

HOG 特徴量を用いた船舶画像認識

正会員 松本 洋平（東京海洋大学）

要旨

HOG(Histogram of Oriented Gradient)特徴量と線形 SVM(Support Vector Machine)を船舶画像認識に適用し評価を行った。HOG 特徴量は様々な対象の画像認識に実績があるが、船舶を対象とした場合の報告はまだなされていない。評価の結果、高解像度の船舶画像に対しては98%/6%(認識率/誤認識率)の認識精度を達成した。

キーワード：情報処理、画像認識、HOG、SVM

1. はじめに

本研究はコンピュータビジョンを用いて船舶を自動的に発見すること、そのための船舶画像認識の実現を目的としている。

船舶の衝突事故の主要因は見張り不十分であるとされている。それらの内訳は必ずしも目視による見張りの不足が原因ではないが、船舶画像認識技術を補助として用いることで、衝突のリスクを削減することができると思われる。

また、近年、離島監視の重要性が高まっている。広大な領海と多数の離島を有する日本国においては、監視網を充実することは容易ではなく、船舶画像認識による補助は有益であると考えられる。

画像認識の研究は様々な対象にたいして数多く行われているが、船舶を対象とした例は少ない。

2 章では物体検出の基礎的な解説と船舶画像認識に関する過去の研究を紹介する。3 章では本研究で船舶画像認識に適用する HOG・SVM について解説する。4 章では評価手法とそれを実現するための学習データの作成について説明する。5 章で評価結果を示し、6 章でまとめる。

2. 関連研究

2.1 物体検出

コンピュータビジョンを用いた物体検出は様々な分野で研究がされており、とりわけ、自動車からの歩行者の検出の研究は多くなされている⁽¹⁾。それらの研究ではおおむね検出窓で走査する方法がとられている。図1に典型的な物体検出アルゴリズムの擬似コードを示す。検出窓(幅 w 高さ h)で画像上の水平・垂直方向(x, y)、スケール方向(s)を走査する。その3重ループの本体が画像認識アルゴリズムである。

画像認識アルゴリズムは(s, x, y)においた検出

```
detectObject(I, O, (w, h)) {
    O = {null}; // initialize list of detected object.
    for each scale s {
        Is = Scale(I, s); // scale the image with factor s
        for each horizontal position x in Is {
            for each vertical position y in Is {
                fv = calculateFeature(Is, (x, y), (w, h));
                if(determine(fv) > th)
                    push(O, (s, x, y));
            }
        }
    }
}
```

図1 物体検出の擬似コード。

窓の切り取る部分画像が対象の物体であるか否かを判定する。

画像認識アルゴリズムは主に2つの部分、すなわち、特徴量の抽出(図1の calculateFeature)と識別(図1の determine)に分けられる。特徴量の抽出としては主成分分析等の統計的なものや、Haar-like 特徴、SIFT 特徴量、HOG 特徴量⁽²⁾などの発見的なものがある。識別手法としては単純ベイズ法、ニューラルネットワーク、SVM⁽⁶⁾などがある。

近年、HOG と SVM を組み合わせた手法が様々な対象の画像認識に実績を挙げている。本研究では HOG・SVM を船舶画像認識に適用し評価する。

2.2 船舶画像認識

船舶を対象とした画像認識の研究を紹介する。

文献(3)は背景差分法を用いて船舶を検出することを試みている。また、文献(4)はオプティカルフローを用いて船舶を検出することを提案している。こ

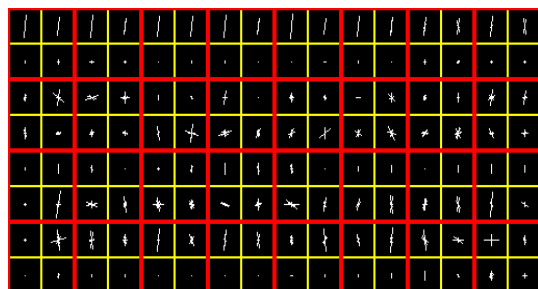
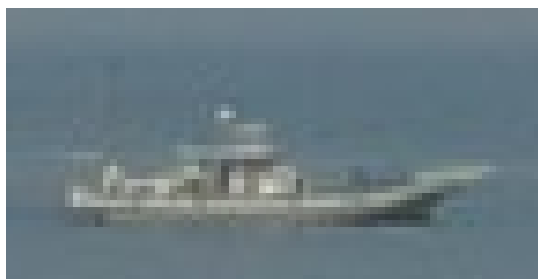


