

1. はじめに

近年、セキュリティや ITS 分野において、安心安全のためにカメラ画像からの人検出が期待されている。従来、人検出器における特徴量には、アピアランスに基づく特徴量が用いられる。近年では、モーションに基づく特徴量を併用することによって検出精度を向上させている。しかし、人が静止状態にある場合、モーションから得られる特徴量が減少するため、検出精度は低下する。そこで、本研究では、従来用いられてきたアピアランスに、時空間特徴を加えた、人の状態変化に不変で高精度な人検出法を提案する。

2. アピアランスと時空間特徴に基づく人検出

提案手法 (図 1) は、アピアランスに基づく特徴として HOG (Histograms of Oriented Gradients)[1], 時空間特徴として、ピクセル状態分析 [2] の結果を特徴ベクトルとして算出する。算出された特徴ベクトルを主成分分析により次元圧縮を行い、カスケード型 AdaBoost を用いて検出する。

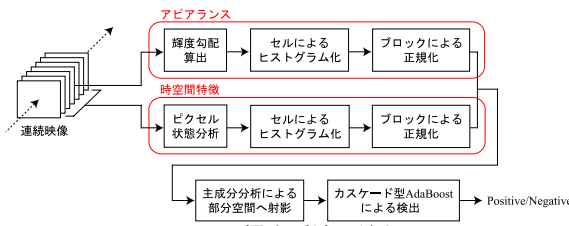


図 1: 提案手法の流れ

3. 特徴量抽出

アピアランスに基づく特徴 アピアランスに基づく特徴として、HOG を用いる。HOG は、局所領域において輝度の勾配方向をヒストグラム化した特徴ベクトルであり、人のきめ細かな輪郭情報を表すことができる。セル領域 (5 × 5 ピクセル) において求めた 9 方向の勾配方向ヒストグラムは、ブロック (3 × 3 セル) 毎に正規化を行う。入力画像を 30 × 60 ピクセルとした場合、正規化後の特徴ベクトルは、3,240 次元となる。図 2(a) に HOG の抽出例を示す。

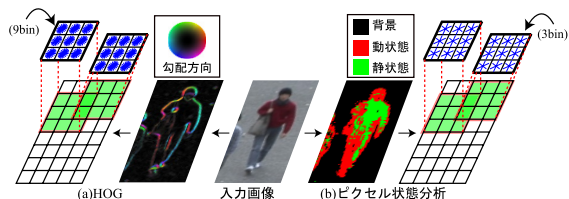


図 2: 提案手法で用いる特徴量

時空間特徴に基づく特徴 時空間特徴に基づく特徴として、ピクセル状態分析を用いる。ピクセル状態分析とは、ピクセルの時間変化をモデル化することにより、各ピクセルを背景、動状態、静状態の 3 状態に判別する。HOG と同様に、セル領域においてピクセルの状態をヒストグラム化し、ブロック毎に正規化する。このとき、特徴ベクトルは 1,080 次元となる。図 2(b) にピクセル状態分析に基づく特徴ベクトル抽出例を示す。

主成分分析による次元圧縮 アピアランス (3,240 次元) と時空間特徴 (1,080 次元) から得られる 4,320 次元の特徴ベクトルを主成分分析により次元圧縮することで、特徴ベクトル間の共起性を表現する。Positive クラス (人画像クラス), Negative クラス (人画像クラス外) の各クラス毎に次元圧縮を行う。本研究では、累積寄与率 99% までの成分を用いることにより、Positive クラスは 765 次元、Negative クラスは 1,265 次元となる。未知入力パターン \mathbf{x} は、2 クラスから得られる射影行列 S_p, S_n を用いて射影され、新たに得られる特徴量 $\mathbf{v} = [v_1, v_2, \dots, v_{m_1+m_2}]$ は式 (1), (2) より算出される。

$$[v_1, \dots, v_{m_1}] = \mathbf{x}^T S_p \quad (1)$$

$$[v_{m_1+1}, \dots, v_{m_1+m_2}] = \mathbf{x}^T S_n \quad (2)$$

4. 識別器の構築

カスケード型 AdaBoost AdaBoost は、複数の弱識別器を重みつき多数決で識別する手法である。識別器をカスケードに構成することで、Negative クラスの多くを最後の識別器まで到達させず、Positive クラスを通過させることができる。これにより FalsePositive の減少が期待でき、かつ計算時間を減少させることが可能である。弱識別器 カスケード型 AdaBoost に用いる弱識別器にはベイズ識別を用いる。式 (3) に示すベイズの定理により、入力画像が Positive クラスに属する確率 $P(C_p|v_i)$, Negative クラスに属する確率 $P(C_n|v_i)$ を算出する。

$$P(C_k|v_i) = \frac{P(v_i|C_k)P(C_k)}{P(v_i)} \quad (k = p, n) \quad (3)$$

$P(v_i|C_k)$ は確率密度分布から得られる確率、 $P(C_k)$ は事前確率、 $P(v_i)$ は出現確率、 C はクラスを表す。事前確率は両クラス共に 0.5 を与える。確率密度分布は、全学習サンプルより得られる特徴量をヒストグラム化したものを用いる。ベイズの定理より算出された確率が式 (4) を満たすとき弱識別器は v_i を Positive クラスと識別する。

$$P(C_p|v_i) > P(C_n|v_i) \quad (4)$$

5. 評価実験

評価実験として、従来手法である HOG を用いた検出法 (HOG) とピクセル状態分析に基づく検出法 (PSA), 提案手法である HOG とピクセル状態に基づく検出法 (HOG+PSA) の精度の比較を行う。評価には、屋外で撮影された 1,800 フレームの画像を用いる。検出器のしきい値を変化させたときの未検出率と誤検出率からなる Detection Error Trade-off (DET) による比較を行う。

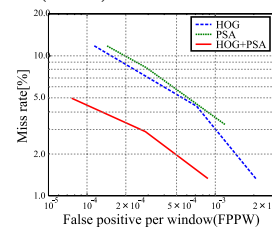


図 3: DET カーブ

図 3 に HOG, PSA, HOG + PSA の DET カーブを示す。提案手法が、どの誤検出率においても精度の良い結果を得ていることがわかる。表 1 は、 2×10^{-4} FPPW における未検出率の比較である。表 1 から、従来法と比較して、HOG に基づく特徴量に、PSA を組み合わせることで、 2×10^{-4} FPPW において未検出率を 5.1% 低下させることができた。図 4 に人の検出例を示す。従来法 (図 4(a)) では未検出であった検出対象を、提案手法 (図 4(b)) は検出可能であることがわかる。

表 1: 未検出率による比較

	未検出率 [%] 2×10^{-4} FPPW
HOG	8.4
PSA	9.9
HOG+PSA	3.3



図 4: 人検出結果

6. おわりに

本研究では、アピアランスに時空間特徴を加えた特徴量に基づく人検出法を提案し、評価実験によりその有効性を確認した。今後は、アクティブカメラやカメラワークへの対応の検討を行う予定である。

参考文献

[1] N. Dalal et al.: "Histograms of Oriented Gradients for Human Detection", Proc. of CVPR, vol. II, pp. 886-893, 2005.
 [2] H. Fujiyoshi et al.: "Layered Detection for Multiple Overlapping Objects", Proc. of IEICE Transactions on Information and Systems, vol. E87-D, pp. 2821-2827, 2004.