Ship detection from optical remote sensing images based on PLSA model

ZHOU Hui[1,2], GUO Jun[1], ZHU Changren[1], WANG Runsheng[1]

1. *ATR Key Laboratory, School of Electronic Science and Engineering, National University of Defense Technology, Hunan Changsha* 410073, *China*;
2. *Beijing Institute of Tracking and telecomunications technology, Beijing* 100094, *China*

Abstract: Ship detection is one of the important areas in remote sensing applications. However, many ship detection approaches often face a difficult dilemma between low detection rate and high false rate, because of the un-matching between object and its features caused by the complicated characteristics of remote sensing images. Therefore, this paper proposes a novel detection algorithm based on Probabilistic Latent Semantic Analysis (PLSA). It firstly describes the object in terms of the probability combination of latent aspects generated by PLSA, then discriminates the latent aspects model of object by statistics recognition method to obtain the final detection result. The generated latent aspects model represents the joint probability of objects and their features, and gives an explanation for the above un-matching problem by the probability distribution of latent aspects. The performance of the proposed algorithm is demonstrated through the ship detection in various optical remote sensing images, and substantiated using quantitative criteria.

Key words: ship detection, optical remote sensing images, probabilistic latent semantic analysis, tempered expectation maximization, local-binary-pattern operator

CLC number: TP751.1 **Document code:** A

Citation format: Zhou H, Guo J, Zhu C R and Wang R S. 2010. Ship detection from optical remote sensing images based on PLSA model. *Journal of Remote Sensing.* **14**(4): 663—680

1 INTRODUCTION

Ship monitoring is one of the traditional tasks for coastal countries. As an active microwave sensor, synthetic aperture radar (SAR) is one of the commonly used techniques for ship monitoring in current years, because it is independent of the weather condition and time of day (or night). Accordingly, lots of SAR image ship detection algorithms are proposed (Wanielik & Stock, 1989; Tello *et al*., 2005; Li & Chong, 2007; Tian *et al*., 2007; Yang *et al*., 2008). However, it is more possible to monitor ship based on optical remote sensing images for real-time surveillance considering the significantly increasing amount of optical sensors and the wide cover and high spatial resolution of the images acquired by these sensors. As a result, some researchers (Zhao *et al*., 2008) have paid attention to this problem.

Information in remote sensing images is often disturbed because of the change of season, weather and work condition of sensor. In the images of complicated ocean area, there are lots of false alarms which have similar characteristics with ship objects, such as ocean waves, clouds, islands, etc. In addition, the ship objects also represent various characteristics under different image resolutions and imaging mechanism. Therefore, it is a challenging difficulty to eliminate these false alarms in ship detection. Traditionally, researchers often consider it as a binary-detection process. For example, Yang *et al*. (2008) and Zhao *et al*. (2008) converted an image into a binary image by a heuristic threshold. Li and Chong (2007) and Tian *et al*. (2007) validated the probable ship after estimating the ship model based on prior distribution. However, because of the simple modeling strategies of these methods, they do not represent diversities between ship objects and false alarms. As a result, they have to face a difficult dilemma between low detection rate and high false rate, or can not lead to qualitatively convincing results while being robust and operational.

A prospective approach to solve this problem, in our opinion, is to translate the ship detection into multi-class problem based on statistics learning. Taking advantage of Support Vector Machines (SVMs) (Vapnik, 1998), we can fulfill ship object detection process by combining multiple effective features. In practical classification, however, there exist two typical phenomena:

Phenomenon 1 corresponds to the case that the samples with the same class category possess different characteristics, such as various ship samples shown in Fig.1(a).

Received: 2009-05-14; **Accepted**: 2009-09-13

Foundation: Foundation of ATR Key Laboratory (No. 9140C8004011007).

First author biography: ZHOU Hui (1981—), male. He is currently pursuing the Ph.D. degree in the ATR Lab, National University of Defense Technology, China. His research interests include image analysis and understanding, pattern recognition, and intelligent processing of remote sensing images. He authored 8 technical papers. E-mail: hui_zhou_atr4@yahoo.com.cn

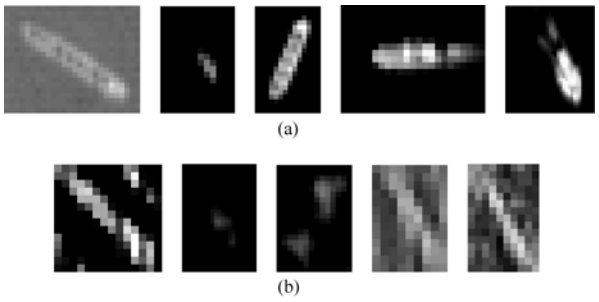


Fig. 1 Disturbance of various samples in real classification
(a) Phenomenon 1; (b) Phenomenon 2

Phenomenon 2 corresponds to the case that the samples with the different class categories possess similar characteristics, such as the samples of cloud and ocean wave shown in Fig.1(b), which have some similar characteristics with ship.

Although the classification-based approach can overcome the effects of phenomenon 1 by analyzing the distribution of sub-class in a class category, it can not deal with phenomenon 2 for associating labeled patterns directly with their extracted features when recognizing these patterns. In other words, when there are fewer training samples, it is difficult to make the SVM classifier closer to real object model; when there are more training samples, it is impossible to fulfill the learning process.

According to the above discussion, this paper first introduces the Probabilistic Latent Semantic Analysis (PLSA) (Hofmann, 2001) into ship detection. When modeling ship object, PLSA does not represent object directly by its extracted features, but generates a latent aspects model by learning "object-feature" co-occurrence frequency matrix (feature frequency matrix for short). Then it uses this model to represent object in terms of a probabilistic decomposition of latent aspects. In this way, we can give an explanation for the previous phenomena by probabilistic diversities of latent aspects in different objects. Modeling object based on PLSA not only simplifies the work of labeling training samples but also reduces the difficulty of the training and predicting of SVM classifier. Meanwhile, we also use an adaptive method to select an optimal temperature control parameter for Tempered Expectation Maximization (TEM) in order to improve the model fitting of PLSA. The performance of the proposed algorithm is demonstrated through the ship detection in various optical remote sensing images, and substantiated using quantitative criteria.

2 PLSA METHODOLOGY

Stemming from a view of Latent Semantic Analysis (LSA), PLSA possesses a solid foundation in mathematical inference and a well-used generative data model. Its main idea is to discover topics in a collection of documents that are represented by the frequencies of words from a vocabulary. The PLSA model has been widely used in applications such as statistical text analysis (Hofmann, 2001) and content-based image retrieval (Quelhas *et al*., 2007; Wang *et al*., 2009), and has been proved to provide better words matching in information extraction. In our case, the documents correspond to the identified objects, the vocabulary corresponds to the feature frequency matrixes of shape and texture, and the topics to be discovered correspond to the different kinds of objects, such as ships, ocean waves, clouds, islands, etc.

Suppose we have a collection of objects $O=\{o_1,o_2,\cdots,o_N\}$ represented by a feature set $F=\{f_1,f_2,\cdots,f_M\}$, and accordingly summarize the data in a rectangular $N\times M$ feature frequency matrix $N=\left(n(o_i,f_j)\right)_{ij}$, where $n(o_i,f_j)$ denotes the number of times the feature f_i occurred in object o_j. Each observation (o_i,f_j) is associated with a group of latent aspects (topics) $Z=\{z_1,z_2,\cdots,z_K\}$, where K is a constant. The PLSA technique uses a graphical model (as shown in Fig. 2) for the joint probability $P(o_i,f_j,z_k)$ of the objects and their features. The filled nodes in Fig. 2 indicate observed random variables, whereas the unfilled node is unobserved. The generative model $P(o_i,f_j)$ for the feature content of objects can be computed by using Eq.(1):

$$\begin{aligned} P(o_i,f_j)&=P(o_i)P(f_j|o_i),\\ P(f_j|o_i)&=\sum_{k=1}^{K}P(f_j|z_k)P(z_k|o_j) \end{aligned} \tag{1}$$

where $P(o_i)$ denotes the probability of observing a particular object o_i, $P(f_j|z_k)$ denotes the conditional probability of a specific feature f_j conditioned on the unobserved topic variable z_k, and $P(z_k|o_i)$ denotes an object specific probability distribution over the latent variable space. We should note that, this model introduces a conditional independence assumption, namely that o_i and f_j are independent conditioned on the state of the associated latent variable z_k (Hofmann, 2001).

