の未検出率を比較した場合、提案手法は従来の共起表現法よりも検出率が 6.8% 向上している。これは、Real AdaBoost を用いることで弱識別器の出力が連続値で表現することが可能となり、演算子で多様な特徴量の捉え方が可能となったため、検出率が向上したと考えられる。

4.4 検出実験

構築した検出器により人検出の実験を行う。人検出を行うために、検出ウィンドウを画像左上からスケールを変化させて複数回ラスタスキャンする。これにより、人のスケールが異なる場合でも検出可能となる。人として検出されたウィンドウは、最後に Mean Shift クラスタリングによる検出ウィンドウの統合処理を行う [14]。図9に複数の場所において撮影した映像に対する人検出例を示す。図9から高精度な人検出を行っていることがわかる。図9(a), (b), (c) は各結果において未検出が多少あるが、人同士の画像中における重なりによる隠れや人の素早い動きによるぶれの影響だと考えられる。また、図9(d), (e), (f) では、幾つか誤検出がみられる。これは、low-level な特徴量として HOG 特徴量を用いているため、人に似た背景を人として誤検出してしまったためである。

4.5 処理時間

提案手法による学習と検出に必要な処理時間を計測する。使用した PC は、Intel Xeon 2.66GHz、メモリ 3.00GB である。

学習には、4.1 節で示したデータベースを使用し、学習回数は 200 回として、約 2 時間を要した。

次に、識別に用いる特徴量の数と精度の関係を提案手法と従来法の比較を行う。4.3 節の実験では、HOG 特徴量+Real AdaBoost では 500 回の学習回数で 500 個の特徴量を使用し、HOG 特徴量+従来の共起表現法と提案手法では 200 回の学習

回数で 400 個の特徴量を使用した。提案手法は、HOG 特徴量+Real AdaBoost と比べて、特徴量の数が少ないにもかかわらず検出性能が向上している。また、従来の共起表現法と比較して、同数の特徴量の数を使用しているが、検出精度が向上している。このことから提案手法は、共起特徴として HOG 特徴量を組み合わせることにより、より識別に適した少数の共起特徴が生成され、高速、高精度に識別を行うことが可能となったと考えられる。

人検出の実験では、HOG 特徴量を高速に算出するために Integral Histograms [15] を用いた。検出する画像サイズが 320×240 ピクセルの場合、12,014 個の検出ウィンドウをポジティブクラスとネガティブクラスに判別する。その結果、1 枚の入力画像を処理するために必要な時間は約 170msec であった。効率良く識別することにより処理時間を削減する Viola と Jones が提案したカスケード型 AdaBoost [1] をすることにより、さらなる高速な人検出が期待できる。

5. 多重解像度間の共起による人検出

4. 章では、一枚の画像から得られる HOG 特徴量の空間的な共起を表現し、人検出における提案する共起特徴の有効性を確認した。組み合わせられる HOG 特徴量は、異なる空間的位置から得られる特徴量であった。提案手法では、空間位置が異なるだけでなく、異なる解像度間から得られる HOG 特徴量の共起の表現も可能である。そこで、本章では多重解像度から得られる HOG 特徴量の共起を表現し、共起特徴により人検出を行う。

本章では、検出対象を人の上半身とし、形状に特徴がある頭部と上半身(頭部から肩にかけての Ω の形状)に分ける。頭部と上半身の画像を多重解像度により表現し、これらの画像から得られる HOG 特徴量を頭部と上半身から 1 個ずつ特徴量を選択することで共起を表現する。

5.1 多重解像度間の共起表現

検出対象である上半身画像の中で、特に形状に特徴がある頭部領域と上半身領域は、特徴を最も表現する解像度が同一であるとは限らない。そこで、本研究では多重解像度から HOG 特徴量を抽出し、頭部らしい特徴と上半身らしい特徴を共起特徴により同時に捉えることを考える。図 10 は、 128×128 ピクセルの上半身画像と 64×64 ピクセルの頭部画像を入力し、半分の大きさまで 0.125 ずつ倍率を縮小した場合の HOG 特徴量を可視化したものである。本研究では、下記のパラメータにより多重解像度画像からの HOG 特徴量を抽出する。

