Ship detection in optical remote sensing image based on visual saliency and AdaBoost classifier*

WANG Hui-li (王慧利)[1,2], **ZHU Ming** (朱明)[1]**, **LIN Chun-bo** (蔺春波)[1], **and CHEN Dian-bing** (陈典兵)[1,2]

1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

2. University of Chinese Academy of Sciences, Beijing 100039, China

(Received 17 January 2017; Revised 10 February 2017)

In this paper, firstly, target candidate regions are extracted by combining maximum symmetric surround saliency detection algorithm with a cellular automata dynamic evolution model. Secondly, an eigenvector independent of the ship target size is constructed by combining the shape feature with ship histogram of oriented gradient (S-HOG) feature, and the target can be recognized by AdaBoost classifier. As demonstrated in our experiments, the proposed method with the detection accuracy of over 96% outperforms the state-of-the-art method.
efficiency switch and modulation.

Document code: A **Article ID:** 1673-1905(2017)02-0151-5

DOI 10.1007/s11801-017-7014-9

Ship detection is essential for oceanic traffic monitor, marine rescue and military application[1]. So far, synthetic aperture radar (SAR) images[2] and optical remote sensing images[3-8] are two main sources for ship detection.Gradually, optical remote sensing image becomes the chief source because of its higher resolution, wider coverage, lower background noise and larger quantity of detail information. In order to extract ship objectives quickly and accurately, the current popular objective detection includes two steps: extracting the candidate regions and distinguishing the real objectives from the false ones.

At the stage of candidate region extraction, Ref.[4] proposed a method of threshold segmentation based on gray-level and edge information, Ref.[5] acquired candidate regions by segmenting the saliency map, while the saliency map is based on the sparse feature of multi-layer sparse coding for image, and Ref.[6] acquired the candidate regions by eliminating on-objective surroundings according to the analysis of sea surface. At the stage of objectives discrimination and confirmation, the ship objectives' complex features and support vector machine (SVM) are commonly used to confirm the final detection targets. Ref.[4] built a 679-dimension vector including the features of the shape feature, the texture feature and other characteristics, Refs.[7] and [8] extracted the local binary pattern (LBP) feature, and Ref.[5] extracted the deformable part mode (DPM) feature. The methods mentioned above can produce good detection results, but the cost of calculation is much larger.

Actually, neither the complexity of approach nor the number of features' dimension can determine the detection effect. Ref.[9] proposes a new ship histogram of oriented gradient (S-HOG) descriptor, which combines the ship features with the traditional histogram of oriented gradient (HOG) features. The advantage of this algorithm is that it can detect the target in the unsupervised condition. However, the homogeneous filter in the stage of extracting candidate region and the selection and setting of many parameters in the target discrimination phase both lead to the degradation of real-time performance and the increase of complexity.

In this paper, we propose a novel detection method based on visual search mechanism and cellular automata evolution model. Compared with Ref.[9], the algorithm proposed in this paper can greatly improve the real-time performance and detection performance. It can not only extract the target region more quickly and accurately, but also describe the target more precisely by combining the S-HOG features with the shape features, and the application of the AdaBoost classifier in the final target confirmation can avoid the influence of parameter setting on the algorithm.

The main flow chart of the ship detection algorithm proposed in this paper can be described in Fig.1. Firstly, saliency map is acquired by the maximum symmetric surround method, and updated according to the local similarity. Secondly, saliency map is segmented by Otsu algorithm to obtain binary image, and then salient regions are extracted from the segmented binary image and

* This work has been supported by the National Natural Science Foundation of China (No.61401425).

** E-mail: zhu_mingca@163.com

filtered roughly by the ship objectives' geometric features. Finally, feature vectors are extracted for every salient region from their shape features and S-HOG features, and the detection objectives are confirmed by thorough discrimination using the AdaBoost classifier.

