

近似 texton 特徴量を用いた高速な人検出法

三井 相和^{†a)} 藤吉 弘亘^{†b)} 秋田 時彦^{††c)}

Fast Human Detection Method Using Approximated Texton Features

Tomokazu MITSUI^{†a)}, Hironobu FUJIYOSHI^{†b)}, and Tokihiko AKITA^{††c)}

あらまし 本論文では, texton に用いられるフィルタをボックスフィルタ (平均化フィルタ) の組合せで近似した, 近似 texton 特徴量による高速な人検出法を提案する. texton による特徴抽出は処理が容易であることから, テクスチャ解析やパターン認識, 一般物体認識など多岐の分野で多く用いられる. しかし, 単純なフィルタリング処理では, フィルタ内の 1 ピクセルごとにアクセスし演算しなければならないため, 特徴抽出の処理コストが非常に高い. そこで, 本論文では特徴抽出の処理コストを削減するために, texton に用いられるフィルタを近似することで高速な人検出を実現する. texton の近似には処理が簡易なボックスフィルタを用いる. このボックスフィルタをピラミッド上に組み合わせることで, 擬似的に texton のフィルタの形状を構築する. また, ボックスフィルタの応答値の算出には, 方形領域の画素値の合計を一括算出することが可能な Integral Image を用いる. これらを組み合わせることで, 近似フィルタの応答値を高速に求めることが可能となる. 評価実験より, 提案手法は人検出の特徴量として有効である Histograms of Oriented Gradients (HOG 特徴量) と同精度であり, 2.3 倍高速な識別が可能であることを確認した.

キーワード 人検出, 局所特徴量, Real AdaBoost, 近似 texton 特徴量, Integral Image

1. ま え が き

デジタルカメラや車載カメラの普及に伴い, 撮影画像の画質改善や運転者のアシストを目的とした人検出の実用化が期待され, FPGA などによるハードウェア化の研究が行われている [1]~[5]. ハードウェア化においては, 高精度かつ高速に動作する人検出アルゴリズムの実現が重要である.

近年で提案されている人検出法は, 画像局所特徴量と統計的学習手法の組合せによる手法 [6]~[10] が多い. 特徴量には, 画像の局所領域のこう配に着目した Histograms of Oriented Gradients (HOG 特徴量) [11] などの特徴量が用いられている [12], [13]. これらの特徴量は, 人の形状特徴を捉えることが可能であるが, 正規化などの処理があるため計算コストがか

かるという問題がある. この問題を解決するためには, いかに画像局所特徴量を簡易に計算するかが重要である. 画像局所特徴量の高速化のアプローチには, Haar-like 特徴量 [14] や SURF (Speeded Up Robust Features) [15] などが挙げられる. Haar-like 特徴量は, 対になる領域の輝度差を特徴量とする手法であり, 高速かつ高精度な顔検出が可能である. SURF では, キーポイントの算出をボックスフィルタにより近似することで高速化を実現している. これらの手法は, 方形領域の画素値の合計を一括算出することが可能な Integral Image [14] を用いることで高速な処理を実現している. 同様に, HOG 特徴量でも, Integral Image をヒストグラムへ応用した Integral Histogram [16] を用いることで高速化が可能である. しかし, 局所領域における正規化が必須となる HOG 特徴量は, ラスタスキャンによる検出で計算コストが増加し, より高速な検出は困難である.

そこで, 我々は単純なフィルタリング処理で特徴量を算出できる texton [17] に着目する. texton は局所領域における様々な特徴量を抽出することが可能であり, また微分フィルタによるエッジ抽出が可能であることから, 人検出への応用が期待できる. しかし, 通

[†] 中部大学大学院工学研究科, 春日井市

Dept. of Computer Science, Chubu Univ., 1200 Matsumoto, Kasugai-shi, 487-8501 Japan

^{††} アイシン精機株式会社, 刈谷市

AISIN SEIKI Co., Ltd., 2-1 Asahi-machi, Kariya-shi, 448-8650 Japan

a) E-mail: mitsui@vision.cs.chubu.ac.jp

b) E-mail: hf@cs.chubu.ac.jp

c) E-mail: akita@rd.aisin.co.jp

常のフィルタリング処理は処理コストが非常に高く、高速な人検出は実現できない。そこで、本論文では高速かつ高精度な人検出手法として、近似 texton 特徴量による人検出法を提案する。texton で用いられるフィルタバンクの各フィルタを、単数または複数のボックスフィルタを組み合わせることで近似する。近似した texton は Integral Image によりボックスフィルタの領域を一括で計算することが可能となり、高速な特徴抽出が期待できる。また、ボックスフィルタ数を変更した全ての特徴量を特徴プールに格納することで、人の識別に有効なボックスフィルタ数を Boosting により選択する。

