A NOVEL SHIP DETECTION METHOD FOR LARGE-SCALE OPTICAL SATELLITE IMAGES BASED ON VISUAL LBP FEATURE AND VISUAL ATTENTION MODEL

Sui Haigang [a], *, Song Zhina [b]

[a] State Key Laboratory of Information Engineering In Surveying Mapping and Remote Sensing, Wuhan University, Wuhan, 430079, China- haigang_sui@263.net.
[b] Remote Sensing and Information Engineering. College of Wuhan University, Wuhan, 430079, China- song.zhina@163.com

ICWG III/VII

KEY WORDS: Ship Detection, Large-Scale, Optical Image, Visual LBP Feature, Visual Attention, Support Vector Machine (SVM)

ABSTRACT:

Reliably ship detection in optical satellite images has a wide application in both military and civil fields. However, this problem is very difficult in complex backgrounds, such as waves, clouds, and small islands. Aiming at these issues, this paper explores an automatic and robust model for ship detection in large-scale optical satellite images, which relies on detecting statistical signatures of ship targets, in terms of biologically-inspired visual features. This model first selects salient candidate regions across large-scale images by using a mechanism based on biologically-inspired visual features, combined with visual attention model with local binary pattern (CVLBP). Different from traditional studies, the proposed algorithm is high-speed and helpful to focus on the suspected ship areas avoiding the separation step of land and sea. Large area images are cut into small image chips and analyzed in two complementary ways: Sparse saliency using visual attention model and detail signatures using LBP features, thus accordant with sparseness of ship distribution on images. Then these features are employed to classify each chip as containing ship targets or not, using a support vector machine (SVM). After getting the suspicious areas, there are still some false alarms such as microwaves and small ribbon clouds, thus simple shape and texture analysis are adopted to distinguish between ships and nonships in suspicious areas. Experimental results show the proposed method is insensitive to waves, clouds, illumination and ship size.

1. INTRODUCTION

1.1 Background

Ship detection of remote sensing images (RS images) is very important in maritime rescue, fishing vessel monitoring, immigration control, the defense of territory, naval battle and so on. As ships are typically constructed from large flat metal sheets and hence are usually radar bright and therefore detectable in synthetic aperture radar (SAR) imagery, and due to SAR's ability to work in all-weather conditions and all-time situations, so much work has been done on SAR images in ship detection. And nowadays, with rapid development of the sensor technology, highly resolution optical satellite images can provide more detailed and easily interpreted characteristics to support automated and real-time ship detecting. Compared with SAR imagery, optical satellite images are more visible thus they can provide more detailed and easily interpreted characteristics to help human interpretation. In addition, it is noticed that so many satellites give numberless data in a very short time, so it can apply to real time ship detection in the daytime. Therefor ship detection based on satellite optical images is an essential part in ship detection.

Nevertheless, the problem of ship target detection is a great challenge in the real optical satellite images: 1) Complex background such as waves, clouds, and islands leads to high loss and false alarms in ship detection. 2) Many detection approaches often face a serious dilemma, as no robust feature set or a good model can be defined for the large interclass variability among diverse kind of ship targets.

1.2 Related Work

As discussed above, A few works have been done on detecting ship target from optical satellite images. For a given scene (image), the target detection task can be simply described as "where is the target" (Li, 2011). In ship detecting, we get regions of interest (ROI) which may have ship targets in the first step likewise. Considering the methods used in optical satellite images ship detection, many researchers firstly separate sea areas from land areas using land masking (usually manual), or use threshold-related algorithms. Then many algorithms and strategies are utilized to discriminate ship target, including gray value statistics (Zhao, 2008), texture analysis (Bi, 2012), fractal discrimination (Guang, 2010), shape examination (Zhu, 2010), and so on. However, these types of methods can only work well when the image background is not complicated and the variability of targets is small. In addition to above machine vision approaches, some biologically-inspired computational models have also started exploring target detection studies in computer vision, usually based on visual cortex, showing some promising results(Li, 2011; Siagian, 2007). But visually "salient" object detection predominantly applied to relatively small images, while performed weak in large scale RS images.

