

## Boosting に基づく特徴量の共起表現による人検出\*

山内 悠嗣<sup>†a)</sup> 山下 隆義<sup>†,††b)</sup> 藤吉 弘亘<sup>†c)</sup>

## Pedestrian Detection by Boosting-Based Feature Co-occurrence Representation\*

Yuji YAMAUCHI<sup>†a)</sup>, Takayoshi YAMASHITA<sup>†,††b)</sup>, and Hironobu FUJIIYOSHI<sup>†c)</sup>

あらまし 本論文では、Boosting に基づく特徴量の共起表現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、入力特徴がオクルージョンなどの影響によって、どちらのクラスとも言いがたい場合にも 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、弱識別器の出力が連続値である Real AdaBoost を用いて、出力を演算子によって結合した共起表現による人検出法を提案する。提案手法は、オクルージョンなどの影響を抑制することができるため、高精度な検出が期待できる。評価実験により、従来法と比較して誤検出率 5.0%において検出率を約 6.8%向上させることができた。

キーワード 人検出, 共起, Real AdaBoost, Histograms of Oriented Gradients

## 1. ま え が き

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 など多岐にわたる分野で実現が期待されている。また、人の追跡や姿勢認識の前処理としても必要な技術であるため、近年最も盛んに行われている研究の一つである。近年の人検出法は、Viola と Jones が提案した顔検出 [1] に代表されるように low-level な特徴と統計的学習による Boosting の組合せによる手法が多い。人検出に有効な low-level な局所特徴量として、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、局所領域におけるこう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 [4] など、エッジに着目した特徴量が多数提案され、高い検出精度を達成している。

更に、検出精度を向上させるために、アピアランスとは異なる特徴量の併用 [5] ~ [7] などが提案されている。しかし、これらの手法は、背景画像や複数枚の連続した画像が必要などの条件がある。そこで、1 枚の画像から高精度な検出を実現するために特徴量間の関係をとらえることができる Shapelet 特徴 [10] や Joint Haar-like 特徴 [9] が提案された。これらの特徴量は、複数の low-level な局所特徴量を Boosting によって組み合わせることにより、特徴量間の関連性をとらえることができる。このため、検出対象物体の構造に基づいた特徴の関係性を評価することにより高精度な検出が可能となる。しかし、Shapelet 特徴量は多数の弱識別器による識別が必要であるため処理コストが大きく、Joint Haar-like 特徴量はオクルージョンなどの影響を受けやすい問題がある。

そこで、本論文では Real AdaBoost [8] の弱識別器の出力を演算子によって結合する新しい特徴量間の共起表現による人検出法を提案する。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。そこで、弱識別器の出力を演算子によって結合することで共起確率特徴を生成する。提案する共起確率特徴は、複数の特徴量を演算子により多

<sup>†</sup> 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

<sup>††</sup> オムロン株式会社, 京都府  
OMRON Corporation, 9-1 Kizugawadai, Kizu-cho, Soraku-  
gun, Kyoto-fu, 619-0283 Japan

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

b) E-mail: takayosi@ari.ncl.omron.co.jp

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

\* 本論文は第 12 回画像の認識・理解シンポジウム推薦論文である。

様なとらえ方で表現することが可能であるため高精度な人検出が期待できる。

## 2. 従来の共起表現法

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、識別に有効な mid-level な特徴量を生成する手法が多数提案されている。Mita らは、複数の Haar-like 特徴の共起に基づく Joint Haar-like 特徴を提案した。Joint Haar-like 特徴量は、Haar-like 特徴量から 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。そのため、Joint Haar-like 特徴では、顔の構造に基づいた特徴量の関係性をとらえることができる。例えば、図 1 では、3 個の Haar-like 特徴によって顔の特徴的な部分を複数とらえることができる。これは、単一の Haar-like 特徴ではとらえられない、各 Haar-like 特徴量間の共起性をとらえることが可能となる。Joint Haar-like 特徴は、Viola と Jones の顔検出と比較して、高精度かつ高速に処理することが可能であり、検出対象や組み合わせる low-level な特徴量に制限がないため、このような共起を表現する方法は非常に有効である。しかし、組み合わせる特徴量の中に、誤った 2 値符号が含まれていた場合、Joint Haar-like 特徴量にも影響するため、最終的な判定に用いられる同時確率の確率分布によっては誤識別する問題がある。特に、検出対象が人など形状変化の激しいものや画像中の重なりによるオクルージョンが発生した場合、これらの問題が多く発生することが考えられる。

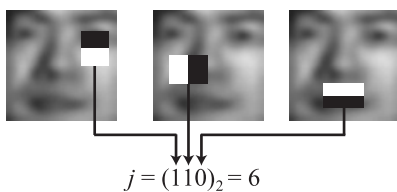


図 1 Joint Haar-like 特徴  
Fig. 1 Joint Haar-like feature.