本論文では、2. で texton 特徴量、3. で近似 texton 特徴量、4. で Integral Image を用いた特徴抽出について述べる。5. で近似 texton 特徴量を用いた学習法について述べ、6. で従来の HOG 特徴量、texton 特徴量と提案手法の比較実験と考察をする。

2. texton 特徴量

texton とは、ガウスフィルタ、LoG フィルタ、微分フィルタなどの様々な方向、サイズによるフィルタの応答を特徴量とするものである。このフィルタの集合をフィルタバンクと呼ぶ。texton は主にテクスチャの解析 [17] をするための特徴抽出法であり、類似画像検索 [18], [19] や、一般物体認識 [20] などに幅広く用いられている。また、フィルタバンクには、微分フィルタや LoG フィルタによりエッジ抽出が可能であることから、人検出においても有効な特徴量であると考えられる。本論文では CIE $L^*a^*b^*$ 色空間に対し、ガウスフィルタから 9 次元、LoG フィルタから 4 次元、ガウス 1 次微分フィルタから 8 次元の計 21 次元を用いる。図 1 に、本論文で利用するフィルタバンクの各フィルタを示す。

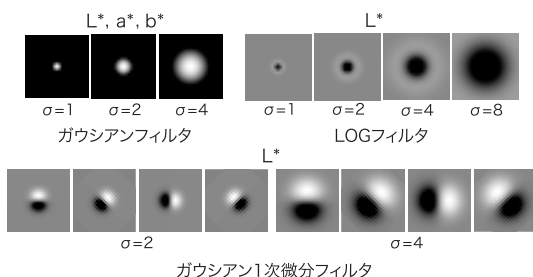


図 1 フィルタバンク

Fig. 1 Filter bank.

2.1 ガウスフィルタ

ガウスフィルタは平均化フィルタにガウス分布の重みを考慮したフィルタである。フィルタの中心座標 (x, y) に対するガウスフィルタ $h_g(x, y)$ は、ガウス分布の広がりや重みを变化させる標準偏差 σ を用いて以下のように表す。

$$h_g(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (1)$$

2.2 LoG フィルタ

LoG (Laplacian of Gaussian) フィルタは、ガウスフィルタによる平滑化と、ラプラシアンフィルタの二つの処理を一度に行いエッジを抽出するフィルタである。ガウスフィルタにより平滑化することで、ノイズを抑えながらエッジを抽出することができる。LoG フィルタ $h_l(x, y)$ は、標準偏差 σ を用いて以下のように表す。

$$h_l(x, y) = \frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^6} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (2)$$

2.3 ガウス 1 次微分フィルタ

ガウス 1 次微分フィルタはガウスフィルタの式を 0 度方向、45 度方向など特定方向のエッジを抽出するフィルタである。 θ 方向のガウス 1 次微分フィルタ $h_d(x, y, \theta)$ は、標準偏差 σ を用いて以下のように表す。

$$h_d(x, y, \theta) = \frac{x \cos \theta + y \sin \theta}{2\pi\sigma^4} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (3)$$

3. ボックスフィルタによる texton 特徴量の近似

前章で述べた texton 特徴量は、処理は容易であるが注目ピクセルとその周辺ピクセルすべてに対し演算するため、高速な特徴抽出は期待できない。また、フィルタのサイズに依存して計算コストが増加する問題がある。そこで、本論文では texton の各フィルタを、単数または複数のボックスフィルタ（平均化フィルタ）で近似し、高速に特徴量を抽出する手法を提案する。

3.1 ボックスフィルタによる近似

提案手法における近似手法の流れについて述べる。図 2 に本手法における近似フィルタ作成の流れを示す。まず、ガウスフィルタの近似をするため、入力情報として標準偏差 σ 、近似に用いるボックスフィルタの数 N を与える。入力情報をもとに近似パターンを網羅的に生成し、真のガウスフィルタとの誤差が最小

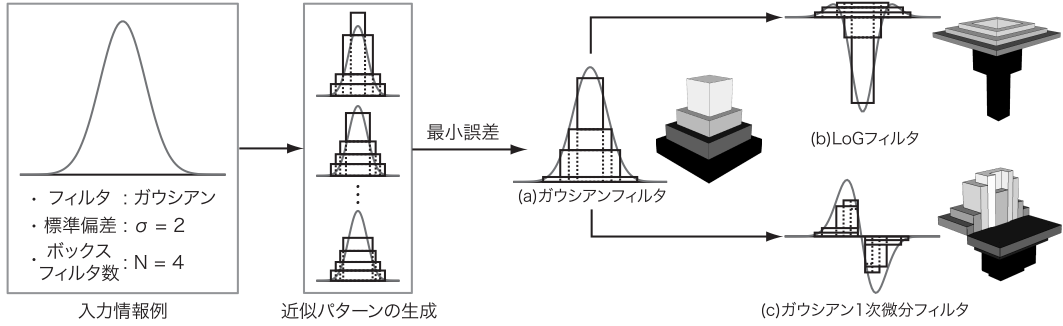


図 2 提案手法によるフィルタ生成の流れ

Fig. 2 Flow of building filters in proposed method.