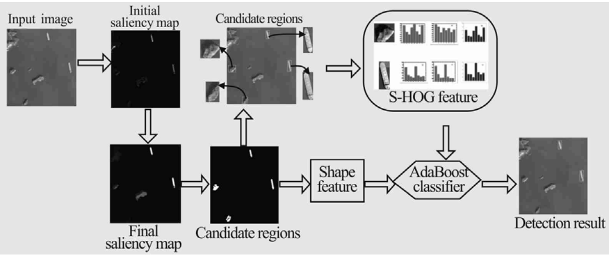


Fig.1 Flow chart of ship detection algorithm proposed in this paper

In this paper, the candidate regions are selected based on the method of visual saliency detection. At present, a lot of novel detection algorithms in the field of saliency detection still suffer from some problems, such as the low resolution of saliency map, the blur of target boundary and the inability of highlighting the whole target. This paper solves the above problems by the combination of the maximum symmetric surround saliency detection method and a synchronization updating mechanism of cellular automata. Finally, the Otsu algorithm is applied to complete the image binarization and candidate target extraction, and then the first time extraction of candidate regions is achieved according to the geometric features of targets.

Ref.[10] proposes a maximum symmetric surround method to calculate the saliency map which can effectively solve the problem of target boundary blur. The specific implementation process of the algorithm contains two main steps. Firstly, the Gaussian differential filter is applied to deal the original image. Secondly, the original saliency map is obtained by combining the image features of color and brightness. The mentioned above algorithm can be described as follows. The pixel saliency value at position (x, y) can be expressed as

$$S(x,y)=\left\|I_\mu(x,y)-I_f(x,y)\right\|, \tag{1}$$

where $I_f(x, y)$ is the corresponding CIELAB image pixel vector in the Gaussian filtered version of the original image, and $I_\mu(x, y)$ is the average CIELAB vector of the sub-image whose center pixel is at position (x, y), which can be given by

$$I_\mu(x,y)=\frac{1}{A}\sum_{i=x-x_0}^{x+x_0}\sum_{j=y-y_0}^{y+y_0}I(i,j), \tag{2}$$

where offsets x_0, y_0 and area A of the sub-image are computed as

$$x_0=\min(x, w-x), \tag{3}$$

$$y_0=\min(y, h-y), \tag{4}$$

$$A=(2x_0+1)(2y_0+1), \tag{5}$$

where w and h are the width and the height of the image.

The method applied above can fast and effectively extract the saliency map with clear target boundary, but it still can not highlight the while salient targets. So we solve the highlighting problem by applying a synchronization updating mechanism based on cellular automata[11].

Before updating the initial saliency map, the original image is segmented into N small super-pixels by the simple linear iterative clustering (SILC) algorithm[12]. Each super-pixel is described by using the mean of color features and coordinates of pixels in its corresponding region. The newly defined neighbors of a super-pixel include not only the adjacent super-pixels, but also the super-pixels sharing common boundaries with its adjacent super-pixels. The cellular automata use each super-pixel as a cell, and the saliency map is updated according to the image's local similarity. The goal of the updating mechanism is set the salience value of the similar region in the image to be the same.

The specific expression of the updating process is described as

$$\boldsymbol{S}^{t+1}=\boldsymbol{C}^*\cdot\boldsymbol{S}^t+(\boldsymbol{I}-\boldsymbol{C}^*)\cdot\boldsymbol{F}^*\cdot\boldsymbol{S}^t, \tag{6}$$

where $\boldsymbol{S}^t$ is the saliency map at time t, specially, $\boldsymbol{S}^0$ is the original saliency map, $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{F}^*$ is the impact matrix, and $\boldsymbol{C}^*$ is the coherence matrix which can be applied to weight the cell's influence on itself updating process. The matrix's element is built as

$$c_i^*=a\cdot\frac{c_i-\min(c_j)}{\max(c_j)-\min(c_j)}+b, \tag{7}$$

where $c_i=1/\max(f_{ij})$ with $i, j=1,\ldots, N$, $[b, a+b]$ is the confidence interval, and a and b are empirically set to be 0.6 and 0.2, respectively.

