

1. はじめに

近年, HOG 特徴量などの局所特徴量と統計的学習手法を組み合わせ、物体検出に関する研究が盛んに取り組まれている。この技術は、監視、防犯、車載、計測など多岐の市場において実用化される段階にある。しかし、実環境において、カメラと人との間の障害物により人体の一部の情報欠落するため、検出精度が低下するという問題がある。そこで、本研究では遮蔽を考慮した Randomized Trees によるパーツベースの物体検出法を提案する。提案手法は、一部のパーツに遮蔽が発生しても、他のパーツを用いての識別が可能となる。

2. Randomized Trees

Randomized Trees[1] は、マルチクラスの識別に対応したアンサンブル学習法である。木構造のフレームワークであるため、高速な学習と識別が可能である。また、ランダム学習により、ノイズに対して強い手法であることから、キーポイント検出やセグメンテーション、物体検出の分野など実用的な問題にも利用され始めている。

2.1 学習

学習は、サブセット作成/分岐ノードの選択/サブセットの分割による 3 つのステップから、複数の決定木を作成する。サブセットの作成では、ラベル付きサンプルの集合からランダムで選択し、 T 個のサブセットに分ける。作成された T 個のサブセットを用いて、それぞれに対し木を作成する。この時、サブセット間でのサンプルの重複を許容するように作成する。これにより、過学習の抑制や、学習の効率化といった効果が得られる。分岐ノードでは、入力されたサンプルが条件を満たせば右、満たさなければ左へ分岐する分岐関数を定義する。ここで、選択される特徴量はランダムで用意した M 次元の中から情報利得 ΔE を求めることで、最適な分岐能力を持つ特徴量を選択する。選択された特徴量を用いてサブセットを左右の子ノードに分割し、上記の処理を一定のノードの深さに到達するまで繰り返すことで木を作成する。

2.2 識別

識別では、未知サンプルを各決定木に入力し、到達した末端ノードが持つ事後確率の平均を算出する。各クラスごとの平均事後確率から、最大となるクラスを識別結果として出力とする。

3. 提案手法

提案手法では、検出対象である人の領域を頭、両腕、両脚、足の 6 個のパーツに分割し、それらを各パーツのラベルが付いたサンプルとして扱い、マルチクラス識別器を構築する。これにより、遮蔽が発生してパーツの一部が隠れている場合でも、他のパーツが人らしさをとらえ、識別が可能となる。以下より提案手法によるパーツベースへの拡張と、学習、識別について述べる。

3.1 パーツベースへの拡張

提案手法では、ポジティブサンプルを N 個の領域に分割し、それぞれを独立したクラスのサンプルとして扱う。ここで、分割された各領域から算出される特徴次元数を等しくするため、分割領域の面積は固定する。提案手法では、 8×8 ピクセルを 1 セルとした、12 セルを 1 つのパーツとして定義する。また、パーツの分割領域数は $N = 6$ とし、学習サンプルの平均勾配画像から人体が各領域でとらえられる領域を定義した。

3.2 学習

提案手法による学習の流れを図 1 に示す。図 1(a) のように分割したそれぞれのパーツを学習サンプルとし、 $c_0 \sim c_{N-1}$ のクラスを与える。ネガティブサンプルについては、全てのパーツにネガティブクラスとして c_N のクラスを与える。これらのサンプルをランダムで選択し、 T 個のサブセットに分ける (図 1(b))。このサブセットを用いて、Randomized Trees のアルゴリズムに従いマルチクラス識

別器を構築する。決定木の各末端ノードは 7 クラスの事後確率を持つこととなる (図 1(c))。

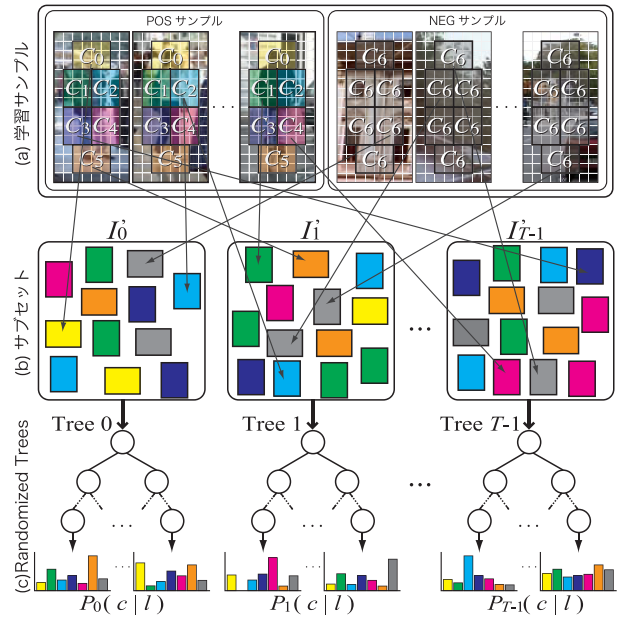


図 1: 学習の流れ

3.3 識別

識別の流れを図 2 に示す。入力された未知サンプルの各パーツを Randomized Trees 識別器に入力する。各パーツの出力値 h_n は (1) 式より求める。