となる近似パターンを、入力情報における近似ガウスフィルタとして採用する（図 2(a)）。次に、生成した近似ガウスフィルタに各微分式を適用し、LoG フィルタ（図 2(b)）とガウス 1 次微分フィルタ（図 2(c)）を自動生成する。このように、基礎となるガウスフィルタを構築すれば、LoG フィルタ、ガウス 1 次微分フィルタを自動的に生成できる。以下より各近似フィルタの生成の詳細を述べる。

3.2 ガウスフィルタの近似

はじめに、ガウスフィルタの構築をする。入力情報として、分布の広がりやを決定する標準偏差 σ と近似に用いるボックスフィルタ数 N を与える。本論文では σ は 1, 2, 4, 8 の 4 種類、 N は 1~6 の 6 種類の近似ガウスフィルタに対し、フィルタのサイズ範囲を、分布の 99.74% を信頼区間とする 3σ までとする。近似ガウスフィルタは以下のようにして決定する。まず、真値のガウス分布に当てはまる近似フィルタを網羅的に作成する。そして、ガウス分布の曲線からはみ出している体積、または満たされない体積の絶対値を誤差としてそれぞれのフィルタを評価する。最も誤差が少ないものを近似ガウスフィルタとして採用する。以下に具体的な近似ガウスフィルタの決定方法を示す。

(1) パラメータとして標準偏差 σ 、近似に用いるボックスフィルタ数 N を与える。

(2) 1 個目のフィルタとして、最下段に半径（サイズ）が $s_i = 3\sigma$ 、重みが w_1 のボックスフィルタを設置する。

(3) 1 個目のフィルタの上に、フィルタサイズが $s_2 = 3\sigma - (N - 1)$ 、重みが w_2 となる 2 個目のボックスフィルタを設置する。

(4) $N - 1$ 個目のフィルタまで、3. と同様の処理

によりフィルタを設置する。

(5) フィルタサイズより求まる面積と、重みである高さより、 $N - 1$ 個まで設置したフィルタの体積 s' を算出する。

(6) 最上段である N 個目のフィルタにおける、重みの合計 w_{top} を $w_{top} = 1.0 - s'$ より算出する。重みの合計が w_{top} 、サイズが $s_{N-1} - 1$ から 0 までの範囲でのボックスフィルタを設置し、真のガウス分布との誤差を算出する。

(7) 上記の処理を各ボックスフィルタのサイズ s 、重み w を変化させ、誤差が最小となるボックスフィルタの組合せを選択する。

ここで、各ボックスフィルタの重み w_n の初期値と増加量は、真値のガウス分布における原点の重みの $1/1000$ とする。また、サイズ（半径）は 1 ピクセルごとに変更する。以上の処理により選択されたボックスフィルタの組合せを、最も適したものとし、与えられた標準偏差 σ 、近似に用いるボックスフィルタ数 N における近似ガウスフィルタとして採用する。ここで、構築される近似ガウスフィルタは、 $\sigma = 1$ の場合、フィルタの最大半径が 3 ピクセルであるため、 $N = 5, 6$ の近似フィルタは構築されない。これに伴い、次節で述べる近似 LoG フィルタについても、 $\sigma = 1$ の場合、 $N = 5, 6$ の近似フィルタは構築されない。図 3 に $\sigma = 2$ の場合の各近似フィルタの例と誤差を示す。

3.3 LoG フィルタの近似

LoG フィルタはガウスフィルタにラプラシアンフィルタを適用したフィルタである。そのため、近似 LoG フィルタは、近似ガウスフィルタにラプラシアンフィルタを適用することで自動生成が可能である。適用するラプラシアンフィルタの式を以下に示す。

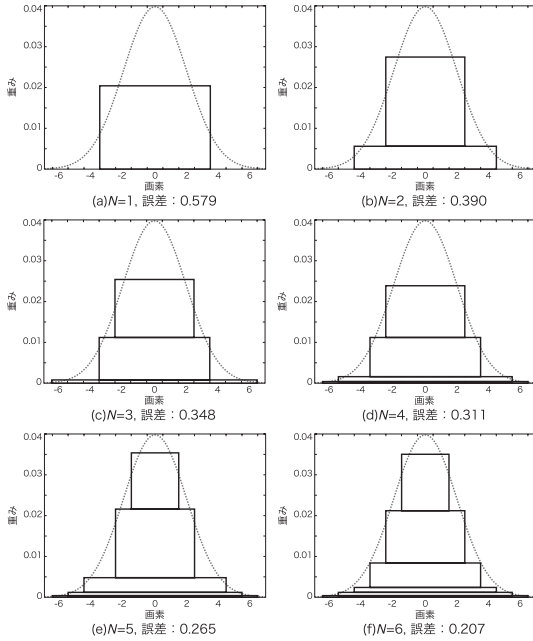


図 3 近似ガウスフィルタ例 ($\sigma = 2$)
Fig. 3 Examples of approximated Gaussian filter ($\sigma = 2$).

