

1. はじめに

物体検出に有効である HOG 特徴量は、局所領域毎に抽出されるため高次元となる。そのため、実装メモリが少ない FPGA などの小規模なハードウェアでは大きな問題となる。そこで、本研究では HOG 特徴量の情報量を削減するために、大小関係を用いて 2 値化した特徴量とワイルドカード (*) を用いた識別に不要なバイナリをマスクする手法を提案する。

2. HOG の 2 値化と問題点

Dalal らにより提案された HOG[1] は、局所領域における輝度の勾配方向をヒストグラム化した特徴量であり、物体の形状を捉えることが可能である。この特徴量は、近傍画素の勾配を局所領域毎にヒストグラム化するため、照明の影響を受けにくく、局所的な幾何学変化に頑健という特長がある。しかし、局所領域毎に特徴量を求めるため高次元となる。

そこで、HOG のメモリ量を削減するために、閾値処理により特徴量を 2 値化する。特徴量を 2 値化した Binarized HOG 特徴量 (B-HOG) は、セル領域における 8 方向分のバイナリを 1 つの特徴量 (バイナリパターン) として観測することにより、セル領域における勾配方向の関係性を捉えることが可能である。例えば、図 1 のような入力画像より得られる HOG 特徴量を 2 値化した場合、(00001011)₂ となる。

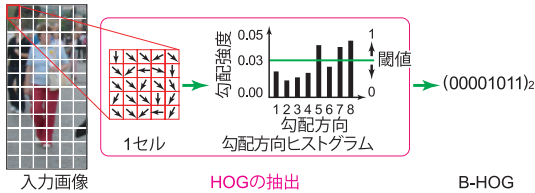


図 1：B-HOG の算出法

B-HOG の利点は、2 値化することにより浮動小数点型 (8 バイト) で表現しなければならない HOG の情報量を、符号なし整数型 (1 バイト) で表現できることである。そのため、B-HOG は、HOG よりメモリ量を 1/8 に削減することができる。しかし、B-HOG を用いる場合、学習サンプルを撮影した環境と実際に検出する環境が異なる時に、環境に合わせた最適な閾値を求めなければならないという問題がある。

3. 提案手法

B-HOG の問題点である閾値の決定法を解決するために、2 つの局所領域から抽出した HOG の大小関係を用いて 2 値化し、閾値を必要としない特徴量を提案する。しかし、2 値化した特徴量は、複数のバイナリを組合わせて構成されるため、識別に不要なバイナリが含まれる。そこで、Real AdaBoost により学習する際に、“0” と “1” の 2 つのバイナリを許容するワイルドカード (*) を導入することにより、識別に不要なバイナリをマスクする学習手法を提案する。

3.1 Relational HOG 特徴量 (R-HOG)

B-HOG の問題点である閾値の決定法を解決するために、閾値を必要としない特徴抽出をする必要がある。そこで、2 つの局所領域から抽出した HOG の値の大小関係を用いて 2 値化する Relational HOG 特徴量 (R-HOG) を提案する。異なる領域から抽出した特徴量の大小関係を用いて抽出するため、一つのバイナリで特徴量間の関係性を表現することができる。

R-HOG を算出するためには、式 (1) に示すように 2 つのセル領域から得られる 8 方向の勾配方向ヒストグラム V_{c_1}, V_{c_2} の大小関係を用いる。このとき抽出される特徴量は、 $B_{c_1 c_2}^{RH} = \{b_{c_1 c_2}^{RH}(1), b_{c_1 c_2}^{RH}(2), \dots, b_{c_1 c_2}^{RH}(8)\}$ となる。

$$b_{c_1 c_2}^{RH}(n) = \begin{cases} 1 & \text{if } v_{c_1}(n) \geq v_{c_2}(n) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

図 2 に示すように、2 つのセル領域における特徴量の大小関係を用いることにより、局所領域間の関係性を捉えるバイナリパターンを作成できる。このとき、R-HOG は全てのセル領域の組み合わせから算出する。しかし、図 3 (a) に示すように、抽出した特徴量が類似している場合、特徴量の大小関係が曖昧であるため、明確に 2 値で表現することが困難となる。

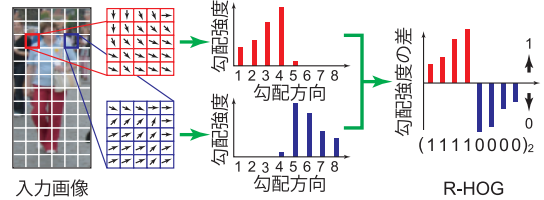


図 2：R-HOG の算出法

3.2 Shifted R-HOG 特徴量 (SR-HOG)

抽出された HOG の値が類似している場合、R-HOG では明確に 2 値で表現することは困難である。そこで、R-HOG を拡張した Shifted R-HOG 特徴量 (SR-HOG) を提案する。SR-HOG は、図 3 の (b) や (c) に示すように、一方のセル領域から抽出される 8 方向の勾配方向ヒストグラム V_{c_2} の勾配方向を s 方向シフトし ($s = 1, 2, \dots, 8$)、8 通りの勾配方向ヒストグラム $V_{c_2 s}$ を作成する。その後、式 (1) と同様に、式 (2) より大小関係と比較して、全シフトに対して特徴量を求めておくため 8 通りの 2 値化した特徴量 $B_{c_1 c_2 s}^{SRH}$ を算出する。

$$b_{c_1 c_2}^{SRH}(n, s) = \begin{cases} 1 & \text{if } v_{c_1}(n) \geq v_{c_2}(\text{mod}((n+s), 8)) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