図2 船舶画像(上)と HOG 特徴量(下)

これらの研究は船舶の認識に時系列データを用いているが、船舶が高速に運動しているか、あるいは、近い位置を航行しているのでなければ、これらの手法は有効に働かない。すなわち、基本的に静止画像から船舶を認識できることが好ましい。文献(5)は色相の分散を用いる方法を提案しており、しかし、その識別精度の評価はされていない。

いずれの船舶画像認識の研究においても認識精度（認識率・誤認識率）の定量的な評価が行われていない。画像認識の評価にはGround Truthとなるデータを多量に必要とし、また、それは人間によって作られなくてはならないため、非常に時間がかかる。しかし、それらのデータは機械学習を用いた画像認識手法においては、学習データとしてそもそも必要である。したがって、本研究では多くの学習データセットを用意した。その結果、認識精度の定量的な評価が可能となっている。

3. 評価する手法

特徴量としてHOG、機械学習アルゴリズムとして線形SVMを用いる。以下、それぞれを簡単に説明する

3.1 HOG 特徴量

HOG 特徴量は勾配画像を基礎とする特徴量であり、まず検出窓内の画像をグレースケール化し、微分画像を計算する。検出窓の内部は複数のセルという領域に分けられ、セル毎に勾配ベクトルの方向のヒストグラムを作る。このとき、ヒストグラムの各階級の投票では勾配ベクトルのノルムで重みをつける。複

表1 各検出窓サイズにおける船舶画像および非船舶画像数

	船舶画像	非船舶画像
128x64	479	938
64x32	1134	2240
32x16	1952	3862
16x8	2280	4496

数の隣接するセルがブロックという単位にまとめられ、各セルのヒストグラムを束ねたベクトルはブロックディスクリプタと呼ばれる。ブロックディスクリプタはノルムが1になるように正規化される。最終的にHOG 特徴量のベクトルとは検出窓内のブロックディスクリプタを束ねたものとなる。

本研究ではHOG 特徴量の計算に勾配ヒストグラムの階級数9、セルのサイズ4x4、ブロック（赤枠）のサイズ2x2を用いている。これは文献(2)などで示された典型的な設定である。対象によって最適なパラメータは異なると考えられるが、今回はその探索は行わない。

図2は64x32の船舶画像に対するHOG 特徴量の計算例を示している。それぞれのセルの勾配ヒストグラムは、階級値を方向で、度数を長さで示している。この場合、検出窓内は128のセルに分割され、それらのセルが、さらに32のブロックに分割される。このとき、最終的なHOG 特徴量は128x9=1152次元のベクトルとなる。

なお、本研究においてHOG 特徴量の計算プログラムはC++で作成した。

3.2 SVM

機械学習アルゴリズムの目的は、この場合、HOG 特徴量の空間で、船舶の分布と非船舶の分布を分離する境界面を探すことである。SVMはクラス間のマージンが最大となるような分離境界面を決定する⁽⁶⁾。

本研究では線形SVMを想定し、カーネルトリックは使わない。また、SVMの計算にはOpenCVのコードを用いた。ソフトマージンのパラメータは自動推定に任せ、その結果、0.1~0.5となっている。