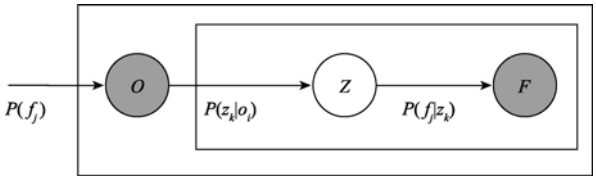


Fig. 2 Graphical model representation of the PLSA

Following the likelihood principle, one determines the parameters of PLSA model by maximization of the log likelihood function in Eq. (2):

$$L=\sum_{i=1}^{N}\sum_{j=1}^{M}n(o_i,f_j)\log P(o_i,f_j) \tag{2}$$

The standard procedure for maximum likelihood estimation in PLSA model is Expectation-Maximization (EM) algorithm. EM alternates two steps: (1) an expectation (*E*) step where posterior probabilities are computed for latent variables, based on the current estimates of the parameters, (2) a maximization (*M*)

step, re-estimate the parameters in order to maximize the expectation of the complete data likelihood. By applying Bayes' formula, one can obtain *E*-step as follows:

$$P(z_k|o_i, f_j) = \frac{\left[P(f_j|z_k)P(z_k|o_i)\right]}{\sum_{l=1}^{K}\left[P(f_j|z_l)P(z_l|o_i)\right]} \tag{3}$$

In the *M*-step, one has to maximize Eq. (2) and then re-estimate parameters $P(f_j|z_k)$ and $P(z_k|o_i)$ as follows:

$$P(f_j|z_k) = \frac{\sum_{i=1}^{N} n(o_i, f_j)P(z_k|o_i, f_j)}{\sum_{m=1}^{M}\sum_{i=1}^{N} n(o_i, f_m)P(z_k|o_i, f_m)} \tag{4}$$

$$P(z_k|o_i) = \frac{\sum_{j=1}^{M} n(o_i, f_j)P(z_k|o_i, f_j)}{\sum_{j=1}^{M} n(o_i, f_j)} \tag{5}$$

The *E*- and the *M*-steps are iterated until the criterion, which is defined in Eq. (6), is less than a threshold or the number of iterations exceeds a predetermined value.

$$E(L) = \sum_{i=1}^{N}\sum_{j=1}^{M} n(o_i, f_j)\sum_{k=1}^{K} P(z_k|o_i, f_j)\log\left[P(f_j|z_k)P(z_k|o_i)\right] \tag{6}$$

3 DETECTION ALGORITHM DESCRIPTION

The proposed ship detection algorithm includes two stages (as shown in Fig. 3) and the arrows with marker at the beginning in Fig. 3 show the detection flow. The first stage is ship candidate region detection, which firstly identifies the ocean region by ocean-land segmentation, then extracts the probable objects according to some prior distributions of ship features, such as gray scale, edge or texture. The second stage is ship candidate region recognition, which aims to distinguish the ship objects in candidate regions by using object recognition approach. Detailed implementations are described in section 3.1 and 3.2.

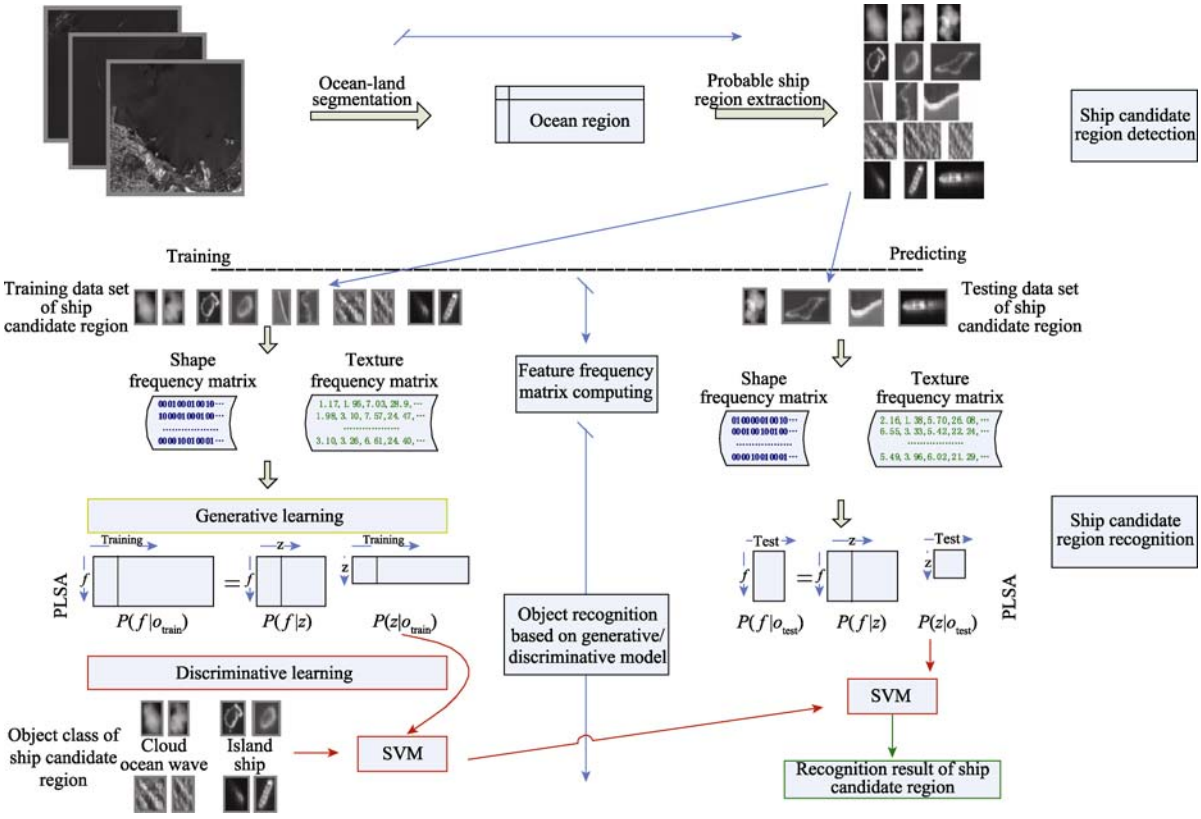


Fig. 3 Overview of the proposed ship detection algorithm

3.1 Ship candidate region detection

In the process of ship detection from optical remote sensing image, some pixels of land area are often detected as ship objects. Therefore, we should firstly partition the ocean and land in order to eliminate the influence of land area. This paper identifies the range of land area by using the information in GIS database. When obtaining the ocean detection area, we apply the following steps to carry out the ship candidate detection based on the characteristics of ship's edge and shape.

Step 1: Image pre-processing. To enhance the distinction between the ship and sea in the image, we choose a new method for computing a mixed image Mixu(*x*,*y*) as defined in Eq. (7):

$$\text{Mixu}(x, y) = u(x, y) + \alpha\text{Magu}(x, y) \tag{7}$$

where $u(x,y)$ denotes the gray value of the image pixel (x,y),

Magu(x,y) denotes the gradient magnitude of this pixel in terms of Sobel operator. α is a coefficient, equaling 1 in our experiments.

