

## 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ている。近年、人検出に有効なLow levelな特徴量として、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)[1]が提案された。さらに、検出精度を向上させるために、複数のLow levelな特徴量を2値に識別し、これらの符号を組み合わせることで共起を表現する手法[2]が提案された。しかし、2値に識別した結果を複数組み合わせるため、間違えた結果が含まれていた場合、最終的に得られる特徴量に影響する問題がある。そこで、Real AdaBoost[3]を用いて、連続値を出力する複数の弱識別器と評価値を演算子により結合する共起確率特徴を提案する。この共起確率特徴量は、複数の特徴量を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。

## 2. 共起確率特徴による物体検出

提案手法は、複数のLow level特徴量をBoostingにより組み合わせることで、特徴量間の関連性を表現する共起確率特徴量を生成し、これによる高精度な物体検出を実現する。

### 2.1. Low level特徴：HOG特徴量

Histograms of Oriented Gradients(HOG)特徴量[1]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、照明変化や局所的な幾何学変化に頑健である。まず、各ピクセルにおいて輝度の勾配強度と勾配方向を算出し、セル領域(図1(b))において、勾配方向ヒストグラムを作成する。算出される勾配方向は $0^\circ - 180^\circ$ に変換し、 $20^\circ$ 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。最後にブロック領域(図1(c))毎に特徴量を正規化する。ブロックは、1セルずつ移動させながら特徴量を抽出することにより、3,240個の特徴量が得られる。

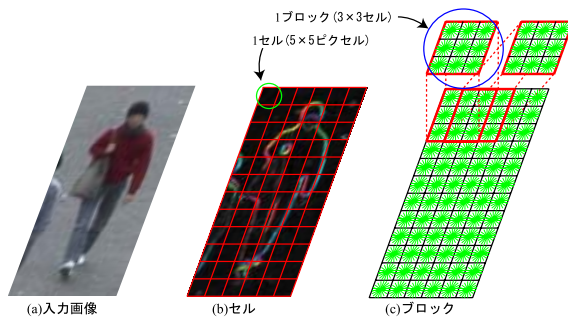


図1: HOG特徴量算出に用いる領域の構造

### 2.2. 共起確率特徴量の生成と学習

提案する共起確率特徴量とReal AdaBoostによる検出器の構築を行う。図2に共起確率特徴量を用いた学習の流れを示す。

1段階目のReal AdaBoostによる特徴Poolの作成  
共起確率特徴量は、複数のHOG特徴量を組み合わせることで生成する。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。

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

$$W_-^j = \sum_{i: j \in J \wedge y_i = -1} D_t(i) \quad (2)$$

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 $\alpha$ を次式より算出する。

$$\alpha = 1 - \sum_j \sqrt{W_+^j W_-^j} \quad (3)$$

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、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 $\alpha$ を用いて、大きい順にソートを行い、 $\alpha$ 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。

Low level特徴の組み合わせによる共起確率特徴量の生成

提案する共起確率特徴量は、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。

共起確率特徴量生成の方法について説明する。まず、Real AdaBoostの弱識別器の出力 $h(x)$ 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、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 $h(x)$ を次式より算出する。

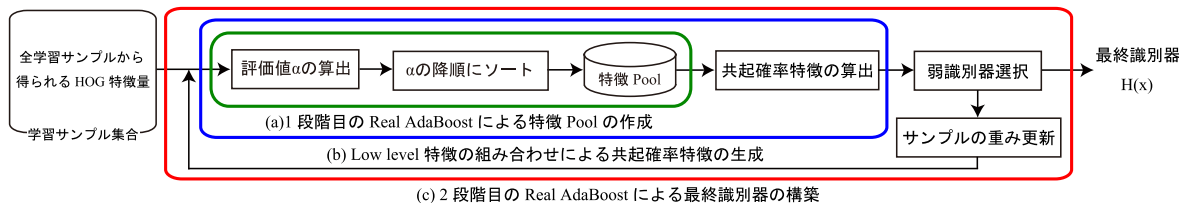
$$h(x) = \frac{1}{2} \ln \frac{W_+ + \epsilon}{W_- + \epsilon} \quad (4)$$

