

階層型 HOG における高次特徴抽出量の PCA 解析

吉武 直樹^{a)} 庄野 逸^{b)}

概要：画像からの一般物体認識は、様々なモデルが提案されており、HOG などで得られた特徴を、SVM などの識別器を用いて分類するという形態が知られている。Arakaki らは、HOG モデルを低次の局所特徴と捉え、低次特徴を統合するような高次特徴を導入することで、より複雑な画像特徴を捉えるモデルを提案しており、特徴次元の大きさを抑えたままで HOG モデルと同程度の識別率を達成している。Arakaki らのモデルでは、高次層での特徴抽出において、ランダムサンプリングを用いていたため、識別能力のパフォーマンスにばらつきを生じていた。そこで、本研究では PCA を用いて高次層の特徴抽出を検討し、物体認識の精度の向上を試みた。

キーワード：物体認識, HOG, PCA

1. はじめに

画像からの一般物体認識は、画像処理や機械学習の分野から様々なモデルが提案されており、HOG (Histogram of Oriented Gradients) などで得られた特徴を、SVM (Support Vector Machine) などの識別器を用いて分類するという HOG モデルが知られている [1]。Arakaki らは、検出の頑強性を向上させるため、脳の視覚処理モデルであるコンボリューションネット [2] の階層性を取り入れることで、HOG モデルの改良を図った [3]。

コンボリューションネットは、局所的な特徴の統合を、徐々に行うことで視覚パターンの識別を行うモデルで、Arakaki らは、HOG モデルから得られる特徴を低次の局所特徴と捉え、高次層を追加することで、より複雑な画像特徴を捉えるモデルを提案しており、特徴次元の大きさを抑えたままで HOG モデルと同程度の識別率を達成している [3]。図 1 はコンボリューションネットと、HOG、階層型 HOG モデルの関係を表した模式図である。

Arakaki らの階層型 HOG モデルでは、高次層 (HS2 層) での特徴抽出において、低次特徴 (HC1 層) の表現を入力とした template matching を行うものとし、学習 template は、入力からのランダムサンプリングを用いていた。この結果として、識別能力のパフォーマンスにばらつきを生じていた [3]。本研究では、template 取得のアルゴリズムとして PCA を用いることにより、パフォーマンスのばらつきを抑えながら物体認識の精度の向上を図る。性能検証のために、歩行者検出用の INRIA Person Data Set [4] を用い、識別性能や Arakaki らのモデルとの比較について検証

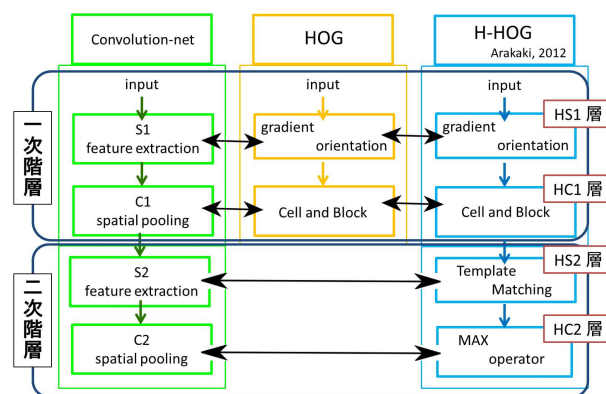


図 1 HOG モデルとコンボリューションネットの比較

と考察を行った。

2. モデル

ここでは、本研究での基本となる階層型 HOG モデルについての説明を行う。

2.1 HOG モデル

HOG モデルの処理は、まず、各画素の勾配強度と勾配方向を算出し [1]、局所領域内での統合を行う。cell と呼ばれる領域に含まれる画素の勾配強度と勾配方向を用いて、勾配方向ヒストグラムを作成する。次に、複数の cell を 1block として block 内で正規化を行う。ここでは、1block を 3cell × 3cell の正方形型にし、L2-norm を用いて正規化を行っている [3]。各 block のヒストグラム要素をすべてを連結させた多次元ベクトルを HOG 特徴量とする。

a) yoshitake@uec.ac.jp

b) shouno@uec.ac.jp

2.2 階層型 HOG モデル

次に階層型 HOG モデルの HS2, HC2 層についての説明を行う。前述の HS1 層と HC1 層は、それぞれ画像勾配を抽出する操作と正規化を行う非線形処理であると捉えることができ、Arakaki らは、これらの HOG モデルにおける操作を、コンボリューションネットの特徴抽出操作と空間プーリング操作とに対応付けた。以降では、高次層に置ける template matching による特徴抽出層を HS2 層、プーリング操作などを行う層を HC2 層と呼ぶ。また、HOG モデルにおける特徴抽出層、空間プーリング層を HS1, HC1 層と呼ぶこととする (図 1 参照) [3]。