$$g_l(x, y) = \frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} \quad (4)$$

近似ガウスフィルタに $g_l(x, y)$ を適用するには、もとなる近似ガウスフィルタの N にかかわらず、一つのフィルタとして扱い、積をとることで適用する。しかし、各座標で与えられる値が異なるためボックスフィルタごとの重みが均一にはならない。そのため、各ボックスフィルタの重みの平均値を該当のボックスフィルタの重みとすることで、均一化をする。また、LoG フィルタに限り、 $N = 1$ は表現が不可能であるため、 $N = 2 \sim 6$ として扱う。

3.4 ガウス 1 次微分フィルタの近似

各ガウス 1 次微分フィルタは、方向 (角度) が異なるのみでありフィルタ内の値は同じである。そのため 0 度方向のガウス 1 次微分フィルタを近似し、45 度ずつ回転させることで各方向の近似ガウス 1 次微分フィルタとする。適用する 0 度方向のガウス 1 次微分フィルタには以下の式を用いる。

$$g_d(x, y, 0) = \frac{x}{\sigma^2} \quad (5)$$

近似ガウス 1 次微分フィルタも、近似 LoG フィルタと同様にもとなる近似ガウスフィルタのもつ重

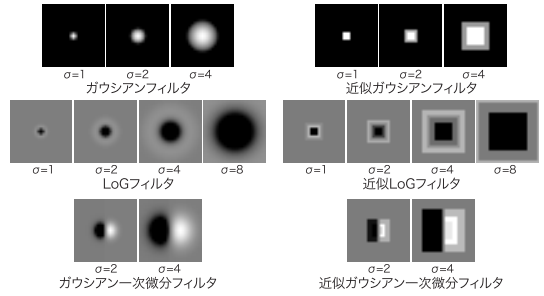
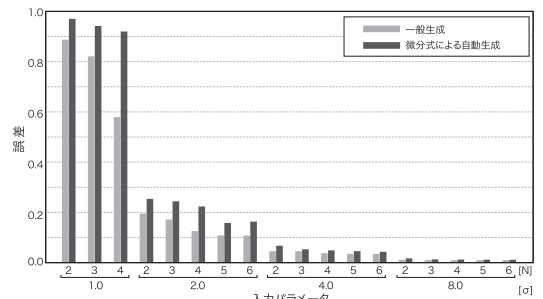
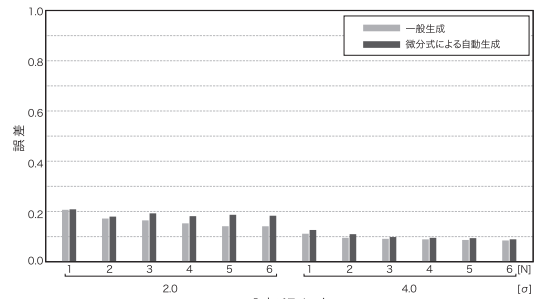


図 4 各フィルタの近似例

Fig. 4 Examples of each approximated filters ($\sigma = 2$).



(a) 近似LoGフィルタの誤差



(b) 近似ガウス一次微分フィルタの誤差

図 5 微分式による自動生成と一般生成による誤差

Fig. 5 Error of automatic generation by the differentiation and general generation.

みと $g_{dx}(x, y)$ の積を算出することで生成する。また、各ボックスフィルタの重みの平均値を該当のボックスフィルタの重みとすることで、均一化をする。微分フィルタの性質として、正領域と負領域の二つの領域に分割されている。そのため、実際に近似ガウス 1 次微分フィルタに用いられるボックスフィルタの数は $2N$ 個となる。最後に、各フィルタの近似例を図 4 に示す。