4. 評価方法

船舶画像認識の評価は次の手順で行った。

- (1) 検出窓と同サイズの船舶画像と非船舶画像を多数用意
- (2) 船舶画像・非船舶画像を学習用と評価用に分割

(3) 学習用データを用いて船舶画像認識アルゴリズムに船舶と非船舶の HOG 特徴量を学習

(4) 評価用データを用いて船舶画像認識アルゴリズムの認識精度を評価

なお、ステップ(3)においては学習用データは鏡対象の画像も用いた。また、(3)と(4)のステップは少ない学習データを有効に活用するため、実際には k-fold CV (Cross Validation)¹を行った。

4.1 学習データの作成

船舶画像と非船舶画像の準備方法を説明する。

2012年7月24~26日、2013年1月25~27日、3月6~8日に東京海洋大学の練習船である汐路丸が勝どき・館山間を往復する間に撮影した2158枚の画像から学習データを作成した。カメラは汐路丸のフライングブリッジの正面方向に取り付けた。カメラの解像度は4912x3264、センサのサイズはAPS-C(23.6mmx15.6mm)、レンズの焦点距離は18mmである

まず、撮影した画像の中の船舶の領域を矩形で指定し、その矩形の左上角と右下角の座標を記録した。これをラベリングデータと呼ぶ。このラベリングデータから128x64、64x32、32x16、16x8の検出窓サイズを想定して船舶画像、非船舶画像を抽出した。

ラベリングデータの船舶を囲む矩形のアスペクト比(横/縦)の平均は2.18であり、検出窓のサイズとしてはもっとも近い整数値として2.0を選んだ。

船舶画像の抽出に際して、窓サイズに満たないサイズの船舶は捨てて、窓サイズより大きい船舶は窓サイズまで縮小した。また、非船舶画像は船舶の存在しない領域からランダムに抽出した。

表1に各窓サイズの船舶画像、非船舶画像数をまとめる。また、図5に128x64の学習データの例を示す。

なお、k-fold CVを行う上で、128x64、64x32のデータセットではk=5とした。また、32x16、16x8のデータセットに関しては、逆にデータ数が多すぎてSVMの学習が収束しない問題が生じたので、学習データの1/kを学習に、残り(k-1)/kを評価に用いることとし、それぞれ、k=4、k=5とした。

5. 評価結果

図4は各窓サイズに対して認識率を横軸に誤認識

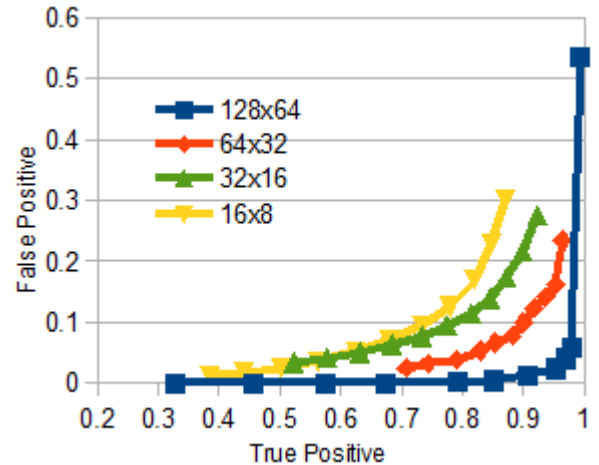


図4 認識率(True Positive)と誤認識率(False Positive)

率を縦軸にとって、SVMの閾値を-1から1の間で動かしながらプロットしたものである。

対象の情報を多く使うことができるという点で、基本的に窓サイズが大きいほど高い認識精度を得られる。具体的に、例えば128x64の閾値の高い領域では認識率98%、誤認識率6%を達している。

ただし、窓サイズが大きいほど学習データが少ないので、評価結果の信頼性および、SVMの訓練度が低い可能性がある。データ数は現在も作成中であり、発表の段階ではより信頼性の高い結果を出せる予定である。

6. まとめ

HOGとSVMを用いた船舶画像認識を多数の学習データを用意して評価した。学習データは現在も作成中であり、今後他の手法も評価していきたい。

学習データはまだ十分であるとはいえないが、高解像度の検出窓(128x64)を用いる場合の認識精度は十分高いといえる。撮影に用いたカメラでは検出窓の幅である128ピクセルは、約300mの距離の10mに相当する。見張りに必要な距離が長ければレンズ焦点距離を延ばすか、高解像度のカメラを用いる必要がある。

