複数の特徴量間の関連性に着目したBoostingによる物体検出

山内 悠嗣† 藤吉 弘亘† 山下 隆義†,††
†中部大学大学院 工学研究科 情報工学専攻
††オムロン株式会社 技術本部
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共起特徴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の弱識別器が連続値である出力を，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12.1%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，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[1]に代表されるようにlow-levelな特徴と統計的学習の組み合わせによる手法が多い．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，検出精度を向上させるため，アピアランスとは異なる特徴量の併用[5][6]や特徴量の改良[7]，Boosting手法の改良[8]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，2値の符号を組み合わせるJoint Haar-like特徴[10]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せたが，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[9]の弱識別器の出力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．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．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．

2 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を演算子により結合した共起特徴を提案する．図1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することにより人の検出を行う．検出ウィンドウより得られる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．

2.1 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，各ピクセルの輝度から勾配強度と勾配方向を算出し，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，$0° - 360°$となるが，$0° - 180°$に変換する．これにより，人の服装に対して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

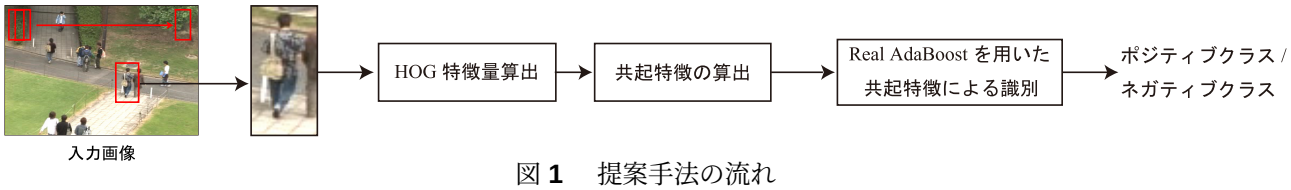

図 **1**　提案手法の流れ

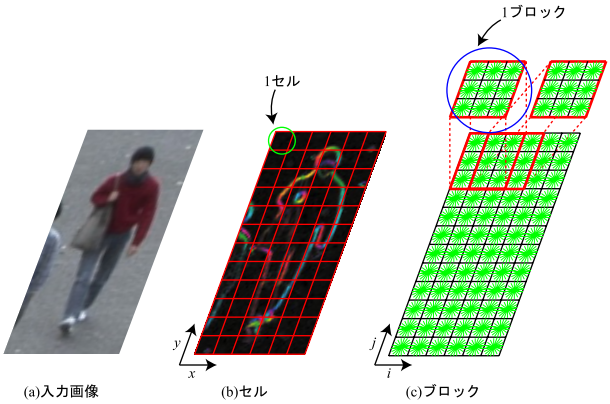

図 **2**　**HOG** 特徴量算出に用いる領域の構造

向の勾配方向ヒストグラムを得る．最後に，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．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．

2.2　共起特徴と Real AdaBoost による識別器構築

HOG 特徴量を用いて共起特徴を生成し，人とそれ以外の判別をする最終識別器の構築を行う．

2.2.1　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．

共起特徴算出の方法について説明する．まず，HOG 特徴量 $\boldsymbol{x}$ を入力とした Real AdaBoost の弱識別器の出力 $h(\boldsymbol{x})$ を算出するために，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．

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

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

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，作成した確率密度分布 $W_\pm$ を用いて，Real AdaBoost の弱識別器の出力 $h(\boldsymbol{x})$ を算出する．学習サンプルより得られる HOG 特徴量 $\boldsymbol{x}$ 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 $h(\boldsymbol{x})$ を次式より算出する．

$$h(\boldsymbol{x}) = \frac{1}{2} \ln \frac{W_+^j + \epsilon}{W_-^j + \epsilon} \tag{3}$$

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．

算出した複数の弱識別器の出力 $h(\boldsymbol{x})$ を用いて共起特徴 C を算出する．次式より N 個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特徴 C を算出する．

$$C = \sum_{n=0}^{N} h_n(\boldsymbol{x}) \tag{4}$$

共起特徴 C は，弱識別器の出力の和であるため，複数の特徴量を総合的に捉える特徴量となる．そのため，部分的な隠れやノイズ等の外乱の影響を受けた場合，HOG 特徴量を入力した場合でも，影響を受けなかった HOG 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．例として，2 個の HOG 特徴量から共起特徴を生成し，最終識別器の弱識別器の出力算出例を図 3 に示す．