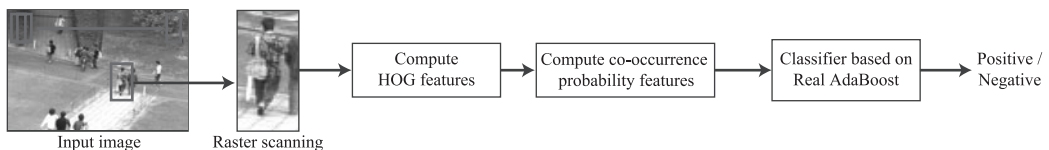


図 2 提案手法による人検出の流れ  
Fig. 2 Flow of pedestrian detection by the proposed method.

そこで、我々は弱識別器の出力が連続値である Real AdaBoost を利用し、弱識別器の出力を演算子によって結合することによって共起確率特徴を算出する。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連続値によって出力する。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。そのため、オクルージョンなどの影響によって low-level な特徴量を用いた識別が間違った場合においても、それ以外の low-level な特徴量が正しい識別を行うことにより、最終的な識別を正しく行うことができる。

## 3. 共起確率特徴による人検出

Real AdaBoost の複数の弱識別器から得られる出力を演算子により結合した共起確率特徴を提案する。図 2 に提案手法による識別の流れを示す。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 low-level 特徴量から共起確率特徴を生成し、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、提案する共起確率特徴について説明する。

### 3.1 HOG 特徴量

本研究では、low-level な特徴量として、人検出において有効な特徴量とされている Histograms of Oriented Gradients (HOG) 特徴量 [3] を用いる。HOG 特徴量とは、局所領域における輝度のこう配方向をヒストグラム化した特徴ベクトルである。近接画素のこう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  $L$  からこう配強度  $m$  とこう配方向  $\theta$  を次式より算出する。

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

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

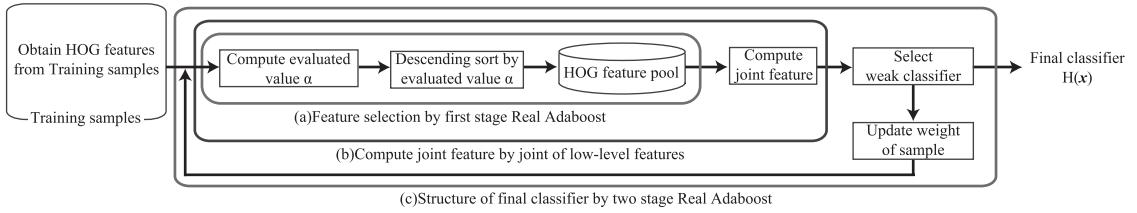


図 4 最終識別器の構築

Fig. 4 Structure of final classifier.

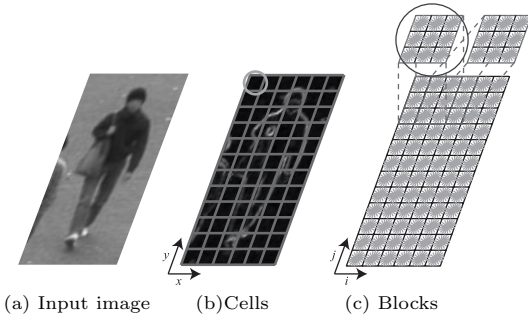


図 3 HOG 特徴量算出に用いるセルとブロック

Fig. 3 Cells and blocks for HOG.