The row-normalized impact factor matrix $\boldsymbol{F}^*$ mentioned above can be applied to weight the neighbor's influence on one cell's updating process. For one cell, it has a greater influence on the cells with similar color features. $\boldsymbol{F}^*$ is built as

$$\boldsymbol{F}^*=\boldsymbol{D}^{-1}\cdot\boldsymbol{F}, \tag{8}$$

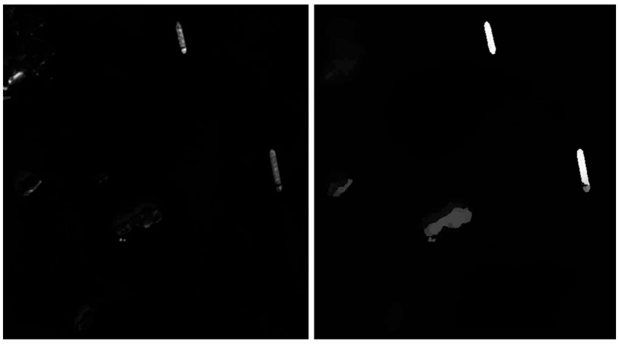
where $\boldsymbol{F}^*$ is the impact factor matrix before normalization, and $\boldsymbol{D}=\mathrm{diag}\{d_1, d_2, \ldots, d_N\}$ is the specific matrix to normalize the original impact factor matrix. The element of matrix $\boldsymbol{F}^*$ can be described as

$$f_{ij}=\begin{cases}\exp\left(\dfrac{-\left\|c_i,c_j\right\|}{\sigma_3^2}\right), & j\in NB(i)\\ 0, & i=j \text{ or otherwise}\end{cases}, \tag{9}$$

where $NB(i)$ is the newly defined neighbor cells, $\|c_i, c_j\|$ is the L2 norm (i.e., Euclidean distance in CIELAB), and for measuring the similarity between super-pixels, σ_3 is a parameter to control the similarity between neighbor cells which is set to be 0.1 in this paper.

Fig.2 shows the original saliency map and the final saliency may. As shown in Fig.2, after updating, the saliency map can better highlight the whole salient object and at the same time guarantee a clear object boundary, which is very suitable for ship segmentation.

The final saliency map needs to be further segmented to obtain candidate regions of the ship targets. Detailed, firstly, due to the difference of saliency value between different images' saliency maps, this paper uses Otsu algorithm to obtain the threshold of saliency map and complete the final saliency map binary. Then, by combining with the geometric features of the ship targets, the candidate regions is extracted.



(a) The initial saliency map (b) The updated saliency map

Fig.2 Comparison of saliency map updating

The detailed geometric features include the area, perimeter of the target and the length and width of the minimum enclosing rectangle, etc. According to the area, length and width of the ship target to be detected, the specific geometric dimensions of the valid candidate area are set as follows: area (40—2 500), long (20—200) and width (10—20).

The candidate regions need the further confirmation by more accurate features to remove false alarms and obtain more accurate detection results.

In this paper, the ship is described by the combination of shape features and S-HOG features. A 27-dimensional eigenvector is constructed, and the target is finally identified and confirmed by the trained AdaBoost classifier.

Feature extraction is the key of target detection, and the result of feature extraction directly influences the consequent detection process. The shape feature and S-HOG feature in this paper not only can effectively distinguish the ship target from the non-ship target, such as ocean waves, clouds and islands, but also are insensitive to the change of ship target size and strong robustness, which are conductive to the subsequent classification and confirmation.

The shape of ship is generally described as long strip with clear and regular contour features. In this paper, the shape feature is described using three features of length-width ratio, compactness and orthogonal-similarity. The length-width ratio can expressed as

$$R_{lw} = \frac{l}{w}, \tag{10}$$

where l and w are the length and the width of circumscribe rectangle. The compactness is defined to describe the degree of circular similarity, which is expressed as

$$Compactness = \frac{P^2}{4\pi A}, \tag{11}$$

where A is the area of the candidate region, and P is the perimeter of the candidate region. And rectangle similarity is expressed as

$$R_{rc} = \frac{A}{S}, \tag{12}$$

where S is the area of circumscribe rectangle.