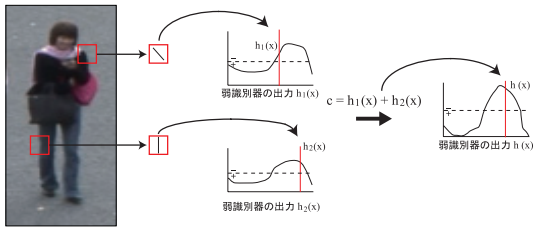

図 **3**　**HOG** 特徴量による共起特徴の算出

2.2.2　特徴量の組み合わせ探索

上記で示した共起特徴は，複数の HOG 特徴量を入力とした Real AdaBoost の弱識別器の出力の和により算出す

る．その際に組み合わされる HOG 特徴量は，最もエラーの少ない特徴量の組み合わせが最適であるが，HOG 特徴量の全組み合わせを評価するためには，学習に膨大な時間が必要となる．本研究では効率的な組み合わせ探索を行う Sequential Forward Selection(SFS)[11] を用いる．この方法は，はじめに最もエラーが少ない特徴量を 1 個選択する．次に，エラーが最も少なくなるように別の特徴量を 1 個追加する．これを N 個の HOG 特徴量を組み合わせる回数繰り返し処理を行うことにより，効率的な特徴の組み合わせを行う．

2.2.3 Real AdaBoost による最終識別器の構築

生成した共起特徴 C を用いて Real AdaBoost[9] により最終識別器の構築を行う．

3 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．

3.1 データベース

文献 [6] 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 4 に学習に用いたサンプルの一部を示す．

図 **4** 学習用サンプルの一部

3.2 実験概要

文献 [6] の評価用データベースを用いて，提案手法 (RJF) と HOG+Real AdaBoost(NR)，特徴量として HOG 特徴量を用いた従来の共起表現法 (BJF)[10] の比較を行う．各手法に用いる特徴量の数は 500 個選択した時点で学習を終了し，各手法においてほぼ同数の特徴量の数を用いて識別を行う．表 1 に，各手法における学習回数 T と 1 個の弱識別器に用いる特徴量の数 N，識別に用いた全特徴量の数を示す．BJF1 及び RJF1 は，1 個の弱識別器が 1 個の特徴量を用いることを表現し，BJF2 及び RJF2 は，1 個の弱識別器が 2 個の特徴量を用いることを表現している．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．

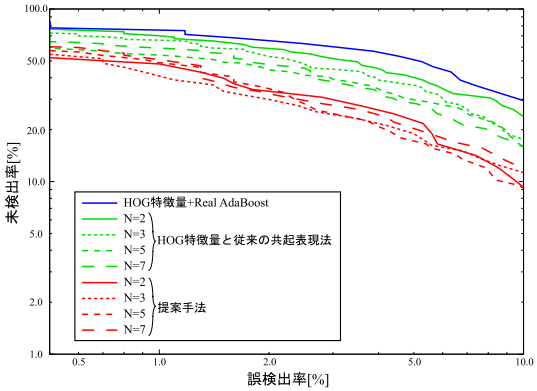

図 **5** DET カーブ

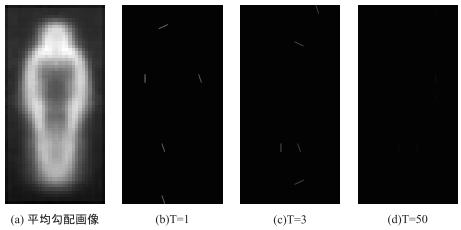

図 **6** 各学習回数において選択された **HOG** 特徴量の可視化例

表 **1** 識別に用いる特徴量の数と学習回数

手法	T	N	全特徴量数
NR	500	1	500
BJF2，RJF2	250	2	500
BJF3，RJF3	167	3	501
BJF5，RJF5	100	5	500
BJF7，RJF7	72	7	504

3.3 実験

図 5 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．そのため，HOG 特徴量+Real AdaBoost よりも共起を表現する手法の方が検出性能が高いことがわかる．共起を表現する方法においても，提案手法が従来の共起を表現する方法に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%における未検出率の比較を行う．提案手法 $N = 5$ は，従来の共起表現法 $N = 7$ よりも検出率が 12.1%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