HS2 層は、template matching を用いて勾配ヒストグラムブロックから再度特徴抽出を行う。Arakaki のモデルでは、HS2 層における template は、HC1 層において得られる画像の部分表現をランダムに抽出したものをを用いることとしていた [3]。以降では、Arakaki らの学習モデルを Random template モデルと呼ぶこととする。

HC2 層では画像から得られる複数の HS2 出力値に対して MAX operator を用いることで最大値のみ取り出す。すなわち、画像にどの程度 template が含まれているかを表す。この値が識別器に与える特徴ベクトルの 1 要素となる [3]。

2.3 統計的学習アルゴリズム

ここでは、本研究でとりいれた、PCA を用いた学習モデルについての説明を行う。

2.3.1 Principal Component Analysis: PCA

PCA とは、高次元の特徴空間に分布する多数の学習用画像から、特徴空間に分布する多数の学習用画像の分布から、その分布をパワーの意味でよく表す直交座標系を抽出する手法である [5]。PCA は、特徴空間で分散が大きい方向順に直交成分を順次取り出す手法で、分散の大きい順に第 1 主成分、第 2 主成分、... と名づけられた特徴方向を抽出する。その上で多次元データを主成分軸上に射影することでデータ表現を変換する [5]。

2.3.2 PCA を使った HS2 層

PCA は n 次元の多変量観測データ x を、PCA 軸による表現 s に変換する。ここでは、 x は、HC1 層での表現、 s は、HS2 層での抽出結果をそれぞれ表す。

$$s = \begin{pmatrix} w_1^T x \\ w_2^T x \\ \vdots \\ w_n^T x \end{pmatrix} = W^T x. \quad (1)$$

ただし、変換行列 W の行要素 w_i は第 i 主成分に対応する [6]。ここで、HS2 層での template matching において、 w_i は template と考えることができる。今回新たに提案する階層型 HOG モデルの HS2 層では、PCA により

求まる template w_i を用いて template matching を行う。以下では、この PCA を使った HS2 層を導入する階層型 HOG モデルを、PCA template モデルと呼ぶ。

3. 計算機実験

本研究では性能の検証のため、INRIA Person Data Set [4] を用いて実験を行った。INRIA Person Data Set は位置・スケール変化があるように画像サイズを 300×600 に切り出し使用した [3]。template 取得用に positive 100 枚、negative 100 枚使い、識別用に positive 300 枚、negative 300 枚用いた。HOG 処理の際の cell サイズは 5 種類、HS2 層で用いる template サイズは、3 種類用いて実験を行った。

実験結果を表 1 に示す。実験は、10 回の識別率の平均を表し、括弧内の数字は標準偏差である。この結果、ランダムサンプリングよりも PCA の方が良い識別率を得ることができ、ランダム性も排除された。

表 1 PCA, ランダムサンプリングによる template 学習に対する識別率の比較。() 内の数字は標準偏差を示す。

	識別率 [%]	特徴次元数
Random template	80.67(1.83)	1380
PCA template	86.33	1380

4. まとめ

本研究では Arakaki らの階層型 HOG モデルに、PCA を使った template 決定方法を採用する新たな階層型 HOG モデルを提案した。その結果、ランダム性が排除され、物体認識の精度の向上を達成することができた。今後の課題として、適した template の解析と、ICA など他手法による template 決定方法による識別率の比較などがあげられる。

参考文献

- [1] Navneet Dalal and Bill Triggs: "Histograms of Oriented Gradients for Human Detection" Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 886-893, 2005.
- [2] J. Mutch, and D.G. Lowe: "Multiclass Object Recognition with Sparse, Localized Feature", Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 11-18, 2006.
- [3] 新垣泰仁: "コンボリューションネットを用いた階層化 HOG モデルによる人物検出の検討", 2011.
- [4] N. Dalal, INRIA Person Dataset. <http://pascal.inrialpes/data/human/>
- [5] デジタル画像処理編集委員会 奥富正敏: "デジタル画像処理", 第二版 財団法人 画像情報教育新興協会, 2008.
- [6] 村田昇: "独立成分分析", 第二版 東京都電機大学出版局, 2005.