3.5 各フィルタの近似精度

前節で述べた手法により、自動生成した LoG フィルタ及びガウス 1 次微分フィルタの近似精度を確認する

ため、ガウスフィルタを構築した一般的な手法を用いて2種のフィルタを生成したものと比較する．図5(a)にLoGフィルタ，図5(b)にガウス1次微分フィルタにおける， σ と N の変化による真値との誤差の推移を示す．図5(a)のLoGフィルタでは， σ が小さい場合，1ピクセルに割り当てられる重みが高い．また，自動生成において正負の入換わりの位置が真のフィルタと異なり，誤差は大きくなってしまことから $\sigma = 1.0$ における誤差が大きくなっていることが分かる．これに対し図5(b)のガウス1次微分フィルタでは，正領域と負領域が異なるボックスフィルタとなっているため， $\sigma = 2.0$ においても誤差は安定していることが分かる．また，各フィルタは一般生成と同様に σ, N の増加に伴い誤差が減少しており，その差は極めてわずかである．これより，複雑な形状のフィルタを設計せずとも，ガウスフィルタと微分式を用いたLoGフィルタ，ガウス1次微分フィルタの生成法は有用なフィルタを自動生成できることを確認した．

4. Integral Image を用いた特徴抽出

提案手法では，各ボックスフィルタの輝度とを高速に求める処理が必要である．輝度とを求める場合，領域が移動するごとに1ピクセルずつ加算処理をしていては高速な特徴抽出は望めない．このような領域の和を繰り返し算出する問題に対し，ViolaとJonesはIntegral Image [14] を利用している．Integral Image は，方形領域内を輝度とを4点の演算のみで一括に算出することができる手法である．提案手法では，各ボックスフィルタに対する輝度との算出に利用する．これによりフィルタサイズに依存しない高速な特徴抽出が可能となる．以下よりIntegral Image と特徴抽出について述べる．

4.1 Integral Image

Integral Image により領域内の輝度とを算出するために，元画像 $I(x', y')$ に対し，以下の式を用いて積分画像 $II(x, y)$ を作成する．

$$II(x, y) = \sum_{x' \leq x, y' \leq y} I(x', y') \quad (6)$$

この積分画像を利用することで，図6(a)のような画像内の方角領域Dにおける画素和を求めることができる．方形領域Dの計算は，領域の端の4点，P1, P2, P3, P4を用いて以下のように求めることができる．

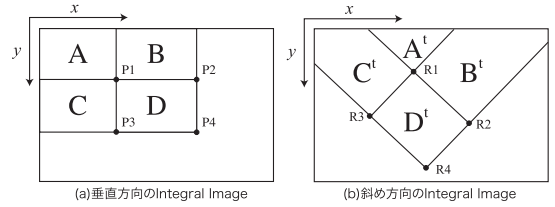


図6 Integral Image
Fig. 6 Integral Image.

$$D = P4 - P3 - P2 + P1 \quad (7)$$

45度，135度の斜めの矩形領域の輝度とを求めるには，斜め方向に対応したIntegral Image [21] を作成する．提案手法では垂直方向と斜め方向のIntegral Image を用意することになる．斜めの方角領域を計算する場合は，元画像 $I(x', y')$ に対して，以下の計算をすることにより，積分画像 $II'(x, y)$ を作成する．

$$II'(x, y) = \sum_{y' \leq y, y' \leq y - |x - x'|} I(x', y') \quad (8)$$

この積分画像 $II'(x', y')$ から領域 D^t を求める場合，垂直方向のIntegral Image と同様に，以下の式により算出する（図6(b)）．

$$D^t = R4 - R3 - R2 + R1 \quad (9)$$

4.2 Integral Image を用いた特徴抽出

Integral Image を用いることで，ボックスフィルタの応答値を4点の演算のみで算出することが可能となり，高速な特徴抽出が可能となる．近似フィルタから算出される特徴量 f は，各ボックスフィルタの応答値 D ，ボックスフィルタの高さである重み w を用いて以下の式により求められる．

$$f(p, t, N) = \sum_{i=1}^N D_t(p, i) w_t(i) \quad (10)$$

ここで $p = (x, y)$ は画像上の座標， t はフィルタの種類， N は近似に用いるボックスフィルタ数である．このように近似texton特徴量 $f(p, t, N)$ は，各ボックスフィルタの応答値 $D_t(p, i)$ に重み $w_t(i)$ を掛け合わせ，和をとるという非常に簡易な処理により算出される．

5. 近似 texton 特徴量による学習法

近似texton特徴量による人検出のための学習法について述べる．図7に本手法における学習の流れを

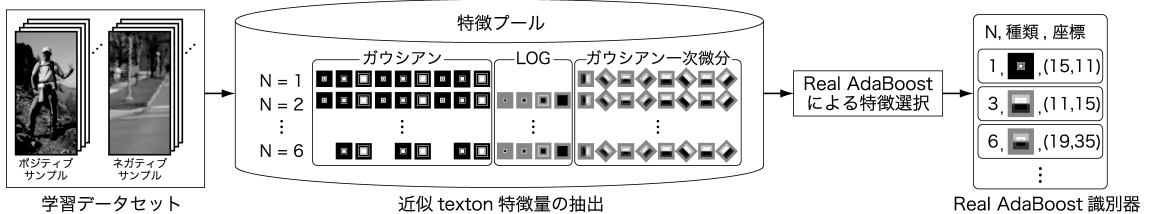


図 7 提案手法による学習の流れ
Fig. 7 Flow of learning by our proposed method.