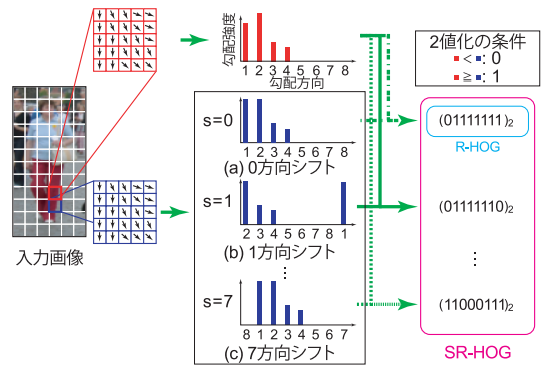


図 3：SR-HOG の算出法

SR-HOG は、勾配方向をシフトしたヒストグラムとの大小関係を算出するため、抽出した特徴量が類似している場合でも大小関係を明確に表現することが可能となる。

3.3 ワイルドカード (*) によるバイナリのマスクング

2 値化した特徴量を抽出した後、統計的学習手法である Real AdaBoost を用いて学習する。識別に有効なセル領域の位置とバイナリパターンを選択すると同時に、バイナリパターン化した特徴量に “*” を導入して識別に不要なバイナリをマスクすることにより、検出精度の向上が期待できる。

例えば、図 4 (a) に示すような 2 つの HOG を抽出した場合、図 4 (b) のようなバイナリパターンが得られる。このとき、図 4 (c) に示すように、“*” を用いて識別に不要なバイナリをマスクすることにより、図 4 (d) のように類似

した複数のバイナリパターンを同時に表現することが可能となる。“*”の数と何ビット目のバイナリに適用するかは、複数の弱識別器を用意して Real AdaBoost により選択する。

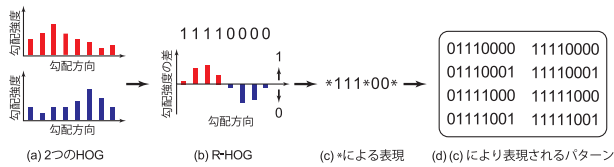


図 4：“*”を用いたバイナリパターンの表現例

本研究で使用する弱識別器は、R-HOG の場合、セル領域の組み合わせ r に対して、図 5 に示すようなバイナリと“*”を組み合わせたパターン o_r を観測する弱識別器候補を全パターン用意する。

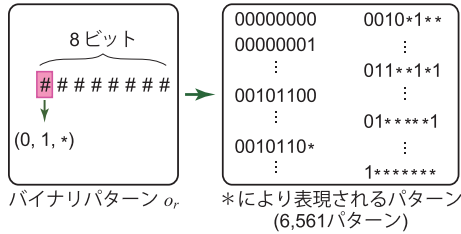


図 5：一組のセル領域における“*”を用いたパターン

1 つの検出ウィンドウ内に存在するセル領域の数が 8×16 セルの場合、セル領域の組み合わせ r は全部で 8,128 ($=_{128} C_2$) パターンとなり、バイナリと“*”を組み合わせたパターン O_r は全部で 6,561 ($= 3^8$) パターン存在するため、弱識別器候補数は約 5 千万 ($= 8128 \times 6561$) となる。

また、SR-HOG の場合、一方のセル領域から抽出した HOG を s 方向シフトすることによりバイナリパターンを抽出するため、弱識別器候補数は R-HOG の 8 倍である約 4 億 ($= 8128 \times 6561 \times 8$) となる。これらの弱識別器候補から、Real AdaBoost により識別に有効なセルの組合せやバイナリの組み合わせを選択して強識別器を構築する。

4. 評価実験

提案手法の有効性を示すために、以下の 2 つの評価実験を行う。1 つ目は大小関係を用いて 2 値化した特徴量の有効性を確認するために、HOG と検出精度を比較する。2 つ目は“*”を導入することの有効性を確認するために、導入前と導入後の検出精度を比較する。

4.1 評価実験の概要

評価実験では、従来法を含む下記に示す方法を比較する。

- HOG 特徴量 (HOG)
- Binarized HOG 特徴量 (B-HOG)
- Relational HOG 特徴量 (R-HOG)
- R-HOG 特徴量 + 勾配方向のシフト有 (SR-HOG)