- バッチサイズ

- 頭部： 64×64 , 56×56 , 48×48 , 40×40 , 32×32 ピクセル
- 上半身： 128×128 , 112×112 , 96×96 , 80×80 , 64×64 ピクセル

- セルサイズ

- 頭部： 8×8 ピクセル
- 上半身： 16×16 ピクセル

- ブロックサイズ

- 頭部： 2×2 セル

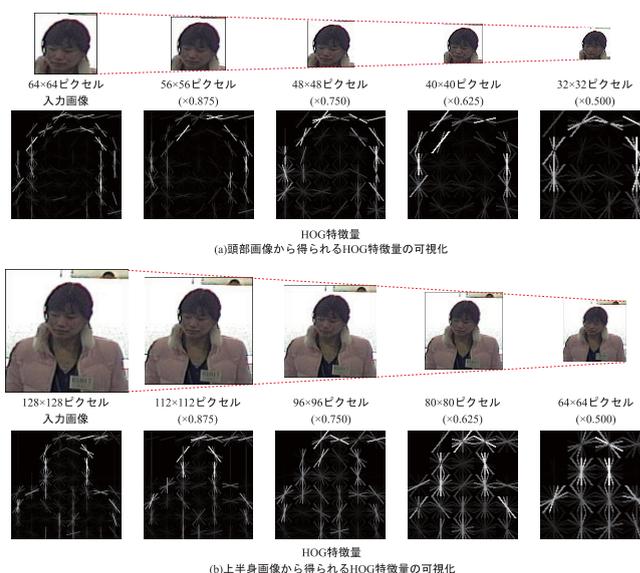


図 10 HOG 特徴量の変化

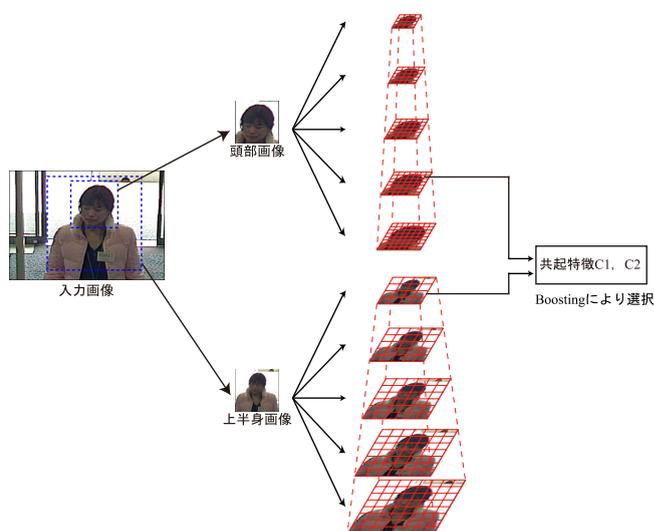


図 11 多重解像度 HOG 特徴量による共起特徴

- 上半身： 2×2 セル

上記のパラメータから HOG 特徴量を算出した場合、頭部と上半身の領域から各 4,860 個得られる。これらの HOG 特徴量を用いて 3.2 節にて述べた共起表現法により多重解像度間の共起を表現を行い人の検出を行う。図 11 に多重解像度間の共起表現の流れを示す。まず、入力画像から頭部と上半身に分け、それぞれダウンサンプリングすることで多重解像度画像を作成し、HOG 特徴量を抽出する。次に、頭部と上半身から得られる HOG 特徴量を 1 個ずつ用いることにより式 9, 式 10 より共起特徴を生成する。これにより、解像度間の共起を表現した共起特徴を生成することが可能となる。

5.2 評価実験

提案手法の有効性を示すために従来法との比較実験について示す。

5.2.1 データベース

学習と評価実験を行うための画像を収集した。データベースのポジティブサンプルには、明るさ、人の向き、画像に対す

人の大きさが異なる場所で撮影した映像から、人の頭部と上半身を切り出した。ネガティブサンプルには、Caltech 256 database [16] の clutter クラスからランダムに切り出した。学習用のポジティブサンプルは 4,585 枚、ネガティブサンプルは 3,738 枚用意した。同様に、評価用のポジティブサンプルは 3,920 枚、ネガティブサンプルは 10,000 枚用意した。学習に用いたデータベースの一部を図 12 に示す。