The features mentioned above are insensitive to the size of targets, and these three features have been proved that they can be used to describe targets accurately, are easy to be achieved and have been widely applied in real-time engineering application.

In optical remote sensing images, the long bar-type shape of ship can be described as the feature that gradients of the two ship sides are symmetrical and generally have high magnitudes in their perpendicular directions. Ref.[9] proposes a new ship S-HOG descriptor by combining the ship features with the traditional HOG features. S-HOG feature makes an improvement than traditional HOG feature by combining of the inherent feature of ship.

S-HOG feature divides the gradient orientations into eight specific gradient bins of 1D—8D, and divides the candidate regions into three blocks of B1—B3. As shown in Fig.3(a), the eight specific gradient bins have the same space angle. As shown in Fig.3 (b), block B1 is the whole candidate region, block B2 is the first half, block B3 is the second half, and these three blocks can thoroughly describe the long bar-type shape.

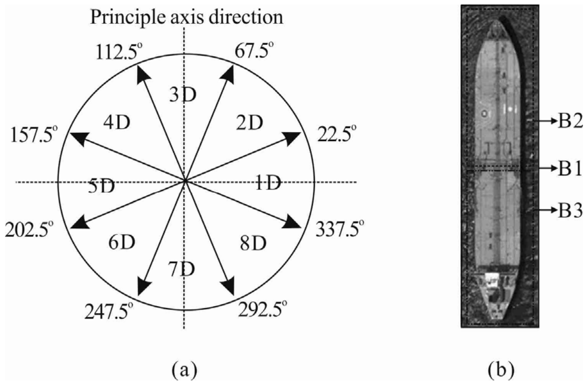


Fig.3 S-HOG feature description: (a) gradient direction interval and (b) the statistical area blocks

As shown in Fig.4(a), for the situation with ship targets, their S-HOG features of bins 1D and 5D are proved to be with high statistical quantities compared with others in all blocks. While as shown in Fig.4(a), for the situation without ship targets (with non-ship targets), the eight bins' statistical quantities do not exist the above appearance, which agree with the analysis of S-HOG theory.

After feature extraction, AdaBoost classifier[13] is used to execute the final identification of the candidate targets. AdaBoost is an iterative algorithm, whose core idea is to train different weak classifiers or basic classifiers for the same training data, and then combine these weak classifiers into a strong classifier. AdaBoost classifier is easy to be achieved and can avoid over-fitting, while still has high classification accuracy[14,15].