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

次に、算出されたこう配強度  $m$  とこう配方向  $\theta$  を用いて、セル領域 (5 × 5 ピクセル (図 3(b))) において、輝度のこう配方向ヒストグラムを作成する。ただし、算出されたこう配方向は、 $0^\circ - 360^\circ$  となるが、 $0^\circ - 180^\circ$  に変換する。この理由は、歩行者の服の明るさとそれに隣接する背景領域の明るさの大小関係が逆転しても不変となり、人の服装に対して影響されないこう配方向が得られるためである。得られたこう配方向は、 $20^\circ$  ずつに分割することにより 9 方向のこう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 (図 3(c))) ごとに特徴量を正規化して抽出する。

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

ここで、 $v$  は HOG 特徴量、 $k$  はブロック内の HOG 特徴量の数、 $\epsilon$  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  $30 \times 60$  ピクセルに正規化した場合、3,240 個の特

量が得られる。

### 3.2 共起確率特徴の生成

共起確率特徴を算出し、人とそれ以外を判別する最終識別器の構築を行う。共起確率特徴を用いた学習の流れを図 4 に示す。ここでは、学習部を三つに分けて説明する。一つ目は、1 段階目の Real AdaBoost による特徴選択 (図 4(a))、二つ目は low-level 特徴の結合による共起確率特徴の算出 (図 4(b))、三つ目は 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

共起確率特徴は、複数の HOG 特徴量を結合させて算出する。ここでは、2 個の特徴量の結合について説明するが、提案手法では  $N$  個の特徴量の結合についても容易に拡張可能である。得られた 3,240 個の HOG 特徴量の全組合せから、Real AdaBoost により最もエラーの少ない 2 個の特徴量の組合せを探索した場合、共起確率特徴の弱識別器候補が  ${}_{3240}C_2 = 5,247,180$  個となるため膨大な学習時間が必要となる。そこで、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  $M$  個の HOG 特徴量のみを組合せを行う。これにより、検出精度を維持しながら、学習時間を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、一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  $D_i$  に基づいて次式より作成する。

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

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

ここで、 $t$  は学習回数、 $i$  は学習サンプルの番号、 $j$  は一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一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．一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，実験より一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，クラスごと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  $\alpha$  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  $\alpha$  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，すべての 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  $\alpha$  を用いて，大きい順にソートを行い， $\alpha$  の大きい方から  $M$  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 特徴の組合せによる共起確率特徴の算出

提案する共起確率特徴は，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とらえ方で観測することができ，単一の特徴のみでは識別困難なパターンに対しても高精度な検出が可能となる．

共起確率特徴の算出法について以下に述べる．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  $h(x)$  を算出する．学習サンプルより得られる特徴量  $x$  の値から，一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  $h(x)$  を次式より算出する．

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

ここで， $\epsilon$  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  $x$  の値から，一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  $h(x)$  を用いて共起確率特徴を算出する．検出ウィンドウが

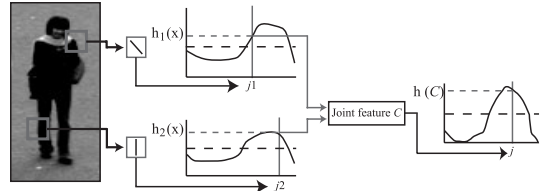


図 5 HOG 特徴量による共起特徴  
Fig. 5 Feature co-occurrence representation by HOG features.

ら得られた 3,240 個の HOG 特徴量を  $x$ ，弱識別器の出力を  $h1(x)$ ， $h2(x)$  とした場合，下記の 2 パターンの結合により共起確率特徴  $C1$ ， $C2$  を次式より表現する．

$$C1 = h1(x) + h2(x) \quad (9)$$

$$C2 = h1(x) \times h2(x) \quad (10)$$

まず，共起特徴  $C1$  は，弱識別器の出力の和であるため，二つの特徴量を総合的にとらえる特徴量となる．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，もう一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，弱識別器の出力の積をとるため，両方の重み付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  $x$  から弱識別器の出力を  $h1(x)$ ， $h2(x)$  を求め，共起特徴  $C1$  若しくは  $C2$  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  ${}_{200}C_2 \times 2$  演算子 = 39,800 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

生成した共起確率特徴を用いて 2 段階目の Real AdaBoost により最終識別器の構築を行う．3.2.1 にて説明した共起確率特徴の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ごとに作成し直す．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 が各ラウンドごとに作成される．

### 3.3 $N$ 個の特徴量を用いた共起特徴への拡張

3.2 では，共起特徴を算出するために 2 個の HOG

特徴量の結合を行ったが、提案手法は  $N$  個の HOG 特徴量についても容易に拡張することが可能である。本節では、 $N$  個の HOG 特徴量の結合を行い、共起特徴量を算出する。