図 12 ポジティブクラスの上半身画像データベース例

5.2.2 実験

評価用データベースを用いて提案手法と多重解像度 HOG [3]+Real AdaBoost, 多重解像度 HOG の共起 [11] との比較を行う。比較には Detection Error Tradeoff (DET) カーブ [13] によって評価を行う。

図 13 に実験結果を示す。図 13 では、提案手法が他の手法と比べ、全体的に原点に近いことから検出性能が高いことわかる。誤検出率 0.1% の未検出率を比較した場合、提案手法は従来法である HOG の共起と比較して未検出率が約 15.0% 向上している。これは、Real AdaBoost を用いることで弱識別器の出力が連続値で表現することが可能となり、演算子で多様な特徴量の捉え方が可能となったため検出率が向上したと考えられる。

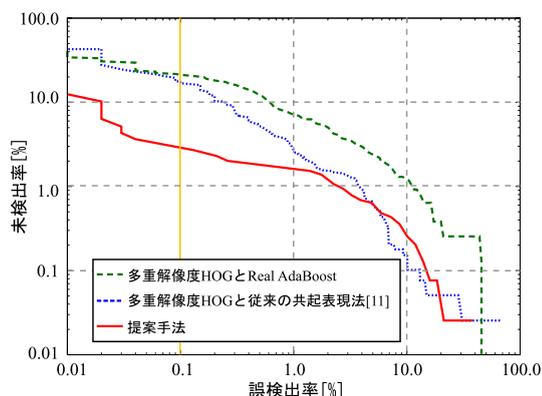


図 13 DET による比較結果

図 14 に複数の場所において撮影した映像に対する人検出例を示す。人同士の重なりによる隠れが発生しても、精度良く検出できていることがわかる。さらに、頭部と上半身の形状を同

時に捉えることで、向きに依存せず人の検出を行うことができている。

5.3 考察

学習時に選択された特徴量から提案手法の考察を行う。図 15 に学習時に選択された HOG 特徴量を可視化した結果を示す。HOG 特徴の勾配方向を 9 方向で表現しており、輝度が高いほど Real AdaBoost における識別に有効な特徴量として選択されたことを表す。図 15(e) に示すように、形状に沿った勾配に対応する HOG 特徴が選択されていることがわかる。また、図 15(a), (b), (c), (d) に最終識別器の各ラウンドにおいて選択された HOG 特徴量を示す。特徴選択の傾向が出やすい学習の初めでは、頭部は高解像度の HOG 特徴が選択され、上半身は低解像度の HOG 特徴量が選択されていることがわかる。このことから頭部は高い解像度の方が特徴を表現しやすく、上半身は低解像度の方が特徴を表現しやすいと考えられる。これを確かめるために学習によって選択された HOG 特徴量を解像度毎にヒストグラムで表現した図 16 に注目する。頭部は、比較的解像度の高い画像から得られる HOG 特徴が選択される傾向があり、上半身は高解像度の HOG 特徴量が最も選択されているが、頭部に比べ低解像度の HOG 特徴量が選択される傾向がある。これは、学習サンプルの平均勾配画像からもわかるように、頭部のように勾配にばらつきが少ないと、高解像度の HOG 特徴量が選択され、上半身のように勾配が分散していると、低解像度の HOG 特徴量が選択されている。今回作成したデータベースの上半身画像は、肩のラインが性別、衣服、年齢によって差異があったため、肩のラインの勾配の分布にばらつきがあった。低解像度画像からの HOG 特徴量は、広い範囲でヒストグラムをまとめることになるため、このばらつきを吸収することができたと考えられる。一方、頭部画像は勾配のばらつきが小さいため、局所領域でも頭部のラインを捉えることができるため、高解像度の HOG 特徴量が選択されたと考えられる。