1.3 Work Flow of Ship Detection by Using Visual LBP Feature and Visual Attention Model

In order to overcome such problems in ship detection, this paper explores an automatic and effective method for ship detection. The work flow of this method is shown in Figure 1.
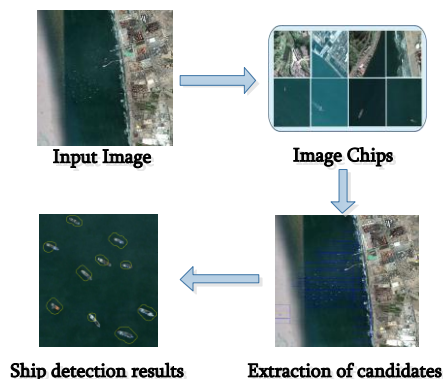
---

* Corresponding author

Figure 1. The work flow structure

Firstly, different from traditional studies, this proposed model simply and directly distinguishes each chip image containing ship targets or not in large images, without sea-land separating pre-processing. In our model, inspired by studies of visual attention, we go further by using combined visual local binary pattern (CVLBP) features and machine learning to predict the whole image areas which may contain ships. Then a Support Vector Machine (SVM) is adopted as the classifier to achieve this decision making task.

After getting the suspicious areas, we focus on finding out the salient regions containing single ship by using the shape and texture statistical features to distinguish ships from these regions.

This paper is structured as 6 sections. Section 2 presents candidate region detection. Section 3 is the single ship detection. The experiments are described and discussed in section 4. Section 5 gives some brief conclusions.

2. PREDICTION OF CANDIDATE REGIONS

The flowchart of the proposed model is shown in Figure 2. Itti-Koch visual attention model can adequately simulate the characteristics of human eyes and quickly find the most "saliency" goals in the scene because it makes full use of a variety of characteristic information such as intensity, color and direction of an image (Itti, 2001). As saliency maps computed from enhanced Itti-Koch model just provide a coarse indication of the structure in the visual contents, we use LBP as saliency map feature extractor to support the target/non-target classification task. LBP is an effective local texture feature descriptor operator which can be used for image classification. What's more, we analyze large images in small chips, mimicking the processing which human image analysts might operate when they deploy multiple eye fixations on a large field imagery, then find ships using SVM.
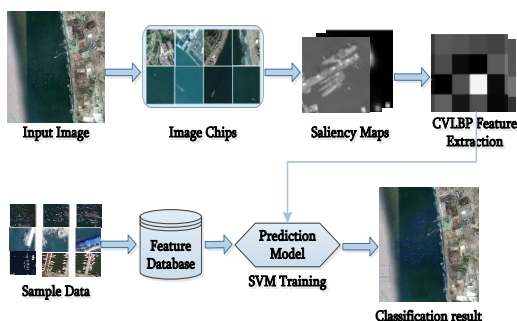

Figure 2. The flowchart of the proposed model

2.1 Saliency Map Computation

The title should appear centered in bold capital letters, at the top of the first page of the paper with a size of twelve (12) points and single-spacing. After one blank line, type the author(s) name(s), affiliation and mailing address (including e-mail) in upper and lower case letters, centred under the title. In the case of multi-authorship, group them by firm or organization as shown in the title of these Guidelines. While in the Itti-Koch model only some biological features (intensity, color, and orientation) were used, here we add gradient features which might be more helpful to distinguish the artificial target. Thus there are four feature channels are employed in this paper: color, intensity, orientation ($0°, 45°, 90°$ and $135°$), and gradient ($0°, 90°$). Then dyadic Gaussian pyramids are used to subsample an image into nine scales, from scale 0 to scale 8. Then for each feature channel among the nine scales, center-surround mechanism is achieved in the model as the difference between fine and coarse scales: The center is a pixel of scale $c \in \{2,3,4\}$, and the surround is the corresponding pixel of scale $s = c + \delta$, with $\delta \in \{3,4\}$. The across-scale difference between two maps, denoted as" Θ " below, is calculated by interpolation to the finer scale and point-by-point subtraction.

The first set of feature maps is intensity contrast, a set of six maps can be got $I(c,s)$, that is

$$I(c,s) = | I(c) \Theta I(s) | \qquad (1)$$

The second set of the color map is similarly constructed for the color channels. Such spatial and chromatic opponent exists for the red/green, green/red, blue/yellow, and yellow/blue color pairs in human primary visual cortex.