Step 2: Image coarse segmentation. First of all, the mixed image Mixu(x,y) is operated with a threshold which is adaptively obtained by Otsu's maximum variance method (Otsu, 1979). Then, morphologic operating is adopted to fill holes within regions, and to eliminate some very thin lines. Lastly, some singular regions are filtered out according to the ratio R_{region} of the length to width of the region bounding rectangle and region size S_{region}. In this paper, the value of R_{region} is 1.5 and the value of S_{region} is related with the image resolution.

Step 3: Image refined segmentation. Although the threshold in coarse segmentation is adaptively set, it is still a global statistic, and then the result of segmentation must have some boundary localization errors. Level set with the Chan-Vese model is adopted to refine image segmentation in this paper. Chan and Vese (2001) propose an algorithm based on the Mumford-Shah model, which could gives an optimal partition of two classes, and the simplified energy function of this algorithm is defined as follows:

$$F(C) = F_o(C) + F_b(C) = \int_{\text{inside}(C)} |u(x,y) - c_o|^2 \,\mathrm{d}x\mathrm{d}y + \int_{\text{outside}(C)} |u(x,y) - c_b|^2 \,\mathrm{d}x\mathrm{d}y \quad (8)$$

where c_o is the mean value inside of the curve C, and c_b is the mean value outside of the curve C. The minimum of the Chan-Vese's model energy will be an optimal piecewise smooth approximation of the edge.

Fig. 4 gives the intermediate results of ship candidate detection. In Fig.4, (a) is the original image, (b) is the new mixed image Mixu(x,y), and (c) is the final detection result. Three small patches from left to right in Fig. 4(d) and (e) correspond to the original images, the initial segmentation results and the refined segmentation results of two objects in Fig. 4(c), respectively. It is easily concluded that the refined segmentation results via level sets approach are closer to the true object edges.

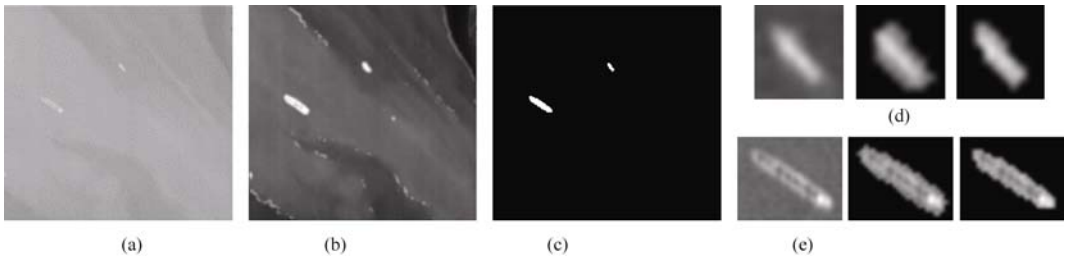


Fig. 4 Detection of ship candidate region

3.2 Ship candidate region recognition

Considering the aim of ship candidate region detection is to locate all probable objects, many false alarms must be produced. Therefore, in order to eliminate these false alarms in the following stage, i.e. ship candidate region recognition, we first employ the PLSA technique to model the candidate objects, and then recognize the generative model by using SVM classifier. Since the PLSA model is generated from the feature frequency matrix, we should explain how to compute this matrix before describing the detailed recognition method.

3.2.1 Feature frequency matrix computing

This paper computes the feature frequency matrix based on two kinds of features-shape and texture. Shape feature, which is computed from the binary image obtained after image refined segmentation, mainly represents the object characteristics. Texture feature, which is computed from the candidate region, represents some context information as well as the object characteristics.

Shape frequency matrix

Considering the different spatial resolutions of remote sensing images, we select the following rotation- and scale-invariant shape features: rectangularity, eccentricity and the first three invariants of Hu moment. In order to form the feature frequency matrix, we quantize each shape feature s into a certain value v_i of vocabulary V according to a nearest neighbor rule:

$$s \mapsto Q(s) = v_i \leftrightarrow \text{dist}(s, v_i) \leqslant \text{dist}(s, v_j), \forall j \in \{1, \cdots, N_v\} \quad (9)$$

where dist(x,y) denotes the Euclidean distance between x and y, and N_v denotes the size of vocabulary V. More specifically, we apply the k-means algorithm to each shape feature extracted from training images, and keep the means as the values of corresponding vocabulary V, then quantize the feature into a v_i of its vocabulary according to Eq. (9). Finally, the shape feature frequency matrix $H(s)$ is constructed by Eq. (10) based on the obtained v_i:

$$H(s) = \left(h^l(s)\right)_{l=1,\cdots,L}, h^l(s) = \left(h_0^l(s), h_1^l(s), \cdots, h_{N_v^l}^l(s)\right) \quad (10)$$

where $h_i(s)$=1 when $Q(s)$=v_i, otherwise $h_i(s)$=0. L denotes the category of shape feature (L=5 in this paper) and N_v^l denotes the vocabulary size of the l-th shape feature.

Texture frequency matrix

Currently, numerous texture descriptors have been presented, and local-binary-pattern (LBP) operator (Ojala *et al*., 2002) has been proved to be a theoretically simple yet very effective multi-resolution statistical texture descriptor in terms of the characteristics of the local structure. However, the extension to the LBP operator based on "uniform" patterns still has some shortcomings: it discards some important texture information and is sensitive to noise. Thus, this paper uses a novel extended LBP operator presented in the previous work (Zhou *et al*.,

2008). It further assigns proper labels to different non-uniform patterns according to the measures of similarity and structure so as to improve the description ability of the LBP operator. For brevity, some detailed definitions can be found in references (Ojala *et al*., 2002; Zhou *et al*., 2008). Meanwhile, the texture feature based on LBP operator is a discrete histogram, and after normalizing this histogram, we can use it as the texture feature frequency matrix.

3.2.2 Object discrimination

The object recognition process includes two steps: training process and predicting process, which are illustrated at the bottom of Fig. 3. Training proceeds in two sub-steps. First, the model parameters $P(f_i|z_k)$ and $P(z_k|o_i^{\text{train}})$ are generated by fitting the PLSA model to the set of training samples based on EM algorithm, and each training sample is then represented by a *K*-Dimension vector $P(z|o_i^{\text{train}})$. Second, a SVM classifier is trained given the vector $P(z|o^{\text{train}})$ of each training object and its class label. We use the LIBSVM package (Chang & Lin, 2001) to implement the SVM classifier and the model parameters of SVM are obtained by cross-validation.

Predicting process also proceeds in two sub-steps. First, $P(z|o^{\text{test}})$ of the test samples are achieved by running EM in a similar manner to that used in training process, but now, only the coefficients $P(z|o^{\text{test}})$ are updated in each *M*-step with the learned $P(f_i|z_k)$ kept fixed. The result is that the test sample is represented by a *K*-Dimension vector $P(z|o_i^{\text{test}})$. Second, each $P(z|o_i^{\text{test}})$ is then predicted by the SVM classifier to get the final recognition result.

In order to avoid overfitting in a standard procedure of EM algorithm, this paper use the Tempered Expectation Maximization (TEM) (Ueda & Nakano, 1988) for model fitting. By applying TEM algorithm only the *E*-step is modified as follows:

$$P(z_k|o_i,f_j)=\frac{\left[P(f_j|z_k)P(z_k|o_i)\right]^{\beta}}{\sum_{l=1}^{K}\left[P(f_j|z_l)P(z_l|o_i)\right]^{\beta}} \tag{11}$$

where $\beta(0<\beta<1)$ is the temperature controlling parameter.

