

### 1. はじめに

人検出は、映像監視やマーケティング分野などでの利用が期待されている。従来、人検出で用いられている Histograms of Oriented Gradients(HOG) 特徴量 [1] は、単一の特徴量であるため識別に困難なパターンが存在する。本研究では、アピランスと時空間特徴を用いた Joint 特徴量による人検出法を提案し、その有効性を示す。

### 2. Joint 特徴量を用いた 2 段階 Boosting

提案手法は、low-level 特徴量を組み合わせた Joint 特徴量と、2 段階に構築した Boosting により人の検出を行う。Boosting には弱識別器の出力値が連続値で得られる Real AdaBoost を用いる。図 1 に提案手法の流れを示す。low-level 特徴量には、アピランス特徴である HOG 特徴量 [1] と時空間特徴である PSA 特徴量 [2] を用いる。

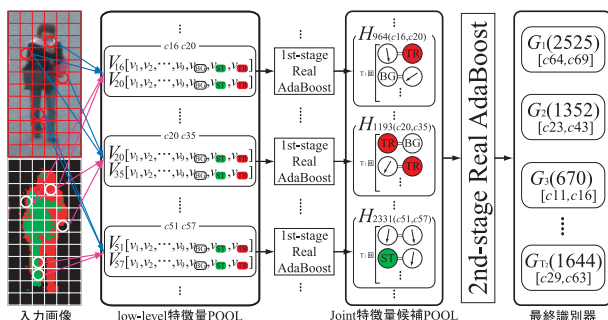


図 1：提案手法の流れ

**アピランス特徴：HOG 特徴量** HOG 特徴量は物体の形状を表すことが可能な特徴量である。5×5 ピクセルで構成されたセル (局所領域) 内での 9 方向の勾配方向をヒストグラム化し、ブロック毎に正規化したものを特徴量として抽出する (図 2(a))。近接画素の勾配をヒストグラム化するため、照明変動や局所的な幾何学変化の影響を受けにくい特長がある。

**時空間特徴：PSA 特徴量** PSA 特徴量とは、ピクセル状態分析 (Pixel State Analysis : PSA)[2] の結果を用い、HOG と同様のセル領域からピクセル状態ヒストグラムを作成することにより算出する (図 2(b))。ピクセル状態分析とは、各ピクセルの時間変化をモデル化し、背景、静状態、動状態の 3 状態に判別する手法である。3 状態に判別することで人らしい動きの情報を捉える事が可能となる。

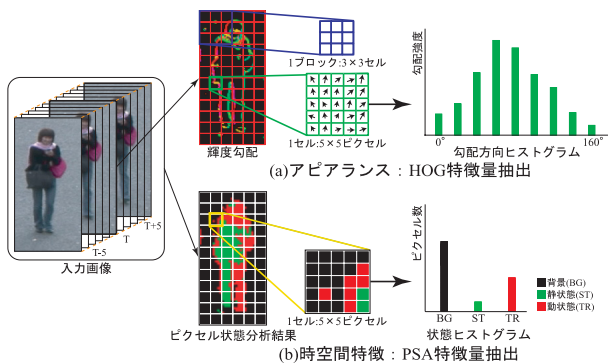


図 2：low-level 特徴量の抽出

本手法は 2 段階に Real AdaBoost を構築する。まず、1 段階目の Real AdaBoost により、全てのセルの組み合わせにおいて、異なる 2 つのセルに含まれる low-level 特徴量を組み合わせた Joint 特徴量の候補プールを生成する。2 つの low-level 特徴量を観測できるため、単一の low-level 特徴量のみでは識別困難なパターンも捉えることが可能となる。次に、生成された Joint 特徴量の候補プールを入力とした 2 段階目の Real AdaBoost により、識別に有効な Joint 特徴量を選択し、最終識別器を構築する。

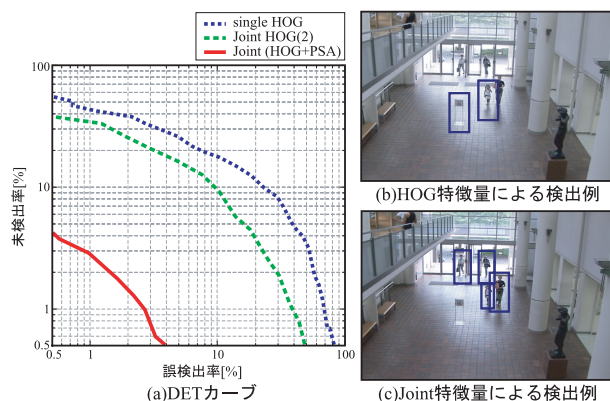


図 3：実験結果

### 3. 評価実験

複数の映像から歩行者を切り出した画像を用いて識別精度の評価を行う。学習用として、ポジティブサンプル 2,054 枚、ネガティブサンプル 6,258 枚、評価用として、ポジティブサンプル 1,000 枚、ネガティブサンプル 1,234 枚用いる。評価には DET カーブを用いる。DET カーブは原点に近いほど精度が良いことを示す。

識別実験結果の DET カーブを図 3(a) に示す。提案手法である HOG 特徴量と PSA 特徴量を組み合わせた Joint 特徴量は、従来法である単一の HOG 特徴量と比較して、誤検出率 1% のとき、39.2% 検出率を向上させることができた。図 3(b), (c) に検出例を示す。図 3(b) の中央で発生した誤検出が、図 3(c) では抑制されている。HOG 特徴量を持つアピランスの情報に動きの情報を加えることで、人に似た形状や複雑な背景に対しても高精度な人検出が可能となった。

図 4 に、学習ラウンド毎に選択された HOG 特徴量と PSA 特徴量の割合と選択された特徴量を示す。学習の初期では PSA 特徴量が多く選択され、約 5 ラウンド以降では HOG 特徴量が多く選択されている。これは、識別時において、まず物体の動きを表す PSA 特徴量により大まかに人か人以外かを判別した後、詳細に識別境界を形成するために HOG 特徴量を選択されたと考えられる。

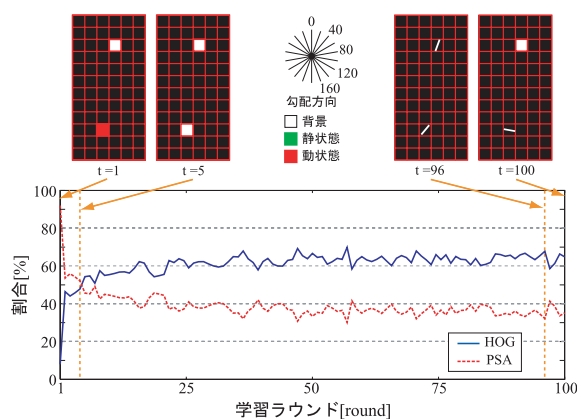


図 4：選択された特徴量の割合

### 4. おわりに

本研究では、アピランスと時空間特徴を用いた Joint 特徴量による人検出法を提案し、高精度な検出が可能であることを示した。今後は、車載カメラからの人検出を行う予定である。

#### 参考文献

[1] N. Dalal *et al.*, "Histograms of Oriented Gradients for Human Detection", CVPR, pp. 886-893, (2005).  
 [2] H. Fujiyoshi *et al.*, "Layered Detection for Multiple Overlapping Objects", IEICE Trans. Inf. & Syst., vol. E87-D, pp. 2821-2827, (2004).