Orientation information is obtained from the intensity using oriented Gabor pyramids $O_1(\sigma, \theta)$, where $\sigma \in \{0...8\}$ and $\theta \in \{0°, 45°, 90°, 135°\}$ σ represents scale and θ is orientation. Thus orientation feature are encoded as a group by calculating local orientation contrast between the center and surrounding scales:

$$O_1(c,s,\theta) = | O_1(c,\theta) \Theta O_1(s,\theta) | \qquad (2)$$

Likewise, in the orientation channel, gradient feature $O_2(\sigma)$ is obtained from intensity adopting Sober operator, where scale $\sigma \in \{0...8\}$ and orientation $x = 0°, y = 90°$. It is defined as follows:

$$O_2(\sigma) = \sqrt{G_x(\sigma)^2 + G_y(\sigma)^2} \qquad (3)$$

In general, 48 feature maps are computed: 6 for intensity, 12 for color, 24 for orientation, and 6 for gradient. Next，each set of feature maps is combined into a " conspicuous map" by a simple summation of these maps after scaling to a fixed dynamic range, so we get four saliency maps with I for intensity, C for color, O_1 for orientation, and O_2 for gradient.

2.2 CVLBP Feature Computation

Local Binary Pattern (LBP) was first proposed by Ojala in 1998 (Ojala, 2002). The original 3×3 neighbourhood window has its center pixel as threshold. If the value of a pixel in the neighborhood is larger than the center pixel, it will be marked

with 1, otherwise 0. Then the marked values adversely affect by the binomial weights given to the corresponding pixels and obtained values are summed for the LBP number of this texture unit. After obtaining the labeled LBP image $f_i(x, y)$, the LBP histogram can be defined as:

$$H_i = \sum_{x,y} I\{f_i(x, y) = i\}, \ i = 0, ..., n-1 \qquad (4)$$

where n = the number of different labels
$I\{A\}$ =1 if A is true and 0 false

Here we use the "Uniform LBP" which is invariant in both gray-scale and rotation.

Figure 3 presents the feature extraction procedure using CVLBP model. In this paper, we employ a circular neighbourhood LBP operator with 12 pixels and a radius of 2.5 pixels (totally 59 bins) to extract CVLBP features of image chips, so each saliency map has 59 dimensional vector, and finally a $4 \times 59 = 236$ dimensional vector can be obtained.
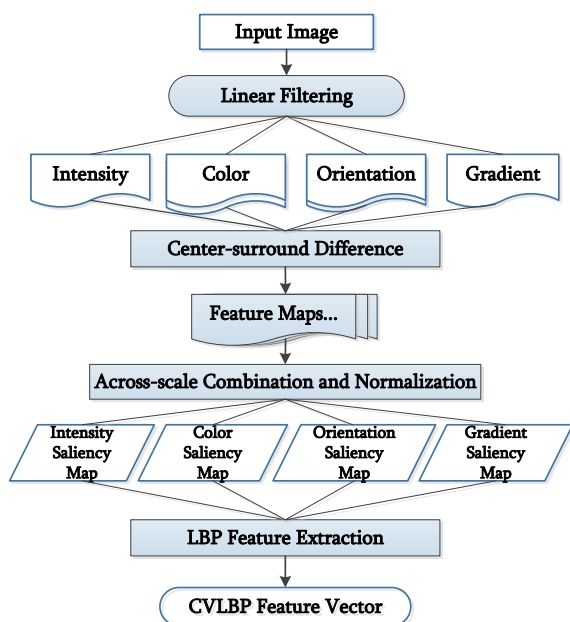

Figure 3. Feature extraction using CVLBP model

2.3 Sample Training and Prediction Using SVM

After getting CVLBP feature of all sample image chips, we manually label some sample chips that contain ship targets with 1 and non-ship targets with 0. After that, all sample image chips are sent to SVM for training. In this paper, SVM were used for its convenience. Considering the high nonlinearity feature of the vectors' distribution, a radial basic function (RBF) based SVM is adopted. Furthermore, for the normalized input, the parameters of SVM can be optimized automatically and no tuning is needed.

3. SINGLE SHIP DETECTION

After getting the suspicious areas, for more accurate salient region extraction of ships, we use iterative localized interactions feature combination strategy relies on simulating local competition between neighbouring salient locations (Itti, 2001) as it can eliminate the order of magnitude differences from

different feature extraction proceedings. The salient are shown in Fig. 6. Then we use the method proposed in (Walther, 2006) to mark the salient regions, it can be accessed by selective attention and subsequently validated as actual objects. So we get the rough region of the salient objects, not simply circle them.