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 $\epsilon = 10^{-6}$ とした。学習サンプルより得られる特徴量 $x$ 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 $h(x)$ と弱識別器の評価値 $\alpha$ を用いて共起確率特徴量を算出する。検出ウィンドウから得られたHOG特徴量を $x$ 、弱識別器の出力を $h1(x)$ 、 $h2(x)$ 、評価値を $\alpha1$ 、 $\alpha2$ とした場合、下記の2パターンの組み合わせより共起確率特徴量 $C1$ 、 $C2$ を次式より得る。

$$C1 = h1(x)\alpha1 + h2(x)\alpha2 \quad (5)$$

$$C2 = h1(x)\alpha1 \times h2(x)\alpha2 \quad (6)$$



(c) 2段階目の Real AdaBoost による最終識別器の構築

図 2：提案手法による学習の流れ

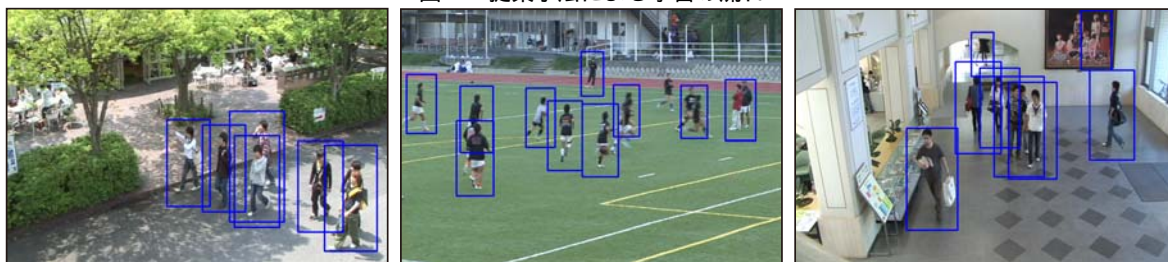


図 3：人検出例

まず、共起確率特徴量  $C1$  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、共起確率特徴量  $C2$  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確率特徴量の値、それ以外は小さい共起確率特徴量の値となる。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、学習の際に算出した評価値を  $\alpha1$ ,  $\alpha2$  から共起確率特徴量  $C1$  もしくは  $C2$  を算出する。特徴 Pool が 200 個の場合、弱識別器の候補は  $200 \times 200 \times 2$  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。

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量を用いて 2 段階目の Real AdaBoost により最終識別器の構築を行う。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。

### 2.3. $N$ 個の特徴量を用いた共起確率特徴量への拡張

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毎に決定する必要がある。そこで、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。hold-out 法は、結合させる特徴量の個数  $N$  を予め決めた上限値  $N_{max}$  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。そして、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。

## 3. 評価実験

本手法の有効性を示すために評価実験を行う。

### 3.1. 実験概要

屋外で撮影した映像から人を画像を切り出してデータベースを作成した。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 枚用意した。比較対象として、HOG 特徴量 [1]、従来の共起表現法 (2 個の Low level 特徴量の組み合わせ)[2]、提案手法により 2 個の Low level 特徴量の組み合わせた共起確率特徴量、提案手法により  $N_{max}$  個の Low level 特徴量の組み合わせた共起確率特徴量 ( $N_{max} = 7$ ) の比較を行う。評価には、Detection Error Tradeoff (DET) カーブによって評価

する。

### 3.2. 実験結果

図 4 に結果を示す。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て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% 向上した。また、Low level 特徴量の最適な個数を推定することで、検出精度が向上した。

提案手法により人検出した例を図 3 に示す。人同士の重なりによるオクルージョンに対しても頑健に検出できていることがわかる。

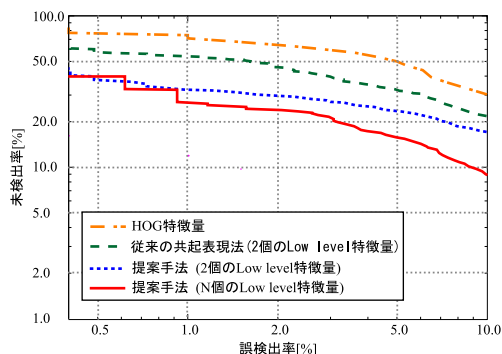


図 4：DET カーブ

## 4. おわりに

Boosting に基づく特徴量の共起表現とこれを用いた人検出法を提案した。今後は、一般物体認識への展開を検討する予定である。

### 参考文献

- [1] N. Dalal *et al.*, "Histograms of Oriented Gradients for Human Detection", CVPR, vol.1, pp. 886-893, 2005.
- [2] T. Mita *et al.*, "Discriminative Feature Co-occurrence Selection for Object Detection", PAMI, vol. 30, no. 7, pp. 1257-1269, 2008.
- [3] R. E. Schapire *et al.*, "Improved Boosting Algorithms Using Confidence-rated Predictions", Machine Learning, No. 37, pp. 297-336, 1999.

### 研究業績

- [1] 山内悠嗣, 藤吉弘巨, Hwang Bon-Woo, 金出武雄, "アピラ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 (MIRU), pp. 1492-1497, 2007.
- [2] Y. Yamauchi and H. Fujiyoshi, "People Detection Based on Co-occurrence of Appearance and Spatiotemporal Features", International Conference on Pattern Recognition, 2008.

(他 学会口頭発表 5 件, 研究会技術報告 1 件)

### 受賞

- [1] MIRU2007 インタラクティブセッション優秀賞
- [2] サマーセミナー 2008 優秀発表賞