### 3.3.1 最適な特徴量の組合せ探索

3.2 で示した共起特徴の生成法を用いて多数の HOG 特徴量の組合せを行った場合、特徴 Pool のサイズを小さくしても長時間の学習時間が必要となる。そこで、効率的な組合せ探索を行う Sequential Forward Selection (SFS) [11] を用いる。SFS は、はじめに最も評価値  $\alpha$  が高い HOG 特徴量を 1 個選択する。次に、共起特徴の評価値  $\alpha$  が最も高くなるように別の特徴量を 1 個追加する。これを  $N$  回繰り返すことにより、 $N$  個の HOG 特徴量を効率的に組み合わせることができる。

### 3.3.2 最適な結合させる特徴量の個数の推定

結合する HOG 特徴量の個数  $N$  を増やすことにより、より強力な弱識別器を作成することができる。しかし、共起の関係を強くするほど過学習となる等の問題が考えられるため、結合させる特徴量の最適な個数を各学習ラウンドごとに決定する必要がある。そこで、文献 [9] と同様の手法により、結合させる特徴量の最適な個数を推定する。文献 [9] では、学習用サンプルとは別に  $N'$  個の検証用サンプルを用いて評価を行う hold-out 法により結合させる特徴量の個数を推定している。hold-out 法は、結合させる特徴量の個数  $N$  をあらかじめ決めた上限値  $N_{max}$  まで増加させ、各  $N$  に対する弱識別器を学習し、 $N_{max}$  通りの弱識別器候補を得る。この中から、その学習ラウンド  $T'$  までの強識別器により検証用サンプル  $(x'_i, y'_i)$  に対する識別誤り率  $E_{T'}$  を最小とする  $N$  を選択している。識別誤り率  $E_{T'}$  は、次式より識別誤りとなった検証用サンプルの数を計数することにより求められる。

$$E_{T'} = \frac{1}{M'} \sum_{n=1}^{M'} I(H_{T'}(x'_i) \neq y'_i) \quad (11)$$

ここで、 $H_{T'}(x)$  は、それまでの学習ラウンドにおける強識別器として式 (12) で表される。なお、 $I(H_{T'}(x'_i) \neq y'_i)$  は、 $H_{T'}(x)$  による識別結果が誤りならば 1、正しい場合には 0 となる。

$$H_{T'}(x) = \text{sign}\left(\sum_{t=1}^{T'} h_t(x)\right) \quad (12)$$

## 4. 評価実験

本章では、提案手法の有効性を確認するために、下記に示す二つの実験を行う。

実験 (1)：提案手法と従来法の比較実験

実験 (2)：最適な特徴量の個数の推定

### 4.1 データベース

評価実験は、文献 [7] で使用されている人画像データベースを使用する。この人画像データベースは、屋外で撮影した映像から人の全身の画像を切り出すことで作成されている。カメラの設置環境は、地上から高さ約 6 ~ 12 m、設置角度を下方約 30° としている。得られた映像は様々な場所で撮影しているため、背景や照明、人の向きなどが異なったデータとなっている。人画像を切り出すために、人の全身を最小方形となるように囲い、最小方形の縦幅と横幅の 10% を余白として横端、縦端に加えている。また、評価用サンプルについては、位置ずれを含んだデータとなっている。上記の方法により切り出された人画像の解像度は、30 × 60 ~ 40 × 80 ピクセルとなった。画像データベースの枚数は、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 枚用いる。評価も学習と同様に切り出された画像を用いて行う。評価用は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 枚用いる。また、hold-out 法により結合する特徴量の個数を推定する実験 2 では検証用サンプルも必要となるため、新たに検証用としてポジティブ、ネガティブサンプルともに 1,000 枚用意した。図 6 に学習用サンプルの一部を示す。

### 4.2 評価方法

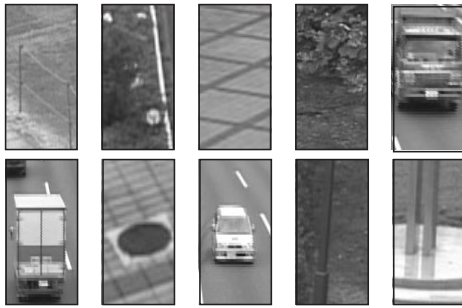
比較には Detection Error Tradeoff (DET) [12] によって評価を行う。DET とは、横軸に誤検出率 (False positive rate)、縦軸に未検出率 (miss rate) を両対数グラフによって表したものである。識別器のしきい値を変化させることによって、誤検出率に対する未検出率の比較を行うことが可能である。原点に近いほど検出性能が良いことを表す。

### 4.3 実験 1

実験 1 では、提案手法の共起表現の有効性を確認するために比較を行う。実験は、共起を表現しない方法 (HOG+Real AdaBoost)、従来の共起表現法 [9] (HOG+Conventional method)、和演算子 (式 (9)) を用いた共起表現法 (Proposed method (+))、



(a) Positive class



(b) Negative class

図 6 学習サンプルの一部

Fig. 6 Examples of training data.