However, a better fitting result based on TEM algorithm can not be produced at any β in practice. Thus, according to Hofmann (2001), this paper uses the following method to select an optimal β:

(1) Set β=1 and perform EM until the perplexity defined in Eq. (12) stops descending.

(2) Decrease β and set $\beta=\eta\beta$, where η=0.95.

(3) As long as the perplexity improves (non-negligible) continue TEM iterations at this value of β, otherwise goto step 2.

(4) Selection stopping on β, i.e., stop when decreasing β does not yield further improvements.

In the above process of iteration, the perplexity is a measure to assess the generalization performance of a model (Hofmann, 2001), and it can be defined as Eq. (12):

$$\text{perplexity}=\exp\left[-\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\tilde{n}(o_i,f_j)P(o_j|f_j)}{\sum_{i=1}^{N}\sum_{j=1}^{M}\tilde{n}(o_i,f_j)}\right] \tag{12}$$

where $\tilde{n}(o_i,f_j)$ denotes the feature frequency matrix of hold-out data. In this paper, we select the same amount of samples from the test data to construct the hold-out data.

4 EXPERIMENT RESULTS

In this section, many different remote sensing data were used to demonstrate the performance of the proposed ship detection algorithm. We firstly introduced the testing environment of the experiments and the quantitative evaluation measures, then investigated the various choices of some important algorithm parameters, lastly compared the results of our algorithm with those of two other commonly used classification methods, i.e. PCA+KNN and Feature+SVM.

4.1 Experiment setup

316 optical remote sensing images acquired by CBERS and SPOT4, the spatial resolutions of which were different from 10m to 5m, were utilized in the experiments. Some typical testing images were shown in Fig. (5), including different false alarms such as clouds, islands, ocean waves, etc. 1653 image patches were obtained by ship candidate region detection, and they contained all ship objects presented in the testing image. 495 image patches were random selected as the training sample set and the rest formed the testing sample set. All experiments involved were executed on a 2.66 GHz Intel Pentium IV workstation and the proposed detection algorithm was implemented

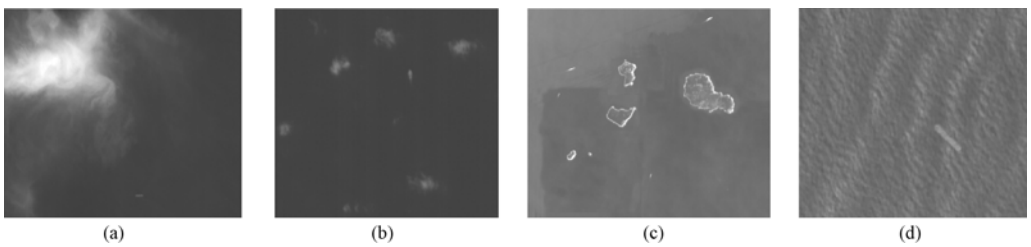


Fig. 5 Typical samples of the testing image

in VC++6.0. Some detailed implementation had been described in section 3 and the setting of important parameters would be specified in section 4.2.

This paper used Recall and Precision of ship object to evaluate the performance of different detection approaches. For the given test image set, $|W_G|$ is the number of ship objects labeled by human, and $|W_M|$ is the number of ship objects detected by certain approach where $|W_C|$ is the number of correctly detected ship objects. Then, Recall and Precision can be defined as:

$$\text{Recall} = |W_C| / |W_G|, \text{Precision} = |W_C| / |W_M| \quad (13)$$

Recall represents the detection completeness of ship objects, while Precision evaluates the tendency of the algorithm for false alarms.

4.2 Algorithm parameters analysis

In this section, we mainly investigated two key parameters involved in our algorithm, temperature controlling parameter β and latent aspect number K. We would carry out a set of experiments under different conditions to select optimal parameter setting.

4.2.1 Optimizing the parameter β

We had carried out four experiments (two experiments for texture and two for shape) to validate the feasibility of the optimal β selecting method. The result of each experiment is shown in Fig. 6(a) and (b), respectively. In this figure, vertical coordinate corresponds to the value of perplexity defined in Eq. (12), horizontal coordinate corresponds to the number of iteration and each curve represents the convergence results at different β. For clear expression, we only marked the result of β=1 (i.e. the result of standard EM algorithm) and that of optimal β, which are emphasized by the curves through circles and pentagrams, respectively. According to the experiment results, the perplexity value of model fitting based on TEM are obviously lower than the value of standard EM, i.e. the model fitting performance of TEM is better than that of EM.

4.2.2 Optimizing the parameter K

In the PLSA technique, K is the number of latent aspects Z, which should be specified by human. It determines the algorithm performance as denoted in references (Quelhas *et al.*, 2007; Wang *et al.*, 2009). Therefore, we should first analyze the diversities of detection results with different number of latent aspect. Shape and texture feature were respectively tested. We test 20 times at each number of K and then computed the mean and variance of each group of results. Fig. 7 shows the experiment results. Meanwhile, to further illuminate the advantage of TEM algorithm, we also gave the detection result of EM algorithm in this figure.

In Fig. 7, the square points and diamond points correspond to the average result with each K based on TEM algorithm and EM algorithm respectively, the line segment between two cross markers of each point represents the double-times variance of the result with corresponding K, and the dashed line denotes the

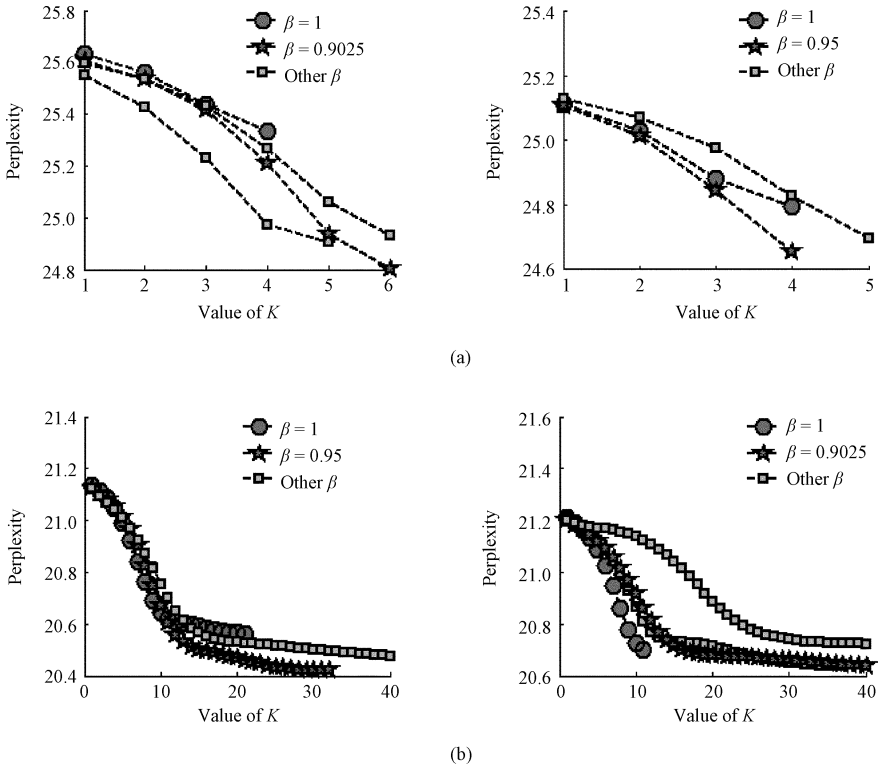


Fig. 6 Model fitting results with different feature occurrence matrixes by using TEM and EM algorithms, respectively
(a) Shape feature; (b) Texture feature