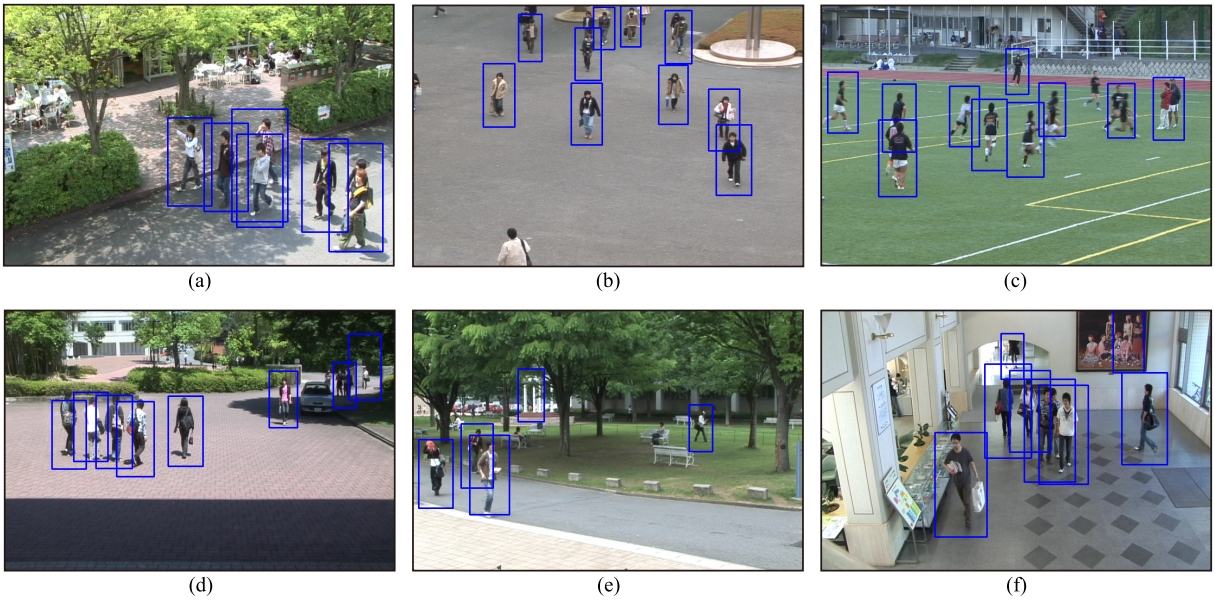

図 **7**　提案手法による人検出例

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 6(a) に全学習ポジティブクラスの平均勾配画像，図 6(a)，(b)，(c) に弱識別器 ($N = 5$) により各学習回数において選択された HOG 特徴量の可視化例を示す．輝度が高いほど Real AdaBoost において識別に有効な特徴量として選択されたことを表す．学習回数 1 回 (図 6(b)) では，人の輪郭に沿った HOG 特徴量が選択されており，学習回数 3 回 (図 6(b))，学習回数 50 回 (図 6(c)) と比較して識別に有効な特徴量として選択されていることがわかる．

3.4　ラスタスキャン方式による人検出

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 [13]．図 7 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 7 から高精度な人検出を行っていることがわかる．図 7(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 7(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うためである．

4　おわりに

本稿では，Boosting に基づく特徴量の共起表現法と共起特徴を用いた人検出法を提案した．提案する共起特徴は，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，多クラスの識別を検討する．

参考文献

[1] P. Viola and M. Jones, “Robust Real-Time Face Detection”, Trans. IJCV, 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.”,Proc. CVPR, vol. 2, pp. 53-60, 2004.

[3] N. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection”, Proc. CVPR, 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”, Proc. ICCV, vol. 1, pp. 90-97, 2005.

[5] P. Viola, M. Jones and D. Snow, “Detecting Pedestrians Using Patterns of Motion and Appearance”, Proc. ICCV, pp. 734-741, 2003.

[6] 山内悠嗣, 藤吉弘亘, Hwang Bon-Woo, 金出武雄, “アピア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.

[7] C. Hou, H. Ai and S. Lao,“Multiview Pedestrian Detection Based on Vector Boosting”, Proc. ACCV, vol. 1, pp. 220-229, 2007.

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

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

[10] T. Mita, T. Kaneko, B. Stenger, O. Hori, “Discriminative Feature Co-occurrence Selection for Object Detection”, IEEE Trans. PAMI, vol. 30, no. 7, pp. 1257-1269, 2008.

[11] S. D. Strearns, “On Selecting Features for Pattern Classifiers”, Proc. ICPR, pp.71-75, 1976.

[12] 鷲見和彦, 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 CVIM 研報, Vol.2005, No.88,pp. 79-98, 2005.

[13] D. Comaniciu and P. Meer, “Mean shift analysis and applications”, Proc. ICCV, pp. 1197-1203, 1999.