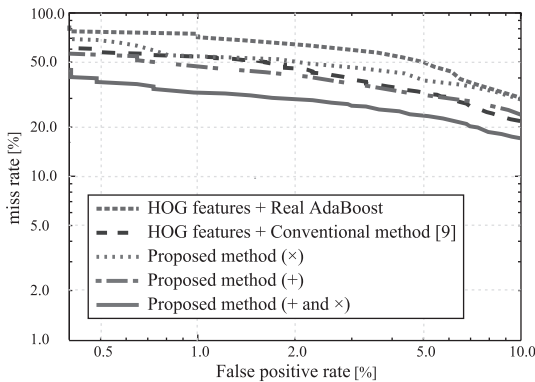


図 7 DET カーブ

Fig. 7 DET curves.

積演算子 (式 (10)) を用いた共起表現法 (Proposed method (x)), 和演算子 (式 (9)) と積演算子 (式 (10)) を用いた共起表現法 (Proposed method (+ and x)) を比較する. 提案手法は, 3.2 で記述した方法によって検出器を構築する. 学習は, 性能評価を平等にするために, 結合する特徴量数  $N = 2$  としたときの共起特徴を用いる. その際の強識別器に使用する特徴量は計 500 個とする.

表 1 誤検出率 5.0%における検出率  
Table 1 Detection rate with miss rate 5.0%.

手法	検出率
HOG + Real AdaBoost	50.2
HOG + 従来の共起表現法	69.4
提案手法 (x)	63.0
提案手法 (+)	69.6
提案手法 (+ and x)	78.3

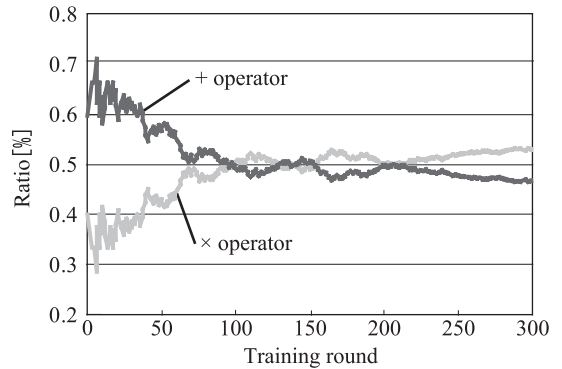


図 8 選択された演算子の割合

Fig. 8 Ratio of selected operators.

図 7 と表 1 に実験結果を示す. 結果より, DET カーブ全体が原点に近いことから, 最も検出性能が高い手法は和演算子と積演算子を用いた共起表現法であることが分かる. まず, 提案する共起表現法を比較する. 検出性能は, 和演算子と積演算子を用いた共起表現法, 和演算子を用いた共起表現法, 積演算子を用いた共起表現法 (x) の順に高いことが分かる. 異なる共起表現法を併用することで検出性能が向上しているため, 和演算子と積演算子は異なる共起表現法であり, Real AdaBoost により最適な共起表現方法を選択することで検出性能が向上していると考えられる. 和演算子と積演算子を用いた共起表現法において, 学習中に選択された演算子の割合を図 8 に示す. 学習の初期では, 和演算子の共起表現法が多く選択され, 学習ラウンド 100 回以降は和と積の演算子の選択される割合が同程度であることが分かる. そのため, 学習の初期では多くの学習サンプルを正識別させるために和演算子が多く選択され, 学習の中盤においては高い重みをもつ学習サンプルに対して正識別するために積演算子が選択されていると考えられる.

次に, 和演算子と積演算子を用いた提案手法と共起を表現しない方法の比較を行う. 誤検出率 5.0%の未検出率を比較した場合, 提案手法は共起を表現しない



図 9 識別可能となったサンプル例

Fig. 9 Example of classifiable data by the proposed method.

表 2 識別に用いる特徴量の数と学習回数

Table 2 Total number of features and training round to use for strong classifier.