6. おわりに

本稿では、Boosting に基づく特徴量の共起表現法とそれを用いた人検出法を提案した。提案手法は、low-level 特徴として HOG 特徴量を入力とした Real AdaBoost の弱識別器の出力である連続値と評価値を、演算子によって結合した共起特徴から Real AdaBoost を用いて人と人以外に判別する。共起特徴は、複数の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な特徴量であるため、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、共起特徴に用いる演算子に差 (-) を加え、多クラスの識別に拡張する予定である。

文献

- [1] P. Viola and M. Jones, "Robust Real-Time Face Detection", International Journal of Computer Vision, 57(2), pp. 137-154, 2004.
- [2] K. Levi and Y. Weiss, "Learning Object Detection from a Small Number of Examples: the Importance of Good Features.", IEEE Computer Vision and Pattern Recognition, vol. 2, pp. 53-60, 2004.
- [3] N. Dalal and B. Triggs, "Histograms of Oriented Gradients

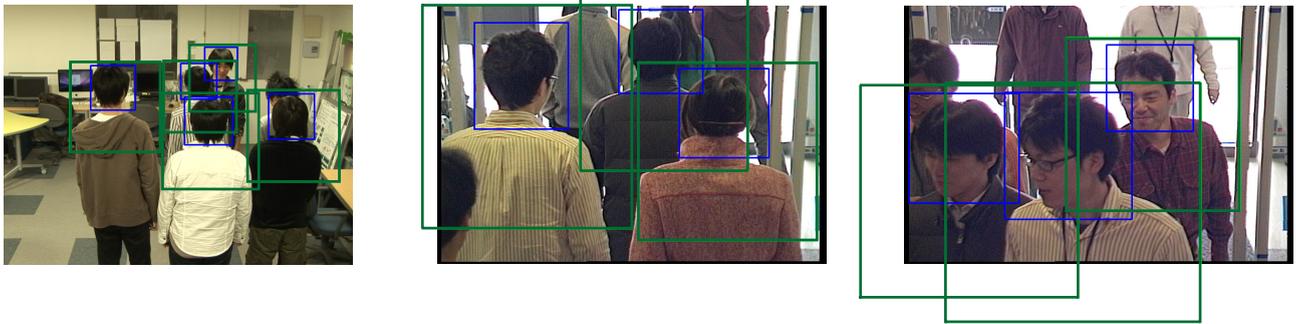


図 14 提案手法による人検出例

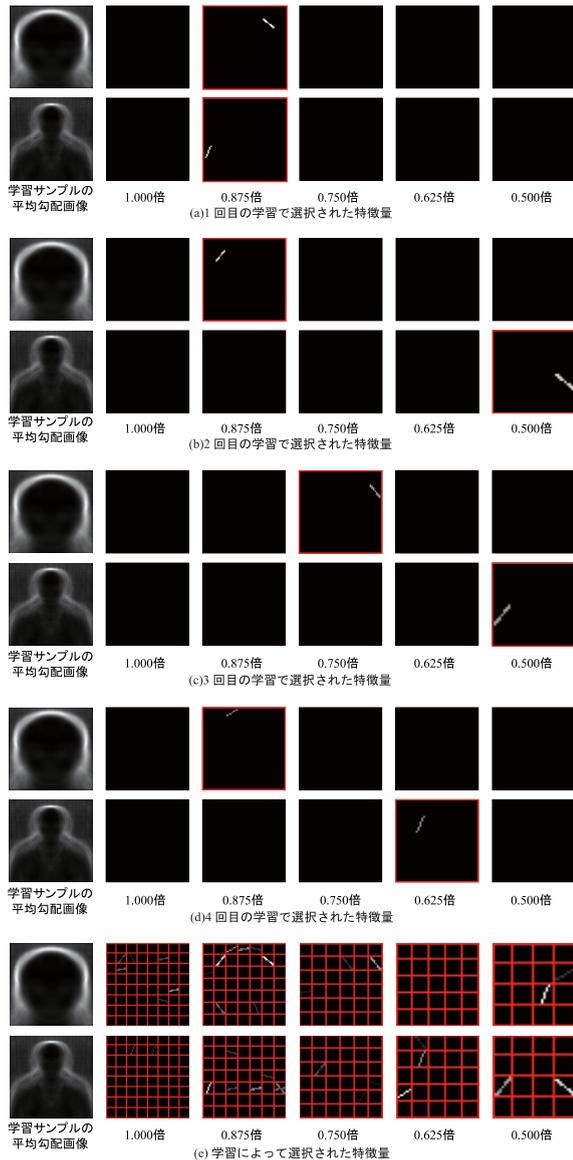


図 15 学習によって選択された HOG 特徴量の可視化

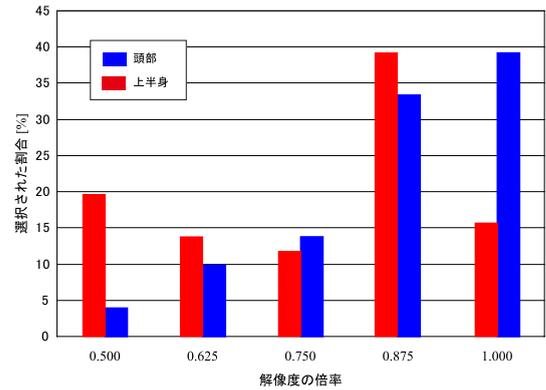


図 16 学習によって選択された HOG 特徴量の解像度

for Human Detection”, IEEE Computer Vision and Pattern Recognition, vol.1, pp. 886-893, 2005.

[4] B. Wu and R. Nevatia, “Detection of multiple, partially occluded humans in a single image by Bayesian combination of edgelet part detectors”, IEEE International Conference on Computer Vision, vol. 1, pp. 90-97, 2005.

[5] P. Viola, M. Jones and D. Snow, “Detecting Pedestrians Using Patterns of Motion and Appearance”, IEEE Interna-

tional Conference on Computer Vision, pp. 734-741, 2003.