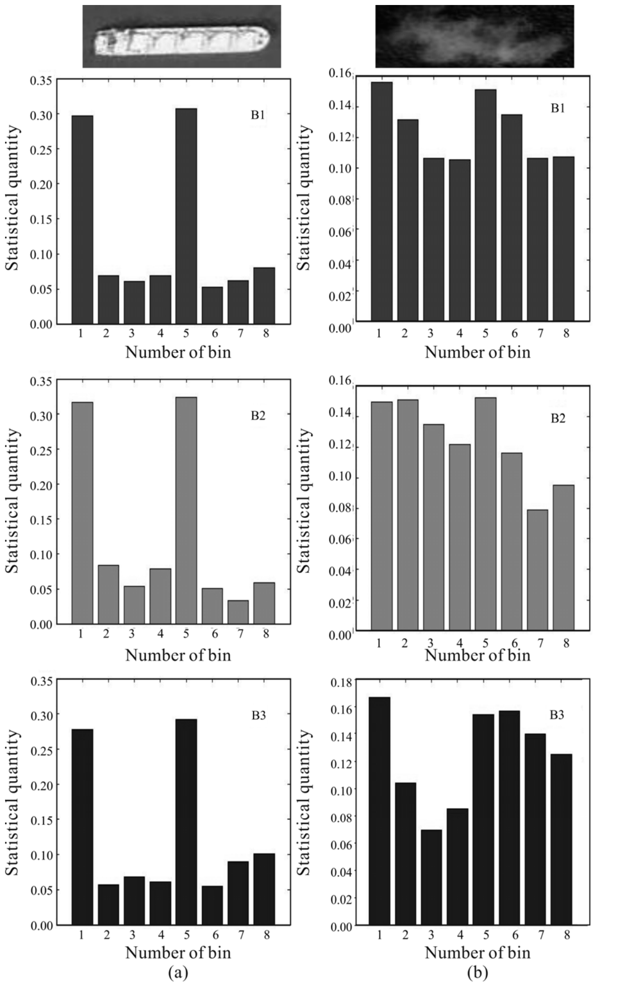


Fig.4 S-HOG feature comparison of positive and negative samples: (a) positive sample and its corresponding S-HOG features for three blocks; (b) negative sample and its corresponding S-HOG features for three blocks

In this paper, an image database is built with 261 remote sensing images at 2 m resolution from the Google Earth database. The size of these images ranges from 2 000 pixel×2 000 pixel to 8 000 pixel×8 000 pixel. These 261 images are divided into two groups, which are the AdaBoost classifier training group containing 232 images and final detection experiment group containing 29 images.

For the AdaBoost classifier training, 407 positive samples and 514 negative samples are intercepted from the 232 images of the AdaBoost classifier training group. The 407 positive samples includes different types of ship target, and the size of positive samples ranges from 20 pixel×10 pixel to 200 pixel×120 pixel. The 514 negative samples includes the non-ship targets, such as clouds, islands, coast line, ocean waves and sea floating objects, and these size also ranges from 20 pixel×10 pixel to 200 pixel×120 pixel.

For final detection experiment, 52 sub-images are intercepted from the 29 images of the final detection experimental group with a uniform size of 1 024 pixel×1 024 pixel. These sub-images contain 238 ship targets with different size and types, and their background includes a variety of complex sea background, such as quiet sea, complex sea surface with ocean waves, and the sea with many islands.

As shown in Fig.5, the algorithm proposed in this paper can accurately detect the ship targets in all kinds of complex sea background. Specially, as shown in the first column of Fig.5, when there are many islands that are similar to the ship target, the algorithm can also distinguish the real ship targets from the false targets accurately.

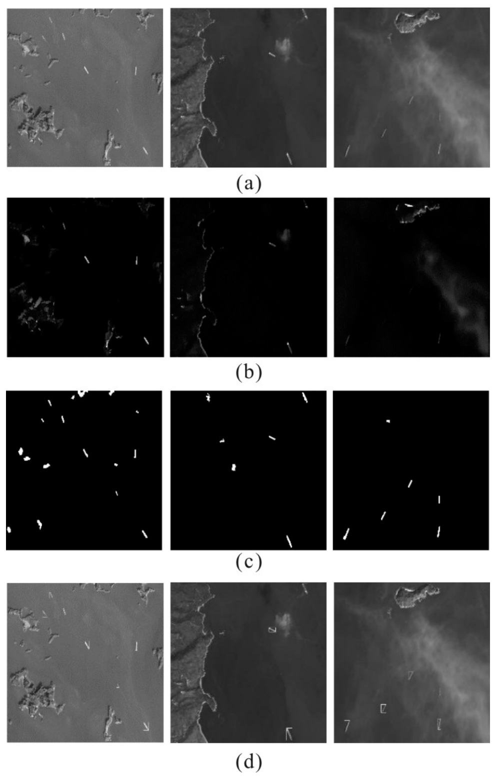


Fig.5 Ship detection results: (a) Original image; (b) Saliency map; (c) Image segmentation with Otsu algorithm; (d) Detection result