弱識別器に用いる特徴量数	学習回数	全特徴量数
1	500	500
3	167	501
7	72	504
$N_{max} = 7$	110	500

方法よりも検出率が 27.8%向上した。最後に、和演算子と積演算子を用いた提案手法と従来の共起表現法の比較を行う。誤検出率 5.0%の未検出率を比較した場合、提案する共起表現法は従来の共起表現法よりも検出率が 6.8%向上した。これは、提案手法では複数の演算子により弱識別器の出力を結合するため、従来法よりも多様な結合の表現が可能になったと考えられる。従来の共起表現法では正しい識別ができないが、提案手法では識別できるようになった例を図 9 に示す。図 9(a) は、オクルージョンが発生している場合や人の形状変化が激しい場合においても正しい識別が可能になった。図 9(b) は、縦方向の強いこう配が人の形状に似ているため、従来法では誤検出していたが、提案手法では人に似た背景に対しても頑健に検出が可能となった。この理由として、従来の共起表現法では、画像から得られた low-level 特徴量に基づいて人と人以外の 2 値に識別するため、誤識別した 2 値符号が組み合わせられた場合、最終的な判定に悪影響を及ぼす可能性がある。これに対して、提案手法は実数を入力する Real AdaBoost の弱識別器の出力を演算子によって結合するため、従来の共起表現法よりも悪影響を及ぼす可能性が低いとみられる。

#### 4.4 実験 2

HOG 特徴量の結合する個数による検出性能の違い

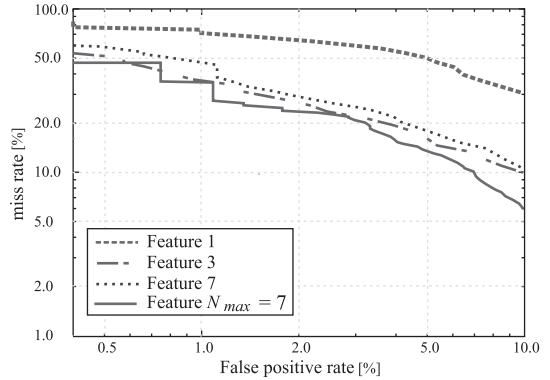


図 10 DET カーブ

Fig. 10 DET curves.

表 3 誤検出率 5.0%における検出率

Table 3 Detection rate with false positive rate 5.0%.

弱識別器に用いる特徴量数	検出率 [%]
1	50.2
3	84.0
7	81.6
$N_{max} = 7$	87.3

を比較する。識別に用いる HOG 特徴量の数が検出精度に影響を与えないように、強識別器に使用する HOG 特徴量の数は 500 個に固定する。表 2 に、学習回数と 1 個の弱識別器に用いる特徴量の数、強識別器に用いる特徴量の数を示す。弱識別器に用いる特徴量が 1 個の場合は共起表現できないため、HOG 特徴量と Real AdaBoost を組み合わせさせた手法となる。また、弱識別器に用いる特徴量数として最適な個数を推定するために hold-out 法により求める手法は、1 個の弱識別器における HOG 特徴量の最大結合個数  $N_{max} = 7$  とする。

図 10 と表 3 に実験結果を示す。実験結果より、全体的に共起表現した手法が良い結果が得られていることが分かる。例えば、 $N_{max} = 7$  とした hold-out 法により最適な結合数を推定した識別器は、共起を表現しない方法よりも、誤検出率 5%において検出率が 37.1%向上している。中でも、hold-out 法により結合させる特徴量の最適な個数を推定する方法が最も精度が高い。結合させる特徴量の個数による違いは、増やすほど検出率が向上するが、7 個結合させると検出率が低下している。この理由は、多数の特徴量を結合させた結果、学習サンプルに過剰適合し、汎化性能が低下したことが考えられる。そのため、hold-out 法によって最適な結合させる特徴量の個数を推定すること

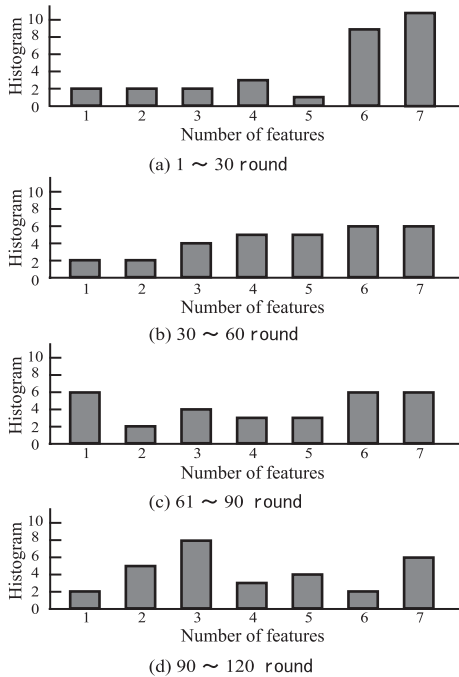


図 11 結合された特徴量の数  
Fig. 11 Number of feature in weak classifier.