To get a high precision of single ship detection, we use texture statistical features to distinguish ships from microwaves and small ribbon clouds. In this paper, we use statistical texture factors such as mean values, variance values, third-order moment values and texture uniformity altogether to build a combined decision model to detect ships in the saliency regions. Therefore, we statistic these values among more than 500 small typical pictures from salient regions and detect ship target by using Decision Tree Analysis (Zhu, 2002).

4. EXPERIMENTAL RESULT AND DISCUSSION

In order to illustrate the effectiveness of our method in detecting ship targets, we design four broad area land-and-sea optical RS images including sea wave, clouds, islands, and various types of ships, and corresponding ship target ground truth is manually marked. We compare our model with LBP, LMP (Zhu, 2010) and Siagian-Itti's gist (Siagian, 2007) features. LMP features is effective to enhance the representation ability of the featureset in shape and gray space. Siagian-Itti's gist features computational model is similar to our research and performs very well in object detection and scene recognition.

We evaluate the detection results by recall R and precision P:

$$R = TP / (TP + FN) \qquad (5)$$
$$P = TP / (TP + FP) \qquad (6)$$

Where TP, FP and FN are the number of true positive, false positive and false negative targets respectively. A higher recall R ratio means we find more genuine ship targets from images, and a higher precision P ratio means lower false alarms in the detecting task. The comparison of predictive recall and precision ratio of four methods is given in Table 1. As can be seen in Table I, our model can performance higher recall ratio and significantly reduce the false alarm compared with other model. LBP and LMP feature only describes local or shape texture feature of an image, and RS images is so complicated that are not a distinguishable method to pick the suspicious areas out of the negative ones especially the similar texture pattern samples. Gist feature has introduced some visual saliency cues of scenes analysis which focus more on overall statistics and contextual information in the entire image, but it can hardly remove those false alarms that have similar characteristics with real ships. So we found that visual attention mechanism can not only detect ship targets in calm sea, but also suitable for ship detection in complex sea containing a large number of sea clutter. Some single ship detection results of the image blocks are shown in Figure.4, we can see our method improve the single ship detection accuracy further more.

Dataset	CVLBP	
	R (%)	P (%)
LBP	87.5	75.1
LMP	93.2	85.5
Gist	92.5	79.1
CVLBP	95	90.3

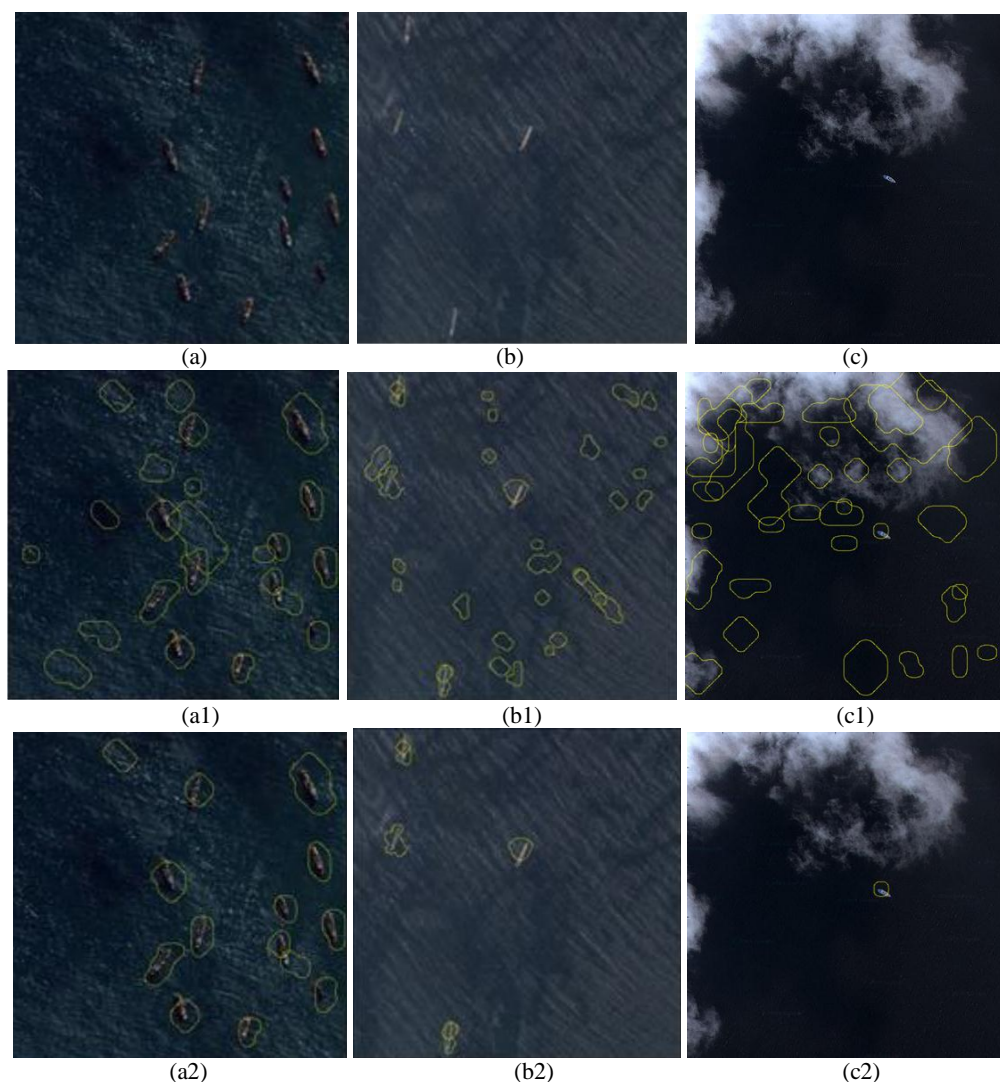Table 1. Comparison Result of Ship Target Detection