示す. 提案手法における学習では, 入力された学習データセットから近似 texton 特徴量を算出し, 特徴プールに格納する. 特徴プールには, 各座標からボックスフィルタの数 $N = 1 \sim 6$, 21 種類の近似フィルタの全パターンを用いた特徴量を格納する. ここで, N が増加するほど, フィルタへの近似精度は向上することを 3.5 で確認したが, 近似精度と識別精度は必ずしも比例関係ではない. フィルタの近似精度が低くても, 人検出に有効な特徴量となる場合も十分に考えられるため, 与えられた学習サンプルをより識別できる最適な N も選択する必要がある. そこで, 学習には Real AdaBoost [22] を用いる. Real AdaBoost は, ランダム以上の性能の弱識別器を組み合わせることで高精度な識別が可能なアンサンブル学習手法である AdaBoost [23] を発展させたアルゴリズムである. AdaBoost は弱識別器が人 (1)/非人 (-1) のどちらかを出力するのに対し, Real AdaBoost は人/非人の確からしさを出力する. これによりサンプルの重みをより最適に決定し, 少ない弱識別器で高精度な識別器が構築できる. 本論文では, この Real AdaBoost を用い, 人検出に有効なフィルタの座標, ボックスフィルタ数, フィルタの種類を選択する.

6. 評価実験

提案手法である近似 texton 特徴量の有効性を評価するために, あらかじめ切り出された画像データセットを用いた識別精度の比較と, 1 フレームにおける処理時間の比較をする.

6.1 データセット

本節では, 識別精度の比較実験に用いるデータセットについて述べる. データセットには, 人検出のベンチマークとして広く知られている INRIA Person Dataset [11] を使用する. 学習, 評価には人の領域のみを切り出した以下の枚数の画像を用いる. 学習用サンプルとして, ポジティブサンプルが 2,416 枚, ネガ



図 8 INRIA Person Dataset の例
Fig. 8 Examples of INRIA Person Dataset.

ティブサンプルが 13,156 枚, 識別実験の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 枚の背景画像からラスタスキャンにより生成した 2,080,938 枚を使用する. 各サンプルは対象である人が低解像度である場合を想定し, 30×60 ピクセルにリサイズをして実験をした. 図 8 に, サンプル例を示す.

6.2 実験概要

比較手法には, 以下に示す手法を用いる.

- 近似 texton 特徴量
- texton 特徴量
- HOG 特徴量

各手法の比較に用いる Real AdaBoost の弱識別器数は, 100 個と 500 個を用いる. 以下より, 各実験の概要について述べる.

6.2.1 識別精度の実験概要

比較対象である HOG 特徴量のパラメータは, セルサイズを 5×5 ピクセル, ブロックサイズを 3×3 セル, こう配方向数を 9 方向とする. 評価には Detection Error Trade-off (DET) カーブを用いる. DET カーブは横軸に False positive per window (誤検出率), 縦軸に miss rate (未検出率) を表し, 原点に近いほど高精度であることを示す.

6.2.2 検出速度の実験概要

HOG 特徴量の抽出には Integral Histogram [16] を用いる。検出速度の比較実験で使用する HOG 特徴量は、3,240 次元全ての特徴量をあらかじめ算出する手法とする。Boosting で選択された各 HOG 特徴量を算出する場合、ブロック領域における正規化項を各特徴量ごとに算出する必要があるため、弱識別器の増加に伴い処理速度が低下する。これに対し、あらかじめ全次元の特徴量を算出する手法は、弱識別器数が増加した場合、効率的に識別をすることが可能となる。また、提案手法と texton 特徴量については、フィルタリングであるため、選択された特徴量のみを算出する。

人検出をするために、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。ラスタスキャンによる検出パラメータは、720×480 ピクセルの入力画像に対し、検出ウィンドウの初期サイズを 30×60 ピクセル、検出ウィンドウの最小スケールを 1.0 倍、最大スケールを 2.0 倍、スケールの変化幅を 0.2 倍ずつ（計 5 段階）、検出ウィンドウのずらし幅を 5 ピクセルとする。この場合における計 63,860 の検出ウィンドウに対して識別処理をした 100 フレームでの平均処理時間を求めた。実験環境は、OS: Windows Server Standard SP2, CPU: Intel Xeon X5470 @3.33 GHz, RAM: 32.0 GByte である。