が有効であり、提案する共起特徴に対しても適応できていることが分かる。

#### 4.5 考 察

hold-out 法を用いて結合させる特徴量の最適な個数を推定した手法において、学習時に選択された HOG 特徴量から考察を行う。まず、hold-out 法によって結合させる特徴量の最適な個数が、各学習ラウンドにおいてどのように変化しているかに着目する。図 11 に、 $N_{max} = 7$  としたときの学習回数 30 ラウンドごとに選択された HOG 特徴量の結合した個数の分布を示す。図 11 から学習ラウンドの初期では、6 個や 7 個といった多数の HOG 特徴量が結合され、学習ラウンドが進むにつれて HOG 特徴量の結合数が減少していることが分かる。

次に、各学習ラウンドにおいてどのような HOG 特徴量が選択されているかに着目する。図 12 に学習サンプルの平均こう配画像と学習ラウンドの初期に選択された HOG 特徴量を可視化した例を示す。学習ラウンドの初期では、人の輪郭に沿った HOG 特徴量が多数選択されている。例えば、学習ラウンド 1 回目では、人の頭部、手、足部の形状をとらえる HOG 特徴量が選択されている。更に、2 回目では人の左半身、3 回目

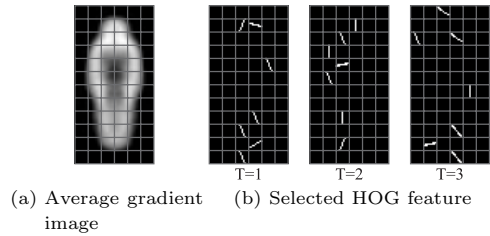


図 12 平均こう配画像と選択された HOG 特徴量の可視化  
Fig. 12 Average gradient image and visualization of selected HOG features.

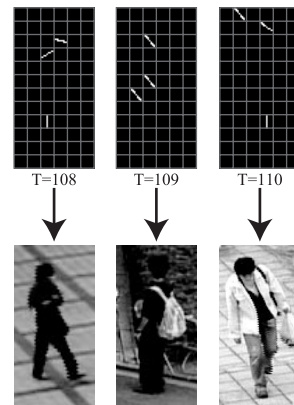


図 13 選択された HOG 特徴量の可視化  
Fig. 13 Visualization of selected HOG features.

では右半身のシルエットをとらえるような HOG 特徴量が選択されている。これは、学習ラウンドの初期では、多数の HOG 特徴量を用いて人の形状をとらえていることから、人の大まかなシルエット形状の情報をとらえることができると考えられる。次に、学習ラウンド終盤に選択された HOG 特徴量に着目する。図 13 に学習ラウンドの終盤 (108 ~ 110 ラウンド) において選択された HOG 特徴量の可視化例とこの弱識別器によって正しい識別が可能な学習サンプルを示す。学習ラウンドの終盤では、少数の HOG 特徴量によって人の輪郭とは異なる位置とこう配の HOG 特徴量が選択されている傾向がある。これらは、学習ラウンドの初期でとらえられていた大域的な情報では得られない、学習サンプルの局所的な情報に着目していると考えられる。これらの学習サンプルは、図 12 のような弱識別器ではとらえにくいため、学習の後半ではこのようなサンプルに対応するために、局所的な情報をとらえられる特徴量が選択されていると考えられる。