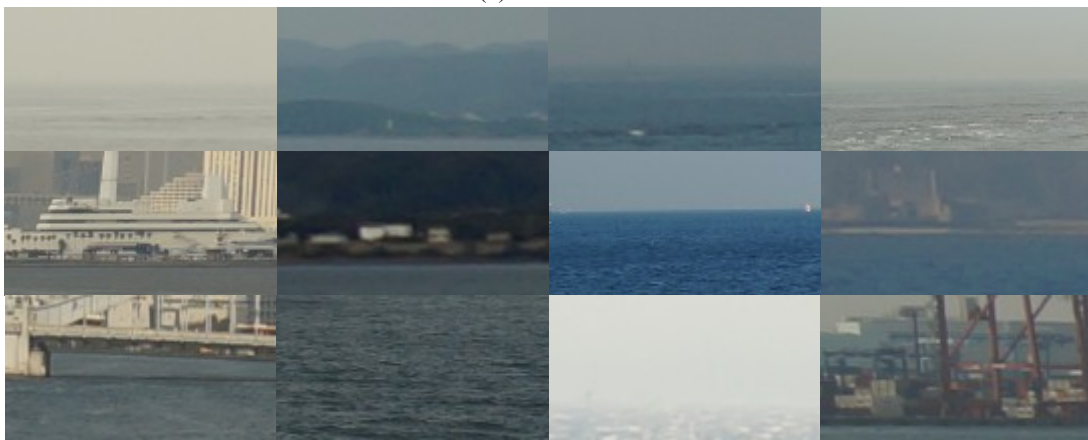
一方で低解像度での認識精度を上げる余地も残されている。まず、今回はHOGのあらゆるパラメータを探索したわけではない。また、HOG・SVMのほかにもさまざまな手法が存在しており、それらを併用することも考えられる。

また、見張りのシステムとしては、認識した船舶を追尾する、距離や速度を測定する、進路を予測する、

¹ 学習データをk分割し、 $S_i (i=1 \sim k)$ を除く学習データを用いて訓練を行い、 S_i の識別精度を評価することを、k度行った平均を評価する。



(a) 船舶画像



(b) 非船舶画像

図5 学習データの例(128x64)

危険船舶を予測する、といったことも実現されねばならない。これに関しても今後研究を進めていきたい。

7. 参考文献

- (1) M. Enzweiler, D. M. Gavrilu, “Monocular Pedestrian Detection Survey and Experiments,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.31, no.12, pp.2179-2195, 2009.
- (2) N. Dalal, B.Triggs, “Histograms of Oriented Gradients for Human Detection,” Proc. IEEE Intl. Conf. CVPR, pp.886-893, 2005.
- (3) 周立輝, 新保雅俊, 大島正毅, ” 航海環境時系列画像からのリアルタイム船舶抽出手法,” 日本航海学会論文集, vol. 117, pp.175-182, 2007.
- (4) 桐谷伸夫, 福戸淳司, ” 海上観測画像を用いたフレーム間差分による動体認識,” 日本航海学会論文集, vol.113, pp.107-113, 2005.

- (5) 新保雅俊, 周立輝, 平澤雅人, 大島正毅, ” HSV分析による航海環境時系列画像からの船舶抽出,” 日本航海学会論文集, vol. 116, pp.69-76, 2007.
- (6) C. Burges, “A Tutorial on Support Vector Machines for Patter Recognition,” Data Mining and Knowledge Discovery 2, pp.121-167, 1998.

謝辞

千田直紀君はたくさんの学習データのラベリングを、鈴木健司君はラベリングツールの開発をしてくれました。また、数多くの画像データが東京海洋大学の練習船・汐路丸の協力で得られました。ここに感謝の意を表します。