6.3 識別精度の比較結果

図 9(a) に弱識別器数 100 個、図 9(b) に弱識別器数 500 個の場合における識別結果の DET カーブを示す。図 9(a) より、FPPW が 5.0% の場合、提案手法は HOG 特徴量と同様に miss rate が約 4.0% であり、同等の識別能力であるが、texton 特徴量が最も高精度である。また、図 9(b) では、弱識別器が増加したことにより識別精度が向上し、提案手法は、texton 特徴量により近い精度で識別が可能であることを確認した。

6.4 検出速度の比較結果

図 10 に各手法における弱識別器数の変化による平均処理時間を示す。また、このグラフは縦軸の検出速度が対数となっている。図 10 より、提案手法と texton 特徴量は、弱識別器数の増加に伴ない、処理コストが増加しているのが分かる。一方、HOG 特徴量は効率良く算出する手法を用いているため、検出速度は弱識別器数に大きく依存せず、250 個以上である場合は HOG 特徴量の方が速く検出できることが分かる。ここで、弱識別器数が 500 個の場合、提案手法は、HOG 特徴量と比較し処理コストは高くなるが、texton 特徴

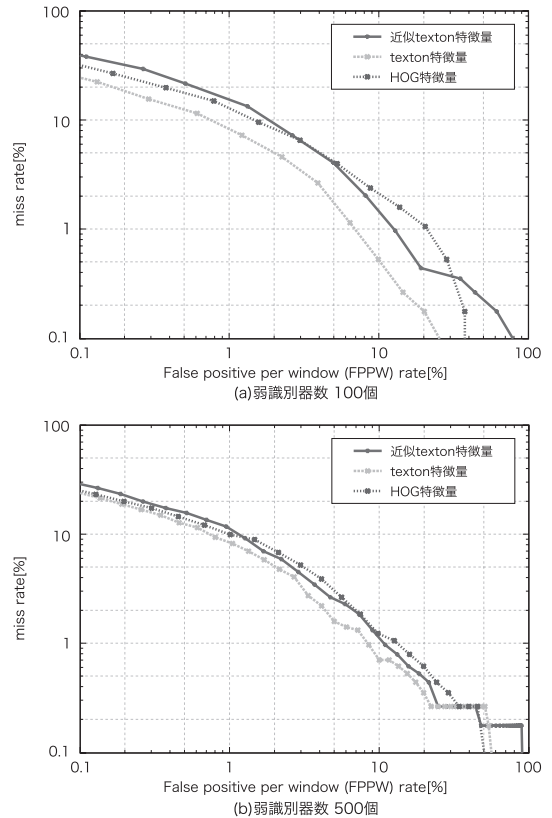


図 9 識別結果の DET カーブ

Fig. 9 DET curve of discrimination result.

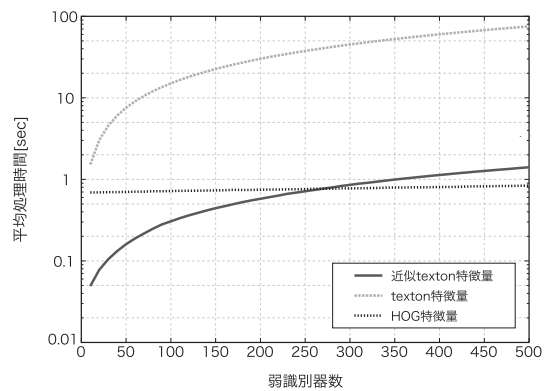


図 10 弱識別器数の変化による平均処理時間の比較

Fig. 10 Comparison of detection speed.

量により近い精度での識別が可能であることが分かる。また、弱識別器数が 100 個の場合、提案手法は HOG 特徴量と同精度であり、十分な識別が可能であることを示した。同条件である弱識別器数が 100 個の場合、提案手法は HOG 特徴量と比較し 2.3 倍、texton 特徴

案手法によるフィルタの選択のされ方, また文献 [25] のように色情報による局所相関なども併用することで, 精度良く人を検出することが可能である.

7. む す び

本論文では, ボックスフィルタで近似した textron 特徴量と Integral Image を用いた高速な人検出法を提案し, その有効性を示した. 提案手法は, 人検出において代表的な特徴量である HOG 特徴量と同精度であり, 2.3 倍高速な識別が可能であることを確認した. 今後は遮へい物などを考慮した人検出法について検討する.