In addition, in order to verify the effectiveness of the proposed algorithm, compared with the method in Ref.[9] and the method with S-HOG and AdaBoost (only the S-HOG feature training AdaBoost classifier), the recall ratio and detection precision are used as the evaluation indices. The recall and precision can be expressed as

$$Recall = \frac{N_{\mathrm{DS}}}{N_{\mathrm{TS}}}, \tag{13}$$

$$Precision = \frac{N_{\mathrm{DS}}}{N_{\mathrm{DS}} + N_{\mathrm{DF}}}, \tag{14}$$

where N_{TS} is the total number of real ships, N_{DS} is the number of real detected ships, and N_{DF} is the number of falsely detected ships. The detection results of these three methods are shown in Tab.1

Tab.1 Comparison of detection results of three algorithms

Method	N_{DS}	N_{DF}	*Recall* (%)	*Precision* (%)
Our method	230	8	96.6	96.6
S-HOG + AdaBoost	220	24	92.4	90.2
Method in Ref.[9]	205	37	86.1	84.7

In this paper, we propose a new method for ship detection in optical remote sensing images. The method is simple, robust, and is able to achieve rapid and efficient ship detection in optical remote sensing image. At the stage of extracting candidate regions, the improved saliency map detection method can effectively highlight the whole salient object, and at the same time, ensure a clear object boundary, which is very suitable for ship segmentation. The 27-dimension eigenvector, including both shape feature and S-HOG feature, is insensitive to the size change of targets and the disturbance of ocean waves. The trained AdaBoost classifier can precisely distinguish the real-ship targets from false targets. Experimental results show that the algorithm proposed in this paper can meet the real-time and accuracy requirements of ship detection applications.

References

[1] Y. Wang, L. Ma and Y. Tian, Acta Automatica Sinaca **37**, 1029 (2011). (in Chinese)

[2] Y. Wang and H. Liu, IEEE Transactions on Geoscience and Remote Sensing **50**, 4173 (2012).

[3] Y. Zhao, X. Wu, L. Wen and S. Xu, Opto-Electronic Engineering **35**, 102 (2008).

[4] C. Zhu, H. Zhou, R. Wang and J. Guo, IEEE Transactions on Geoscience and Remote Sensing **48**, 3446 (2010).

[5] Z. Li, D. Yang and Z. Chen, Multi-Layer Sparse Coding Based Ship Detection for Remote Sensing Images, IEEE International Conference on Information Reuse and Integration, San Francisco, 122 (2015).

[6] G. Yang, B. Li, S. Ji, F. Gao and Q. Xu, IEEE Geoscience and Remote Sensing Letters **11**, 641 (2014).

[7] Z. Song, H. Sui and Y. Wang, Automatic Ship Detection for Optical Satellite Images Based on Visual Attention Model and LBP, IEEE Workshop on Electronics, Computer and Applications, Ottawa, 722 (2014).

[8] F. Yang, Q. Xu, F. Gao and L. Hu, Ship Detection from Optical Satellite Images Based on Visual Search Mechanism, IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, 3679 (2015).

[9] S. Qi, J. Ma, J. Lin, Y. Li and J. Tian, IEEE Geoscience and Remote Sensing Letters **12**, 1451 (2015).

[10] R. Achanta and S. Süsstrunk, Saliency Detection Using Maximum Symmetric Surround, IEEE International Conference on Image Processing, Hong Kong, 2653 (2010).

[11] Y. Qin, H. Lu, Y. Xu and H. Wang, Saliency detection via Cellular Automata, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 110 (2015).

[12] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua and S. Süsstrunk, Slic Superpixels, EPFL Technical Report, 149300 (2010).

[13] R. E. Schapire and Y. Singer, Machine Learning **37**, 297 (1999).

[14] J. Sochman and J. Malas, AdaBoost with Totally Corrective Updates for Fast Face Detection, Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 445 (2004).

[15] P. Wang, C. Shen, N. Barnes and H. Zheng, IEEE Transactions on Neural Networks and Learning Systems **23**, 33 (2012).