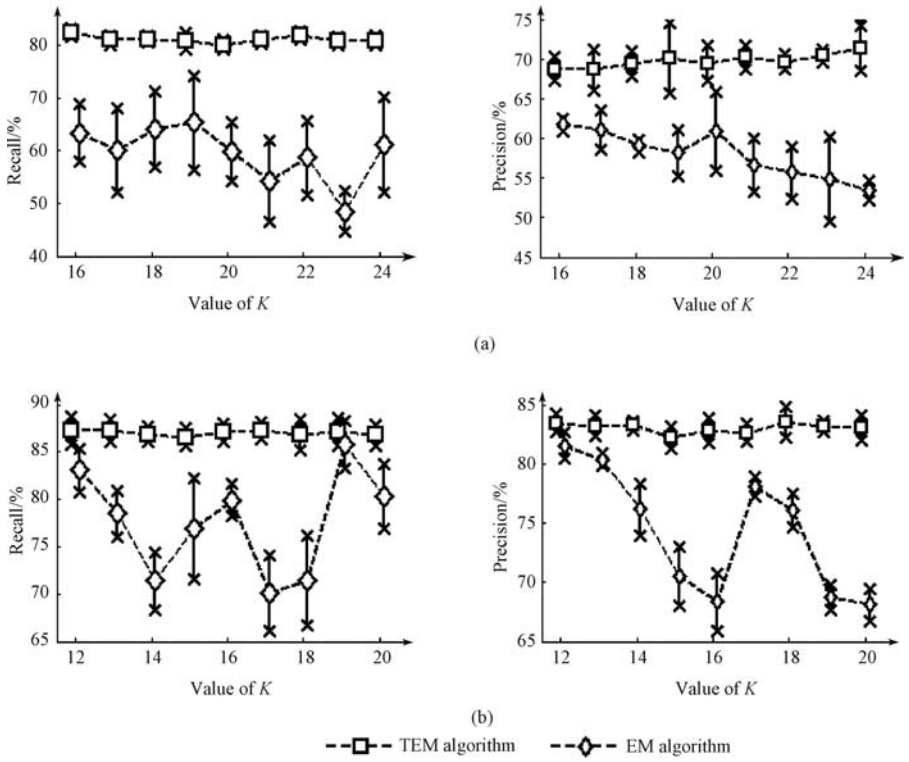


Fig. 7 Detection results with different *K* by using TEM and EM algorithms, respectively
(a) Shape feature; (b) Texture feature

changing tendency of the detection results with K varying. According to Fig. 7, the mean of Recall and Precision based on TEM is obviously larger than that based on EM. On the other hand, the variance of the former is almost lower than that of the latter. In other words, the performance of TEM algorithm is much better and more robust than that of EM algorithm. On the whole, the detection results of TEM algorithm change little with the various *K*.

In Fig. 8, we also gave the average time for fitting the training set and the testing set with different *K*, respectively. This figure shows that the model fitting time of TEM algorithm is often longer than that of EM algorithm. This is mainly because TEM algorithm has to estimate the optimal temperature controlling parameter β and implements additional exponential operation in *E*-step. Therefore, to balance the performance and computational burden, *K* was set as 32 in the following experiment, where the number of shape latent aspect was 19 and the number of texture latent aspect was 13.

4.3 Detection performance analysis

In order to further evaluate the performance of the proposed algorithm, we compared it with two other existing approaches. The first is the traditional PCA+KNN approach in which principle component analysis is carried out on the extracted features and then the analysis results are recognized by KNN classifier (the number of neighborhood is 7). The second is the commonly used Feature+SVM approach in which extracted features are directly recognized by SVM classifier and then final results

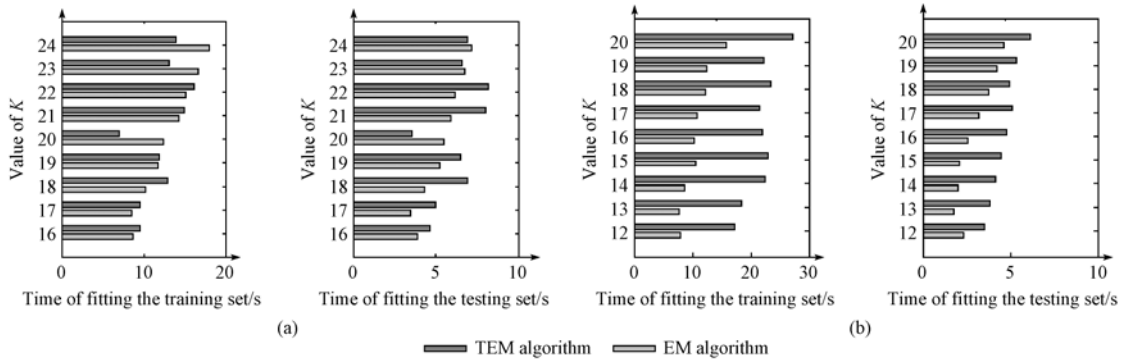


Fig. 8 Model fitting times with different *K* by using TEM and EM algorithms, respectively
(a) Shape feature; (b) Texture feature

are obtained. Meanwhile, we also test two different detection strategies: (1) binary detection, i.e., ship objects are labeled as one category and other false alarms are assigned into the same category in training process; (2) multiple classification, i.e., each kind of false alarms as well as ship objects are label as one category in training process. Table 1 and Table 2 list the detection results on different feature sets under strategy 1 and 2, respectively.

Table 1 Performance evaluation of different methods based on strategy 1

Feature category	Recall /%			Precision /%		
	PCA+KNN	Feature+SVM	Our method	PCA+KNN	Feature+SVM	Our method
Shape	79.39	79.15	**81.27**	55.27	**62.39**	60.52
Texture	86.52	90.10	**91.52**	60.09	67.46	**68.27**
Shape + Texture	88.23	**89.39**	88.51	60.28	70.08	**71.55**

Table 2 Performance evaluation of different methods based on strategy 2

Feature category	Recall /%			Precision /%		
	PCA+KNN	Feature+SVM	Our method	PCA+KNN	Feature+SVM	Our method
Shape	75.78	79.14	**82.38**	68.65	70.29	**71.24**
Texture	76.26	84.89	**87.41**	81.80	83.98	**84.97**
Shape + Texture	77.34	86.48	**88.77**	82.98	81.45	**83.03**

Totally speaking, strategy 1 can yield higher Recall but much lower Precision. All Precisions under strategy 1 are no larger than 70%. However, increasing the training samples of false alarms cause Recall decreasing quickly. The main reason is that the used detection algorithm could not efficiently model ship and false alarms under strategy 1 and then can not yield better Recall while keeping Precision higher. Therefore, we paid our attention to analyze the results under strategy 2. According to Table 2, the detection results on texture feature are better than which on shape feature for the reason that some false alarms, such as cloud and ocean wave, indeed have similar shape characteristics with those of ship objects. In other words, the phenomenon 2 described in the introduction often exists in shape feature. On the other hand, combining the shape and texture feature will improve Recall of each approach for the reason that the information provide by them is complementary. However, for the effect of the previous phenomenon, Precision of each approach decreases.

Table 2 also indicates that the proposed algorithm outperforms the other two approaches. Compared with PCA+KNN approach, our algorithm increases Recall with 7 points at least on different feature sets while producing higher Precision. Compared with Feature+SVM approach, our algorithm increases Precision with about 3 points while producing comparable Recall. Although the increase for Feature+SVM approach is not very large, the proposed algorithm still improve the whole performance of detection considering its less time of training and testing in model discriminating which is listed in Table 3. In addition, PCA+SVM belongs to a non-parameter approach and it does not need offline training, so Table 3 do not list the training time of this approach. However, the testing time of PCA+SVM is the longest of the three approaches and its performance is the worst.