#### 4.6 提案手法による人検出の実施例

評価実験では、切り出したデータベースを用いて提



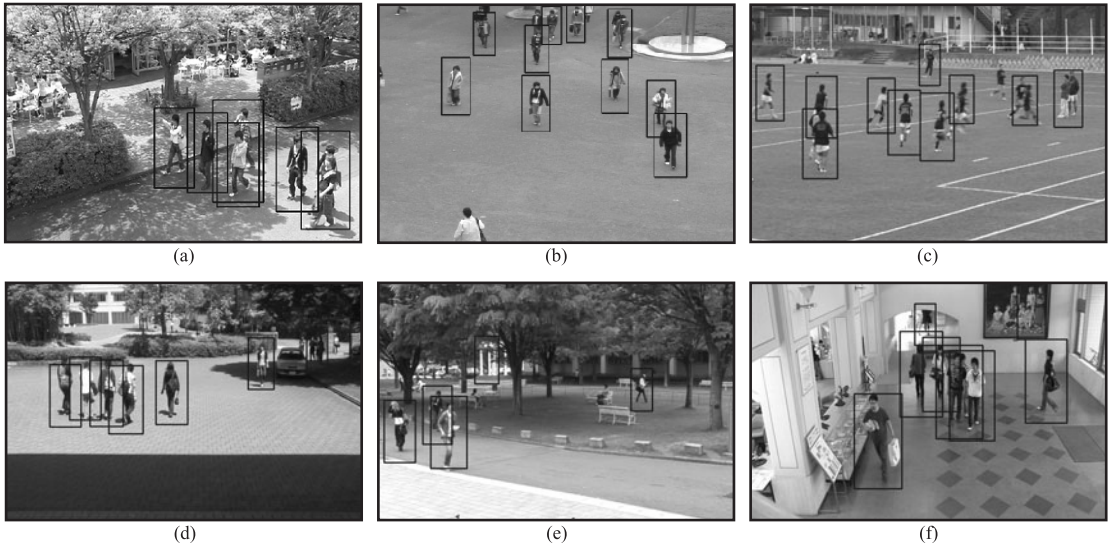


図 14 提案手法による人検出例  
Fig. 14 Examples of pedestrian detection.

案手法の有効性を確認した。本節では、1枚の画像から提案手法による人検出の実施例を示す。人検出を行うには、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 [13]。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。人の大きさや画像中の人同士の重なりによるオクルージョンに対して頑健な検出ができていることが分かる。

## 5. む す び

本論文では、Boosting に基づく特徴量の共起表現による人検出法を提案した。提案する共起特徴は、弱識別器の出力が連続値である Real AdaBoost を利用し、弱識別器の出力を演算子によって結合させることによって算出する。そのため、オクルージョンなどの影響によって low-level な特徴量を用いた識別が間違った場合においても、それ以外の low-level な特徴量が正しい識別を行うことによって、最終的な識別を正しく行うことができるため、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

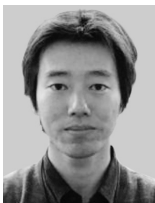
- [1] P. Viola and M. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol.57, no.2, pp.137–154, 2004.
- [2] K. Levi and Y. Weiss, "Learning object detection from a small number of examples: The importance of good features," *IEEE Comput. Vis. Pattern Recognit.*, vol.2, pp.53–60, 2004.
- [3] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Comput. Vis. Pattern Recognit.*, 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 International 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, 藤吉弘直, 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] R.E. Schapire and Y. Singer, "Improved boosting algorithms using confidence-rated predictions," *Mach. Learn.*, no.37, pp.297–336, 1999.
- [9] T. Mita, T. Kaneko, B. Stenger, and O. Hori, "Dis-

- criminative feature co-occurrence selection for object detection,” IEEE Trans. Pattern Anal. Mach. Intell., vol.30, no.7, pp.1257–1269, 2008.
- [10] P. Sabzmeydani and G. Mori, “Detecting pedestrians by learning shapelet features,” IEEE Comput. Vis. Pattern Recognit., pp.1–8, 2007.
- [11] S.D. Stearns, “On selecting features for pattern classifiers,” International Conference on Pattern Recognition, pp.71–75, 1976.
- [12] 驚見和彦, 関真規人, 波部 斉, “物体検出—背景と検出対象のモデリング,” 情処学研報 (CVIM2005), vol.2005, no.88, pp.79–98, 2005.
- [13] D. Comaniciu and P. Meer, “Mean shift analysis and applications,” IEEE International Conference on Computer Vision, pp.1197–1203, 1999.  
(平成 20 年 10 月 10 日受付, 21 年 3 月 2 日再受付)



山内 悠嗣 (学生員)

2007 中部大・工・情報卒・同年より同修士課程に在籍・画像処理, パターン認識に関する研究に従事・情報処理学会会員.



山下 隆義

2002 奈良先端科学技術大学院大学情報科学研究科修士課程了・同年オムロン(株)入社・顔及び人画像センシング技術の研究に従事・IEEE-CS, 情報処理学会各会員.



藤吉 弘巨 (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 中部大学講師を経て 2004 より同大准教授・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・工博・計算機視覚, 動画画像処理, 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・情報処理学会, 電気学会, IEEE 各会員.