実験に使用するデータセットは、人検出に使用する INRIA person dataset[1] と、車両検出に使用する文献 [2] のデータセットである。評価には、Detection Error Tradeoff (DET) カーブを用いる。また、B-HOG を算出する際に用いる閾値 th は予備実験にて決定し、人の場合は $th = 0.09$ 、車両の場合は $th = 0.10$ とする。

4.2 実験 1: R-HOG と SR-HOG の有効性

実験 1 では、大小関係を用いて HOG を 2 値化した R-HOG と SR-HOG の有効性を確認する。各データセットの実験結果を図 6 に示す。図 6 より B-HOG と R-HOG を比較すると、R-HOG の方が高精度であり、R-HOG と SR-HOG を比較すると、SR-HOG の方が高精度であることが分かる。最後に、SR-HOG は、HOG とほぼ同程度の検出精度であることがわかる。

4.3 実験 2：“*”によるマスクの有効性

実験 2 では、Real AdaBoost で識別器を学習する際に、“*”を導入して識別に不要なバイナリをマスクすることの有効性を確認する。本実験では、実験 1 の結果より提案

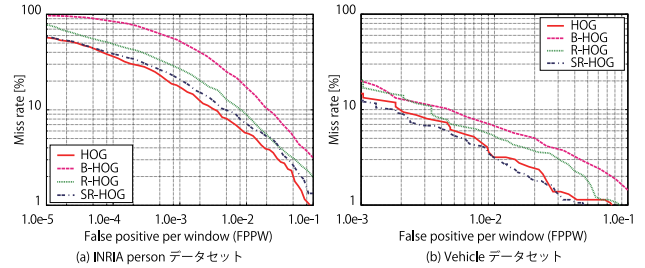


図 6：R-HOG と SR-HOG の有効性

手法の中で最も検出精度がよい SR-HOG に着目する。各データセットの実験結果を図 7 に示す。図 7 より、学習時に“*”を導入することで精度が向上し、HOG と同程度以上の検出精度であることがわかる。これは、各ポジティブサンプルの状態に依存して反転しやすく、識別に有用な情報を含んでいないバイナリが“*”として表現されるからであると考えられる。

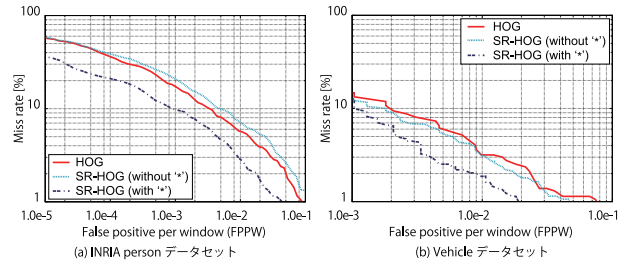


図 7：“*”によるバイナリ選択の有効性の実験結果

4.4 メモリ量と処理時間の比較

メモリ量と処理時間について比較する。弱識別器数を 500 個としたときの、1 検出ウィンドウ当たりのメモリ量を表 1 に示す。B-HOG と R-HOG を比較するとほぼ同じであり、R-HOG と SR-HOG を比較すると、勾配方向をシフトして 2 値化する SR-HOG は、R-HOG よりも約 2 倍のメモリ量が必要となる。また、HOG と SR-HOG を比較すると、SR-HOG は HOG を 2 値化しているため、メモリ量を約 75.0%削減することができた。

表 1：メモリ量 [KB] の比較

特徴量	HOG	B-HOG	R-HOG	SR-HOG
メモリ量	3.91	0.50	0.49	0.98

次に、1 検出ウィンドウ (64×128 ピクセル) における特徴量の抽出から識別に必要な処理時間を比較する。HOG の処理時間は 5.39×10^{-7} [ms] であり、SR-HOG は 2.70×10^{-7} [ms] であった。SR-HOG は、HOG の正規化処理を必要としないため、HOG よりも約 50.0%の処理時間を削減することができた。

5. おわりに

本研究では、HOG の情報量を削減するために、大小関係を用いて 2 値化した特徴量とワイルドカード (*) を用いた識別に不要なバイナリをマスクする手法を提案した。HOG を 2 値化することによりメモリ量の削減を実現した。また、識別に不要なバイナリをマスクすることにより、特徴量の情報量を削減しても高精度な検出を実現した。

参考文献

- [1] N.Dalal and B.Triggs: “Histograms of oriented gradients for human detection”, CVPR, pp.886–893, 2005.
- [2] 高木雅成, 山内悠嗣, 藤吉弘巨: “Geometric Context を用いた特徴量間の共起による物体検出の高精度化”, 電子情報通信学会論文誌 D, vol. J93-D, no. 8, pp.1429–1438, 2010.

研究業績

- [1] 松島千佳, 山内悠嗣, 山下隆義, 藤吉弘巨: “人検出のための Real AdaBoost に基づく HOG 特徴量の効率的な削減法”, 情報処理学会研究報告 CVIM167, pp.167–32, 2009. (他 研究会技術報告 2 件)

受賞

- [1] 情報処理学会 CVIM 研究会 卒論セッション 優秀賞 (他 1 件)