5 CONCLUSION

This paper proposes a novel ship detection algorithm for optical remote sensing images based on model generating and discriminating. On the basis of quick detection of ship candidate region and effective extraction of various features, it first generates the description model of ship candidate region in terms of latent aspects by using the PLSA technique, and then constructs the discrimination model of ship candidate region by using SVM classifier. For one thing, the proposed algorithm can overcome the difficulty of representing object caused by the un-matching between the object and its extracted features. For another, it can relieve the complexity of the classification by SVM in the process of ship candidate region recognition, and decrease the time of training and testing for discrimination model. Meanwhile, the designed method of adaptive parameter

Table 3 Training and testing time for object recognition by using different methods

Feature category	Training time /s			Testing time /s		
	PCA+KNN	Feature+SVM	Our method	PCA+KNN	Feature+SVM	Our method
Shape	—	10.297	**7.266**	0.515	0.250	**0.203**
Texture	—	7.328	**4.875**	0.453	0.375	**0.324**
Shape + Texture	—	14.187	**7.437**	0.647	0.421	**0.391**

selection can obtain the optimal β of TEM algorithm so as to significantly improve the performance of model fitting. As the experiments demonstrated, the proposed algorithm can provide satisfactory ship detection results for optical remote sensing images in various complicate conditions.

Although more accurate, the proposed algorithm still face to address some problems, such as how to confirm the optimal number of latent aspect and how to quantify the extracted features to effectively compute feature frequency matrix. In addition, another method to improve the detection performance, we believed, is to introduce some proper prior constraints into the PLSA model.

REFERENCES

Chan F T and Vese A L. 2001. Active contours without edges. *IEEE Transactions on Image Processing*, **10**(2): 266—277

Chang C C and Lin C J. 2001. LIBSVM: A Library for Support Vector Machines. available at http://www.csie.ntu.edu.tw/~ cjlin/libsvm

Hofmann T. 2001. Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, **42**(2): 177—196

Li X W and Chong J S. 2007. A ship detectionMethod of SAR images based on wavelets and K-distribution. *Journal of Test and Measurement Technology*, **21**(4): 350—354

Ojala T, Pietikäinen M and Mäenpää T. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(7): 971—987

Otsu N. 1979. A threshold selection method from gray-level histogram. *IEEE Trans. on Systems, Man and Cybernetics*, **9**(1): 62—66

Quelhas P, Monay F, Odobez J M, Gatica Perez D and Tuytelaars T. 2007. A thousand words in a scene. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(9): 1575—1589

Tello M, Martínez C L and Mallorqui J J. 2005. A novel algorithm for ship detection in SAR imagery based on the wavelet transform. *IEEE Transactions on Geoscience and Remote Sensing Letters*, **2**(2): 201—205

Tian S R, Sun G Y, Wang C and Zhang H. 2007. A ship detection method in SAR image based on gravity enhancement. *Journal of Remote Sensing*, **11**(4): 452—459

Ueda N and Nakano R. 1988. Deterministic annealing EM algorithm. *Neural Networks*, **11**(2): 271—282

Vapnik V. 1998. Statistical Leaning Theory. New York, Wiley

Wang Y, Mei T, Gong S and Hua X S. 2009. Combining global, regional and contextual features for automatic image annotation. *Pattern Recognition*, **42**: 259—266

Wanielik G and Stock D J R. 1989. Use of radar polarimetric information in CFAR and classification algorithms. Proceeding of International Radar Conference

Yang W D, Zhang T X and Song C J. 2008. A novel algorithm for ship detection in low-resolution SAR images. *J. Huazhong Univ. of Sci. & Tech*. (*Natural Science Edition*), **36**(2): 78—81

Zhao Y H, Wu X Q, Wen L Y and Xu S S. 2008. Ship target detection scheme for optical remote sensing images. *Opto-Electronic Engineering*, **35**(8): 102—106

Zhou H, Wang R and Wang C. 2008. A novel extended local-binary-pattern operator for texture analysis. *Information Sciences*, **178**(22): 4314—4325

引入 PLSA 模型的光学遥感图像舰船检测

周 晖[1,2], 郭 军[1], 朱长仁[1], 王润生[1]

1. 国防科技大学 电子科学与工程学院 ATR 重点实验室, 湖南 长沙 410073;
2. 北京跟踪与通信技术研究所, 北京 100094

摘 要: 提出一种基于概率潜在语义分析(Probabilistic Latent Semantic Analysis, PLSA)的检测算法, 首先通过 PLSA 将目标表述为潜在成分的概率组合, 然后利用统计模式识别方法对获取的潜在成分概率进行判别, 从而完成最终的检测。其中, 生成的潜在成分反映了目标与特征之间相互出现的频率关系, 并以潜在成分在目标中概率差异的形式对上述不对应现象给出了直观描述。实验结果表明, 所提出的算法对多种复杂情况下的光学图像舰船检测具有很好的准确性和鲁棒性。

关键词: 舰船检测, 光学遥感图像, 概率潜在语义分析, 回火期望极大算法, 局部二进制模式算子

中图分类号: TP751.1 文献标识码: A

引用格式: 周 晖, 郭 军, 朱长仁, 王润生. 2010. 引入 PLSA 模型的光学遥感图像舰船检测. 遥感学报, **14**(4): 663—680
Zhou H, Guo J, Zhu C R and Wang R S. 2010. Ship detection from optical remote sensing images based on PLSA model. *Journal of Remote Sensing*. **14**(4): 663—680

1 引 言

舰船监视是各海岸地带国家的传统任务, 利用卫星图像对重点海域和重要港口进行监控, 可大大提高海防预警以及海运监测管理、调度的能力。合成孔径雷达(SAR) 作为一种主动式微波传感器, 具有全天时、全天候工作的特点, 成为目前海上舰船监视的主要手段, 许多 SAR 图像舰船检测算法(Wanielik & Stock, 1989; Tello 等, 2005; 李晓伟&种劲松, 2007; 田巳睿等, 2007; 杨卫东等, 2008)也随之提出。但是随着光学传感器的数量与日俱增, 且考虑到光学遥感图像具有覆盖幅度宽、成像分辨率高等特点, 基于光学图像的舰船检测可能更能满足实时监视的要求, 因此引起了学者们(赵英海等, 2008)的注意。

遥感图像中的信息通常会受季节、天气、传感器状态等多种因素影响, 复杂海面遥感图像更是存在大量与舰船特征属性类似的虚警干扰, 如海浪、云层、小岛等, 且舰船目标也因图像分辨率和成像机理不同而呈现出不一致的特征, 所以, 如何滤除这些虚警干扰一直是舰船检测的主要困难之一。传统上, 一般将其作为一个二元检测问题, 例如启发式的设置门限进行阈值化处理(杨卫东等, 2008; 赵英海等, 2008)、或者借助假定的概率分布估计舰船模型的参数后进行疑似目标验证(李晓伟&种劲松, 2007; 田巳睿等, 2007), 但这类方法由于建模方式比较单一, 无法充分刻画舰船目标和干扰对象之间的特征差异, 导致检测结果要么在低检测概率和高虚警概率之间徘徊, 要么检测类型受限、算法通用性不高。

我们认为解决该问题的一条途径是: 将舰船目标检测转化为基于统计学习的多元分类问题。利用支持向量机(Support Vector Machines, SVMs)(Vapnik, 1998)强大的学习性能, 通过融合多种有效特征来完成舰船目标的检测。然而, 在实际分类中, 通常存在很多“同类别异特性”和“同特性异类别”的现象。前者对应着属于相同模式类的样本具有不同特征属性的情况, 如图 1(a)所示的不同类型舰船样本; 后者则对应着具有相同特征属性的样本分属不同模式类的情况, 如图 1(b)所示的与舰船目标具有类似特