Figure 4. Single ship detection results
(a), (b) and (c) are original images; (a1), (b1) and (c1) are results of only-visual attention method; (a2), (b2) and (c2) are results of our method.

5. CONCLUSION

In this paper, we apply a ship detection method based on a new feature extraction method, combining visual attention and visual LBP mechanism (CVLBP feature). Experiments based on complicated optical RS images show our method has good performance on ship detection. It is due to our use of two complementary sets of model: Sparse saliency highlighted in sea utilizing visual attention and detail description using LBP features. However, for the purpose of practical application of ship detection, what we have done is just show the suspicious ship candidates, we have to classify single ship targets respectively, and this is mainly work we will study in the future.

REFERENCES

Zhicheng Li, Laurent Itti, 2011. Saliency and gist features for target detection in satellite images. *IEEE Trans. Image Processing,* 20(7), pp. 2017 - 2029.

Zhao Ying-hai, Wu Xiu-qing, Wen Ling-yun, Xu Shou-shi, 2008. Ship target detection scheme for optical remote sensing images," *Opto- Electronic Engineering,* 35(8), pp. 105-109.

Fukun Bi, Bocheng Zhu, Lining Gao, and Mingming Bian, 2012. A visual search inspired computational model for ship detection in optical satellite images. *IEEE Trans. Geoscience and Remote Sensing*, 9(4), pp. 749 - 753.

Yang Guang, Lu Qichao, and Gao Feng, 2011. A novel ship detection method based on sea state analysis from optical imagery. Image and Graphics (ICIG), 2011 Sixth International Conference, pp. 466-471.

Changren Zhu, Hui Zhou, Runsheng Wang, and Jun Guo, 2010. A novel hierarchical method of ship detection from spaceborne optical image based on shape and texture features. *IEEE Trans. Geoscience and Remote Sensing*, 48(9), pp. 3446 - 3456.

Zhicheng Li, Laurent Itti, 2011. Saliency and gist features for target detection in satellite images. *IEEE Trans. Image Processing*, 20(7), pp. 2017 - 2029.

C. Siagian, L. Itti, 2007. Rapid biologically-inspired scene classification using feature shared with visual attention. *IEEE Trans. Pattern Analysis and Machine Intelligence*. 29(2), pp. 300–312.

L. Itti, C. Koch, 2001. Feature combination strategies for saliency-based visual attention systems. *Journal of Electronic Imaging*, 10(1), pp. 161-169.

T. Ojala, M. Pietikänen, T. Mäenpää, 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(7), pp. 971–987.

L. Itti and C. Koch, 2001. Feature Combination Strategies for Saliency-Based Visual Attention Systems. *Journal of Electronic Imaging*, 10(1), pp. 161-169.

ACKNOWLEDGEMENTS