

1. はじめに

人検出に用いられている Histograms of Oriented Gradients(HOG) 特徴量 [1] は、物体の形状特徴を表すことで高精度な検出が可能であるが、ラスタスキャンによる検出手法では、多くの検出ウィンドウを処理するため計算コストが増加し、リアルタイム処理が困難である。そこで、本研究では texton 特徴量 [2] の filter bank をボックスフィルタで近似した人検出手法を提案し、その有効性を示す。

2. ボックスフィルタによる texton 特徴量の近似

2.1 texton 特徴量

texton 特徴量は複数のフィルタの応答値を特徴量として使用する。このフィルタの集合を filter bank と呼ぶ(図 1)。本研究では CIE L*a*b*画像に対し、ガウシアンフィルタから 9 個、LoG フィルタから 4 個、ガウシアン 1 次微分フィルタから 8 個の計 21 個を用いる。

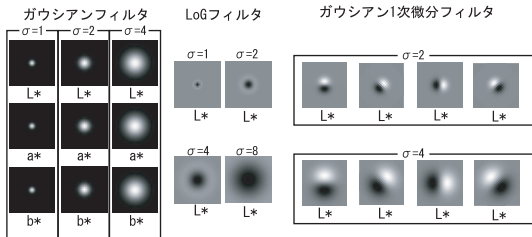


図 1: texton 特徴量の filter bank

2.2 ボックスフィルタによる近似フィルタ

ボックスフィルタとは、矩形領域の平均化フィルタである。提案手法では、単数又は複数のボックスフィルタを用いて texton 特徴量の各種のフィルタを近似し、特徴抽出に用いる。まず基本となるガウシアンフィルタを指定した数のボックスフィルタを用いて近似パターンを作成する。このとき、ガウシアンフィルタとの誤差が最小となるものを選出する。次に、近似ガウシアンフィルタを微分し、LoG フィルタとガウシアン 1 次微分フィルタを近似したフィルタを得る。図 2 に、ガウシアンフィルタ (a)、LoG フィルタ (b)、ガウシアン 1 次微分フィルタ (c) の近似例を示す。

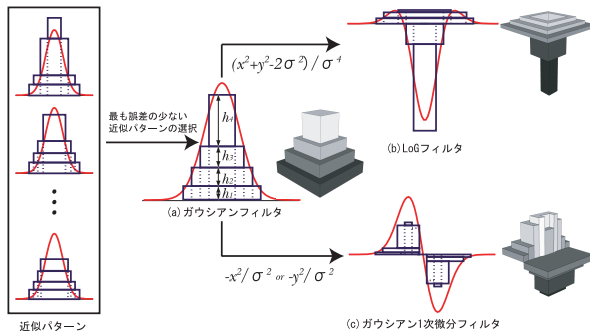


図 2: 各近似フィルタの構成

2.3 Integral Image を用いた高速化

Integral Image[3] は、予め画像 $I(x, y)$ から積分画像 $II(x', y')$ を作成しておくことにより、図 3(a), (b) の領域の和 D, D^t を式 (1), (2) を用いて求めることができる。

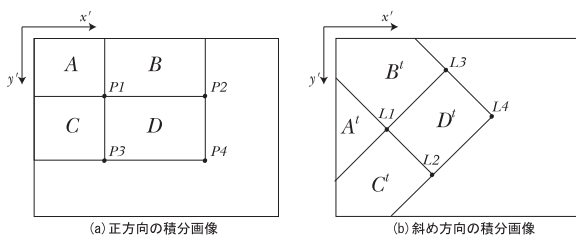


図 3: 積分画像の生成

$$D = P1 + P4 - P2 - P3 \quad (1)$$

$$D^t = L1 + L4 - L2 - L3 \quad (2)$$

Integral Image を用いると、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、高速な特徴抽出を実現できる。正方向の積分画像(図 3(a))を用いる場合、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和 D_i 、ボックスフィルタの高さである重み h_i を用いて、式 (3) より求められる。また、斜め方向の積分画像(図 3(b))を用いる場合も式 (3) と同様に算出する。

$$f = \sum_{i=1}^N D_i h_i \quad (3)$$

3. 評価実験

評価実験では、予め切り出された画像を用いて、提案手法、texton 特徴量、HOG 特徴量の精度比較を行う。データベースは、INRIA Person dataset を用いる。学習用サンプルとして、ポジティブサンプル 2,416 枚とネガティブサンプル 1,3161 枚、評価用サンプルとして、ポジティブサンプル 1,126 枚とネガティブサンプル 9,060 枚を用いる。学習には Real AdaBoost を使用し、評価には Detection Error Trade-off(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。図 4(a) に評価実験による結果の DET カーブを、図 4(b) に検出例を示す。図 4(a) より、誤検出率 5.0% の場合に、未検出率が texton 特徴量では約 1.7%、提案手法では約 2.1% となり、両手法の差は約 0.4% であることから、提案手法である近似 texton 特徴量は人検出に有効である。

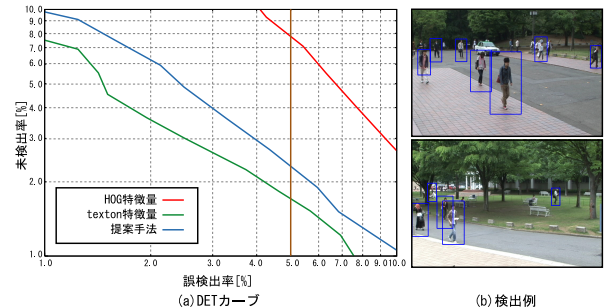


図 4: 検出結果の DET カーブと検出例

次に、HOG 特徴量の最高精度である識別率 97.64% と同精度における提案手法の検出速度を比較する。表 1 に、1 フレームにおける処理時間を示す。表 1 より、提案手法はインテグラルヒストグラム (IH) を用いた HOG 特徴量より約 7 倍高速であり、6fps でリアルタイムの人検出が可能である。

表 1: 同精度の場合における検出速度の比較

| 手法 | 弱識別器数 | 検出速度 [sec] |
|---------|-------|------------|
| 提案手法 | 41 | 0.16 |
| texton | 45 | 7.22 |
| HOG(IH) | 879 | 1.08 |

4. おわりに

本研究では、ボックスフィルタと Integral Image を用いた近似 texton 特徴量による人検出の手法を提案し、その有効性を確認した。今後は、更なる人検出法の高速化について検討する予定である。

参考文献

[1] N. Dalal and B. Triggs: "Histograms of Oriented Gradients for Human Detection", CVPR, pp. 886-893 (2005).
 [2] T. Leung and J. Malik: "Representing and recognizing the visual appearance of materials using three-dimensional texton", IJCV, 43, 1, pp. 29-44 (2001).
 [3] P. Viola and M. Jones: "Rapid Object Detection using a Boosted Cascade of Simple Features", CVPR, pp. 511-518 (2001).