收稿日期: 2009-05-14; 修订日期: 2009-09-13
基金项目: 国防科技大学 ATR 重点实验室基金(编号: 9140C8004011007)。
第一作者简介: 周晖(1981—), 男, 国防科技大学电子科学与工程学院博士研究生, 主要研究方向为图像分析与理解、模式识别、遥感数据智能处理; 发表论文 8 篇。E-mail: hui_zhou_atr4@yahoo.com.cn。

征的朵云、海浪等虚警样本。尽管上述方法可以通过分析模式类中子模式的分布情况，给出“同类别异特性”识别问题的解决方法，然而，它将标注的模式类与提取的特征向量直接相关联的识别方式，对“同特性异类别”识别问题还无法有效解决。也就是说，当训练样本较少时，训练得到的 SVM 分类器无法逼近实际的目标模型，而当样本不断增加时，会让学习问题变得困难和不现实。

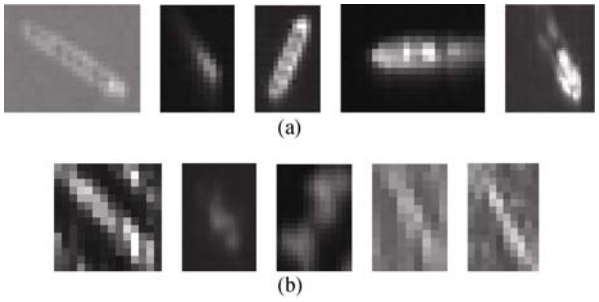


图 1　实际分类中的样本相互干扰示意图
(a)“同类别异特性”现象; (b)“同特性异类别”现象

鉴于上述分析，首次将概率潜在语义分析(Probabilistic Latent Semantic Analysis, PLSA)(Hofmann, 2001)引入到舰船目标的检测过程之中。在用于舰船目标建模时, PLSA 不是将提取的特征直接用于目标描述，而是通过学习“目标-特征”共现频率矩阵(以下简称“特征频率矩阵”)生成一个潜在成分(latent aspects)模型，利用该模型将目标表述为这些潜在成分的概率分解。这样，可以用不同目标中具有相应潜在成分的概率差异，直观描述“同类别异特性”或“同特性异类别”现象。利用 PLSA 进行目标建模，不仅简化了前期的标注工作，更降低了后期 SVM 识别模型在训练与判别工作中的难度。同时，考虑到 EM 算法在计算 PLSA 模型参数时存在的过拟合缺陷，本文使用回火期望极大算法(Tempered Expectation Maximization, TEM)计算参数，使得数据的 PLSA 模型拟合效果更优且稳健。测试结果表明，本文所提检测算法对多种复杂情况下的光学遥感图像舰船检测问题具有较强的实用性和鲁棒性。

2　PLSA 理论基础

PLSA 作为潜在语义分析(Latent Semantic Analysis, LSA)的变种，拥有更坚实的数学基础和易于利用的数据生成模型，其基本思想是根据词汇表中单词出现的频率发掘文本集中的主题。PLSA 模型常用于统计文本分析(Hofmann, 2001)和图像内容检索(Quelhas 等, 2007; Wang 等, 2009)，现已被证实能够为信息提取提供更好的词汇匹配。在本文中，文本集对应着待识别的目标，词汇表对应着目标的形状和纹理特征频率矩阵，而待发现的主题则对应着不同的目标类型，如舰船、海浪、朵云、岛屿等。

假定我们得到一组由特征集 $F=\{f_1, f_2, \cdots, f_M\}$ 描述的目标集 $O=\{o_1, o_2, \cdots, o_N\}$，并由此组成一个 $N\times M$ 的特征频率矩阵 $N=\left(n\left(o_i, f_j\right)\right)_{ij}$，$n\left(o_i, f_j\right)$ 表示特征 f_i 在目标 o_j 中出现的频率。此外，每一对可视数据 $\left(o_i, f_j\right)$ 与一组潜在成分(主题) $Z=\{z_1, z_2, \cdots, z_K\}$ 相关，K 为人工指定的一个常数。那么, PLSA 方法就是使用图 2 所示的图模型来刻画目标与对应特征及潜在成分之间的联合概率 $P\left(o_i, f_j, z_k\right)$，图中填充的节点表示可观测到的随机变量，未填充的则表示不可观测到的随机变量。而用于表述目标中特征内容的生成模型 $P(o_i,f_j)$可以通过式(1)计算得到：

$$\begin{aligned} P\left(o_i, f_j\right) &= P\left(o_i\right) P\left(f_j \middle| o_i\right) \\ P\left(f_j \middle| o_i\right) &= \sum_{k=1}^{K} P\left(f_j \middle| z_k\right) P\left(z_k \middle| o_j\right) \end{aligned} \tag{1}$$

式中，$P(o_i)$表示目标 o_i 的出现概率；$P\left(f_j \middle| z_k\right)$ 表示潜在成分 z_k 在特征 f_j 上的分布概率，也可以解释为特征 f_i 对潜在成分 z_k 的贡献度；$P\left(z_k \middle| o_i\right)$ 表示目标 o_i 中的潜在成分 z_k 分布概率，也可以解释为目标 o_i 中具有潜在成分 z_k 的概率。需要指出，该模型是建立在以下假设基础之上：o_i 和 f_j 之间是条件独立的，并且潜在成分 z_k 在目标和特征集上的分布也是条件独立。

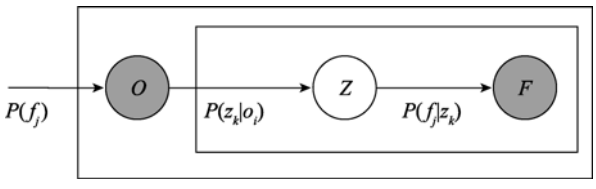


图 2　PLSA 的图模型表述

根据极大似然估计原则，通过求取式(2)定义的对数-似然函数的极大值，计算 PLSA 模型的参数：

$$L=\sum_{i=1}^{N}\sum_{j=1}^{M} n\left(o_i, f_j\right)\log P\left(o_i, f_j\right) \tag{2}$$

在 PLSA 模型中，极大似然估计的标准过程是

期望最大(Expectation Maximization, EM)算法。EM 算法交替于两个步骤：(1) *E*-步，利用当前估计的参数值计算潜在成分变量的后验概率; (2) *M*-步，基于所给后验概率，更新参数值。对于 PLSA 模型，应用贝叶斯公式即可得出 *E*-步如下：

$$P\left(z_k\left|o_i, f_j\right.\right)=\frac{\left[P\left(f_j\left|z_k\right.\right)P\left(z_k\left|o_i\right.\right)\right]}{\sum_{l=1}^{K}\left[P\left(f_j\left|z_l\right.\right)P\left(z_l\left|o_i\right.\right)\right]} \tag{3}$$

极大化式(2)，重新估计参数 $P\left(f_j\left|z_k\right.\right)$ 及 $P\left(z_k\left|o_i\right.\right)$，得 *M*-步：

$$P\left(f_j\left|z_k\right.\right)=\frac{\sum_{i=1}^{N}n\left(o_i, f_j\right)P\left(z_k\left|o_i, f_j\right.\right)}{\sum_{m=1}^{M}\sum_{i=1}^{N}n\left(o_i, f_m\right)P\left(z_k\left|o_i, f_m\right.\right)} \tag{4}$$

$$P\left(z_k\left|o_i\right.\right)=\frac{\sum_{j=1}^{M}n\left(o_i, f_j\right)P\left(z_k\left|o_i, f_j\right.\right)}{\sum_{j=1}^{M}n\left(o_i, f_j\right)} \tag{5}$$