$$h_n = \frac{1}{T} \sum_{t=0}^{T-1} (P_t(c_n|I_t) - P_t(c_N|I_t)) \quad (1)$$

ここで、 n は入力のパーツの番号、 t は木の数を示している。到達した末端ノードのヒストグラムから、 n 番目のパーツのクラス c_n と、ネガティブクラス c_N の確率の平均の差を算出する。最終出力値 H は、(2) 式より各パーツの出力値の平均により算出される。

$$H = \frac{1}{N} \sum_{n=0}^{N-1} h_n \quad (2)$$

H は、-1.0 から 1.0 までの実数値を出力し、正の値であるほど人らしいことを示す。これにより、各パーツにおけるマルチクラス識別器の応答を 2 クラス問題として扱う。

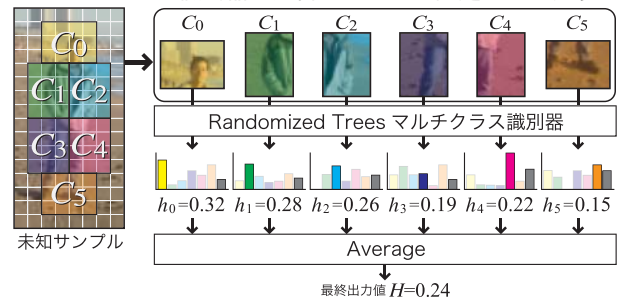


図 2: 識別の流れ

3.4 分岐関数

本研究では、物体検出に有効な勾配方向ヒストグラムを特徴量として決定木の分岐を行う。分岐関数には通常のしきい値関数 (図 3(a)) と、比較関数 (図 3(b)) を用いる。しきい値関数では、ランダムで選択した特徴量 $g(i)$ としきい値 t を用いることで、勾配の強度をとらえる。比較関数では、ランダムで選択した 2 つの特徴量 $g_1(i), g_2(j)$ の大小関係を用いることで、局所領域間における勾配の関係性をとらえる。ここで、 g は各セルから算出される勾配方向ヒストグラム、 i, j は勾配方向ヒストグラムの要素番号を示す。本研究では、情報利得 ΔE によりどちらか最適な関数を各ノードで決定する。

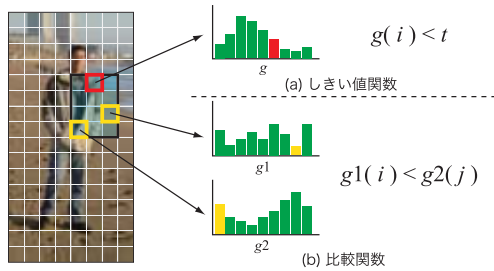


図 3: 分岐関数

4. 評価実験

提案手法の有効性を示すために 2 つの識別実験を行う。

4.1 実験概要

本研究では、従来法との基本性能の比較する実験 1 と、実験 1 と同様の識別器を用いて、擬似的に遮蔽物を発生させたサンプルの評価し、遮蔽に対する頑健性を評価する実験 2 の 2 つの実験を行う。擬似遮蔽物は、提案手法における各パーツの中から 1 つパーツ領域をランダムで選択し、矩形を評価サンプル画像に描画する。これにより、遮蔽が発生した環境を想定した実験が可能である。評価手法は、提案手法 (勾配方向ヒストグラム + パーツベース Randomized Trees), 勾配方向ヒストグラム + Randomized Trees, HOG + Real AdaBoost の 3 手法である。学習パラメータは、全ての手法に対して、入力画像を 64×128 ピクセル, セルサイズを 8×8 ピクセル, 勾配方向ヒストグラムの要素数を 9 とした。また比較手法である HOG 特徴量のブロックサイズは 2×2 セルとした。

データセットは、人検出のベンチマークで用いられる INRIA Person Dataset を使用する。学習用サンプルとして、ポジティブサンプルが 2,146 枚, ネガティブサンプルが 13,156 枚, 評価用サンプルとして、ポジティブサンプルが 1,176 枚, ネガティブサンプルが 453 枚の背景画像からスライディングウィンドウにより取得したものを使用する。また、実験 2 で使用する評価用ポジティブサンプルは、全てに擬似遮蔽物を発生させた。

4.2 結果

図 4(a) に実験 1, 図 4(b) に実験 2 の結果を示す。結果は横軸を誤検出率, 縦軸を未検出率とする ROC カーブで評価する。図 4(a) より, パーツベースの識別をする提案手法は, 従来の Randomized Trees と比較し, 誤検出率 10% において検出率が約 4% 向上した。図 4(b) の遮蔽サンプルに対する評価より, 提案手法は, 従来の Randomized Trees と比較し, 誤検出率 10% において検出率が約 15% 向上した。さらに, HOG + Real AdaBoost と比較し, 検出率が約 3% 向上した。