[6] N. Dalal, B. Triggs and C. Schmid, “Human Detection Using Oriented Histograms of Flow and Appearance”, IEEE European Conference on Computer Vision, vol. 2, pp. 428-441, 2006.

[7] 山内悠嗣, 藤吉弘巨, Hwang Bon-Woo, 金出武雄, “アピランスと時空間特徴の共起に基づく人検出”, 画像の認識・理解シンポジウム (MIRU2007), pp. 1492-1497, 2007. .

[8] C. Hou, H. Ai and S. Lao, “Multiview Pedestrian Detection Based on Vector Boosting”, Eighth Asian Conference on Computer Vision, vol. 1, pp. 220-229, 2007.

[9] T.Hastie, R.Tibshirani and J.Friedman, “The Elements of Statistical Learning - Data Mining, Inference, and Prediction”, Springer-Verlag, 2001.

[10] R. E. Schapire and Y. Singer, “Improved Boosting Algorithms Using Confidence-rated Predictions”, Machine Learning, No. 37, pp. 297-336, 1999.

[11] T. Mita, T. Kaneko, B. Stenger, O. Hori, “Discriminative Feature Co-occurrence Selection for Object Detection”, IEEE Pattern Analysis and Machine Intelligence, vol. 30, no. 7, pp. 1257-1269, 2008.

[12] P. Sabzmeydani and G. Mori : “Detecting Pedestrians by Learning Shapelet Features”, IEEE Computer Vision and Pattern Recognition , pp1-8, 2007.

[13] 鷺見和彦, 関真規人, 波部齊, “物体検出 - 背景と検出対象のモデリング - ”, 情報処理学会研究会 研究報告 (CVIM2005), Vol.2005, No.88, pp. 79-98, 2005.

[14] D. Comaniciu and P. Meer, “Mean shift analysis and applications”, IEEE International Conference on Computer Vision, pp. 1197-1203, 1999.

[15] F. Porikli, “Integral Histogram: A Fast Way to Extract Histograms in Cartesian Spaces”, IEEE Computer Vision and Pattern Recognition, pp829-836, 2005.

[16] Caltech256, “http://www.vision.caltech.edu/Image_Datasets/Caltech256/”