文 献

- [1] V. Nair, P.O. Laprise, and J.J. Clark, "An FP-GAbased people detection system," *EURASIP J. Applied Signal Processing*, vol.2005, Issue 7, pp.1047–1061, 2005.
- [2] A. Ess, B. Leibe, K. Schindler, and L.V. Gool, "A mobile vision system for robust multi-person tracking," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1–8, 2008.
- [3] K. Khattab, J. Dubois, and J. Miteran, "Cascade boosting-based object detection from high-level description to hardware implementation," *EURASIP J. Embedded Systems*, 2009.
- [4] 門田亮二, 中村行宏, 宮本龍介, "実時間歩行者認識に向けた HOG 特徴抽出のハードウェア実装," *信学技報*, vol.109, no.78, pp.43–48, 2009.
- [5] 山中悠歩, 山崎俊彦, 相澤清晴, "FPGA による histogram of oriented gradients アルゴリズムの高速化," *情報処理学会創立 50 周年記念全国大会*, 2010.
- [6] B. Wu and R. Nevatia, "Detection of multiple, partially occluded humans in single image by Bayesian combination of edgelet part detectors," *IEEE Conference on International Conference on Computer Vision*, vol.1, pp.90–97, 2005.
- [7] B. Leibe, E. Seemann, and B. Schiele, "Pedestrian detection in crowded scenes," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.878–885, 2005.
- [8] O. Tuzel, F. Porikli, and P. Meer, "Human detection via classification on riemannian manifolds," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1–8, 2007.
- [9] T. Watanabe, S. Ito, and K. Yokoi, "Co-occurrence histograms of oriented gradients for pedestrian detection," *Advances in Image and Video Technology*, pp.37–47, 2009.
- [10] V. Chandrasekhar, G. Takacs, D. Chen, S. Tsai, R. Grzeszczuk, and B. Girod, "CHoG: Compressed histogram of gradients: A low bit-rate feature descriptor," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.2504–2511, 2009.
- [11] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.886–893, 2005.
- [12] C. Hou, H. Ai, and S. Lao, "Multiview pedestrian detection based on vector boosting," *Proc. Asian Conference on Computer Vision*, pp.210–219, 2007.
- [13] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," *Proc. ACM International Conference on Image and Video Retrieval*, pp.401–408, 2007.
- [14] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.511–518, 2001.
- [15] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," *Comput. Vis. Image Understand.*, vol.110, no.3, pp.346–359, 2008.
- [16] F. Porikli, "Integral histogram: A fast way to extract histograms in cartesian spaces," *IEEE Conference on Computer Vision and Pattern Recognition*, vol.1, pp.829–836, 2005.
- [17] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textron," *Int. J. Comput. Vis.*, vol.43, no.1, pp.29–44, 2001.
- [18] B. Georgescu, I. Shimshoni, and P. Meer, "Mean shift based clustering in high dimensions: A texture classification example," *IEEE Conference on International Conference on Computer Vision*, vol.1, pp.456–463, 2003.
- [19] J. Hays and A.A. Efros, "IM2GPS: Estimating geographic information from a single image," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1–8, 2008.
- [20] J. Winn, A. Criminisi, and T. Minka, "Object categorization by learned universal visual dictionary," *IEEE Conference on International Conference on Computer Vision*, vol.2, pp.1800–1807, 2005.
- [21] R. Lienhart and J. Maydt, "An extended set of Haar-like features for rapid object detection," *Proc. International Conference on Image Processing*, vol.1, pp.900–903, 2002.
- [22] R.E. Schapire and Y. Singer, "Improved boosting algorithms using confidence-rated prediction," *Mach. Learn.*, vol.37, no.3, pp.297–336, 1999.
- [23] Y. Freund and R.E. Schapire, "Experiments with a new boosting algorithm," *Proc. Thirteenth International Conference on Machine Learning*, pp.148–156, 1996.
- [24] D. Comaniciu and P. Meer, "Mean shift analysis and applications," *IEEE Conference on International Conference on Computer Vision*, pp.1197–1203, 1999.
- [25] S. Walk, N. Majer, K. Schindler, and B. Schiele,

“New features and insights for pedestrian detection,”
IEEE Conference on Computer Vision and Pattern
Recognition, pp.13–18, 2010.

(平成 22 年 11 月 30 日受付, 23 年 3 月 7 日再受付)



三井 相和 (学生員)

2009 中部大・工・情報卒. 同年より同修士課程に在籍. 画像処理, パターン認識に関する研究に従事.



藤吉 弘亘 (正員)

1997 中部大学大学院博士後期課程了. 1997~2000 米カーネギーメロン大学ロボット工学研究所 Postdoctoral Fellow. 2000 中部大学講師を経て 2004 より同大准教授. 2006 米カーネギーメロン大学ロボット工学研究所客員研究員. 工博. 計算機視覚, 動画画像処理, パターン認識・理解の研究に従事. 2005 年度ロボカップ研究賞. 2009 年度情報処理学会論文誌コンピュータビジョンとイメージメディア優秀論文賞, 2009 年度山下記念研究賞. 情報処理学会, 電気学会, IEEE 各会員.



秋田 時彦

1982 岡山大・工・電子卒, 同年よりアイシン精機(株)に在籍. 2010 香川大学大学院博士後期課程了, 工博. 車両制御システム及び車両周辺環境認識システムの研究開発に従事. 自動車技術会, システム制御情報学会各会員.