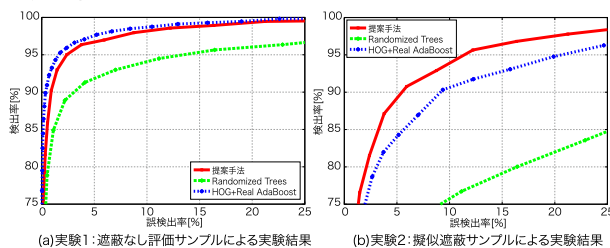


図 4: 実験結果の ROC カーブ

4.3 考察

遮蔽サンプルに対する提案手法の効果を確認するため, 遮蔽の有無による提案手法と HOG+Real AdaBoost における最終識別器の出力値の比較を図 5 に示す。HOG+Real AdaBoost では遮蔽の発生により負の出力値となるのに対し, 提案手法では正の出力値を保っている。これは, Randomized Trees のノイズに対する頑健性と, パーツごとの識別による効果であると考えられる。

次に, 1 つのサンプルを識別するための処理時間と, 各識別器において必要なメモリ量を表 1 に示す。処理時間は, 全評価サンプルから特徴抽出と識別をした時の平均処理時間を算出した。提案手法は, HOG+Real AdaBoost と同等の処理時間で識別が可能となる。メモリ量の比較では,

サンプル				
提案手法	遮蔽なし	0.230	0.136	0.108
	遮蔽あり	0.122	0.098	0.089
HOG+RealAdaBoost	遮蔽なし	17.733	8.733	-2.793
	遮蔽あり	-0.781	-1.689	-5.351

図 5: 遮蔽の有無による出力値の違い

各識別器における必要な情報を格納するためのメモリ量を示している。7 クラスの情報を格納している提案手法は 2 クラスである Randomized Trees と比較するとメモリ量が約 2 倍増加している。しかし, HOG+Real AdaBoost と比較すると約 1/4 倍のメモリ量であることがわかる。これらの処理時間とメモリ量から, 提案手法は FPGA などの小規模ハードウェア化を想定した場合でも有効であると言える。

表 1: 1 サンプルの処理時間と必要なメモリ量

手法	処理時間 [ms]	メモリ [KB]
提案手法	1.383	120
Randomized Trees	1.161	56
HOG+Real AdaBoost	1.397	513

5. 実遮蔽データを用いた評価実験

実環境において遮蔽が発生した場合の性能を評価するため, 実遮蔽データセットを用いた評価実験を行う。実遮蔽データは, 車載カメラからの歩行者検出システムを想定したものであり, 実際の車両に取り付けられたカメラから歩行者を撮影したものである。実遮蔽データセットは, 学習用サンプルとして, ポジティブサンプルが 1,948 枚, ネガティブサンプルが 8,067 枚, 評価用サンプルとして, ポジティブサンプルが 614 枚, ネガティブサンプルが 5,250 枚を使用する。実験結果の図 6(a) より, 提案手法は従来の Randomized Trees と比較し, 誤検出率 10% において検出率が約 17% 向上した。さらに, HOG + Real AdaBoost と比較し, 検出率が約 6% 向上し, 検出率約 100% を達成した。また, 図 6(b) に提案手法, 図 6(c) に HOG+Real AdaBoost による検出例を示す。提案手法は, 人体の一部の情報が欠落しても正しく人として検出できている。

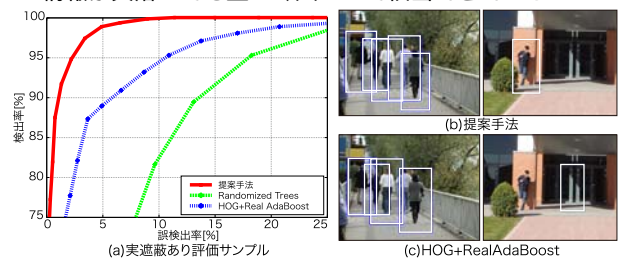


図 6: 実遮蔽データによる ROC カーブと検出例

6. おわりに

実環境下で起こりうる遮蔽問題による検出率低下を抑制するための手法として, Randomized Trees を用いたパーツベースの物体検出法を提案した。物体の領域を N 個のパーツに分割し, Randomized Trees によりマルチクラス識別器を学習することで, 擬似遮蔽データに対する識別率の低下を抑制する効果があり, 実遮蔽データに対しても高精度な識別が可能であることを確認した。

参考文献

[1] L. Breiman, "Random forests", Machine Learning, No.45(1), pp. 5-32, 2001.

研究業績

[1] 三井相和, 山内悠嗣, 藤吉弘巨, "Joint 特徴量を用いた 2 段階 Boosting による物体検出", 電子情報通信学会論文誌, Vol. J92-D, No. 9, pp. 1591-1601, 2009.
(他 国際会議論文 1 件, 研究会技術報告 4 件, 学会口頭発表 2 件, 解説論文 1 件)

受賞

[1] MIRU2009 インタラクティブセッション賞