在使用随机数初始化之后，交替实施 *E* 步骤和 *M* 步骤进行迭代计算直至式(6)定义的 *E*(*L*)达到收敛。

$$E\left(L\right)=\sum_{i=1}^{N}\sum_{j=1}^{M}n\left(o_i, f_j\right)\times \left(\sum_{k=1}^{K}P\left(z_k\left|o_i, f_j\right.\right)\log\left[P\left(f_j\left|z_k\right.\right)P\left(z_k\left|o_i\right.\right)\right]\right) \tag{6}$$

3 检测算法描述

本文所提舰船检测算法的工作流程可分为两部分，如图 3，其中起点带标记的箭头显示了检测的进行方向。 第一部分是“ 舰船候选区检测 ”：通过海陆分割确定检测区域，然后根据舰船目标在图像中一些特性的先验分布，比如灰度、边缘或者纹理，从检测区域中提取可疑目标；第二部分是“ 舰船候选区识别 ”：使用目标识别算法从得到的目标候选区域中确认舰船目标。

3.1 舰船候选区检测

在使用光学图像对舰船目标进行检测时，部分陆地区域的像元也会被检测为舰船目标，因此，在进行目标检测前需要海陆分割，剔除陆地区域的影响。本文使用 GIS 数据库中的海岸线信息和陆地区域信息确定陆地区域的范围，并从图像中将陆地区域剔除。当获得海洋检测区域后，根据舰船目标的边缘和形状信息，采用如下步骤进行舰船候选区检测：

步骤 **1** 图像的预处理 为了增强舰船目标与海洋区域在图像中的差异，本文采用一种新的混合

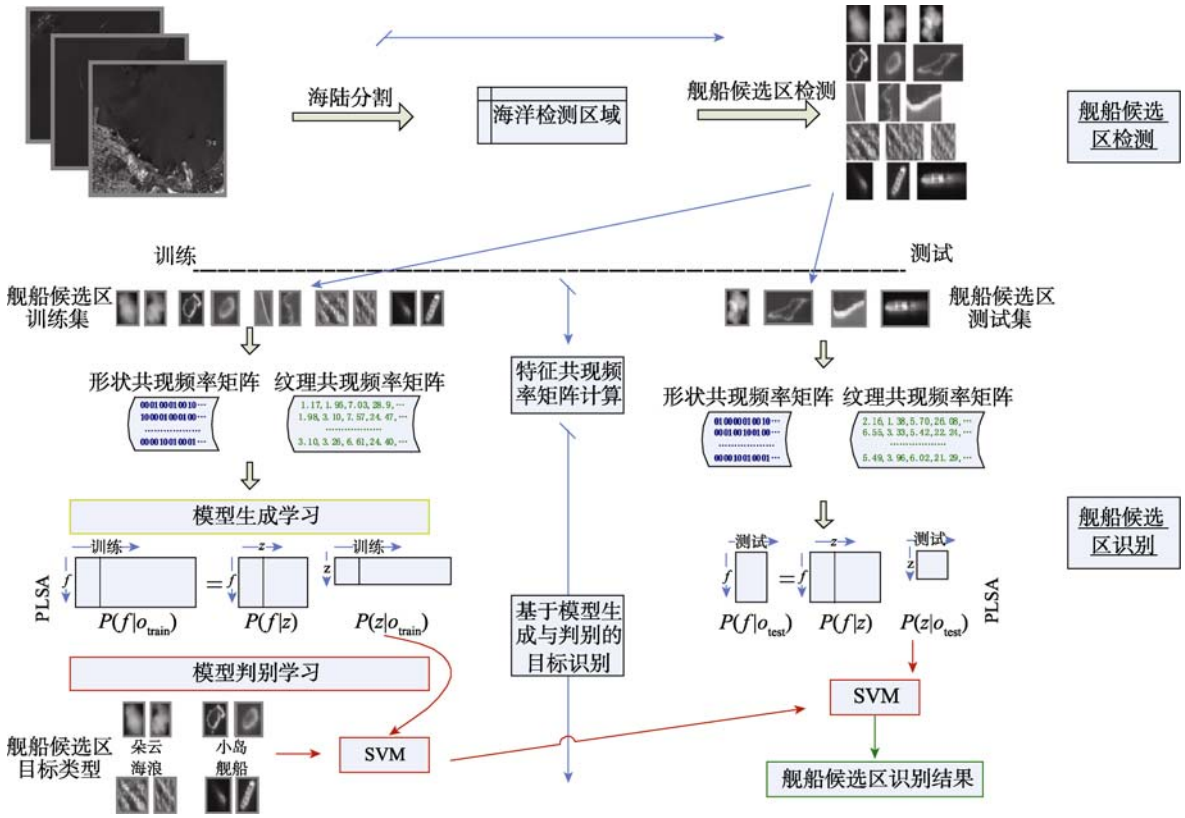


图 3 本文舰船检测算法总览

图像 $\mathrm{Mixu}(x,y)$ 计算方法，如式(7)。

$$\mathrm{Mixu}(x,y)=u(x,y)+\alpha\mathrm{Magu}(x,y) \tag{7}$$

式中，$u(x,y)$ 表示输入灰度图像中位于 (x,y) 的像元值；$\mathrm{Magu}(x,y)$ 则表示根据 Sobel 算子计算得到的该位置的梯度幅值。α 为幅值系数，本文设置为 1。

步骤 **2**　图像的粗分割　首先利用 Otsu 最大方差方法(Otsu，1979)自适应得到阈值，然后作用于混合图像 $\mathrm{Mixu}(x,y)$。针对这个分割结果利用形态学滤波滤除孤立和过于狭长的区域，然后根据区域最小外接矩形的长宽比 R_{region} 和区域大小 S_{region} 滤除过于奇异的检测结果，从而得到舰船候选区的初始分割。在本文中，R_{region} 的值取 1.5，而 S_{region} 的值与分辨率有关。

步骤 **3**　图像的精分割　在步骤 2 中，尽管采用了自适应的方法来设置阈值，但它还是一个基于全局性的统计，因此得到的候选区域边缘往往与真实的目标有所偏差。本文采用基于 Chan-Vese 模型的水平集方法来对边缘进行修正。Chan 和 Vese (2001)基于 Mumford-Shah 模型提出一种二类目标的最优分割算法，其简化的能量方程如下：

$$\begin{aligned}F(C)&=F_o(C)+F_b(C)=\\&\int_{\mathrm{inside}(C)}\left|u(x,y)-c_o\right|^2\mathrm{d}x\mathrm{d}y+\\&\int_{\mathrm{outside}(C)}\left|u(x,y)-c_b\right|^2\mathrm{d}x\mathrm{d}y\end{aligned} \tag{8}$$

其中 c_o 为曲线 C 内部区域的灰度均值，c_b 为曲线 C 外部区域的灰度均值。Chan-Vese 模型的能量最小值将对应着目标边缘的分段最优近似，也就是舰船候选区域边缘的修正结果。

图 4 给出了舰船候选区检测的中间结果，其中图 4(a)是输入的海域图像，图 4(b)是混合图像 Mixu(*x*,*y*)的计算结果，图 4(c)是舰船候选区的检测结果，图 4(d)和图 4(e)则分别对应着图 4(c)中两个目标的修正前后对比图。在图 4(d)和图 4(e)中，3 个切片依次是目标的原始图像、粗分割结果和精分割结果，可以看出通过水平集算法精炼后的分割结果更加逼近目标真实边缘。

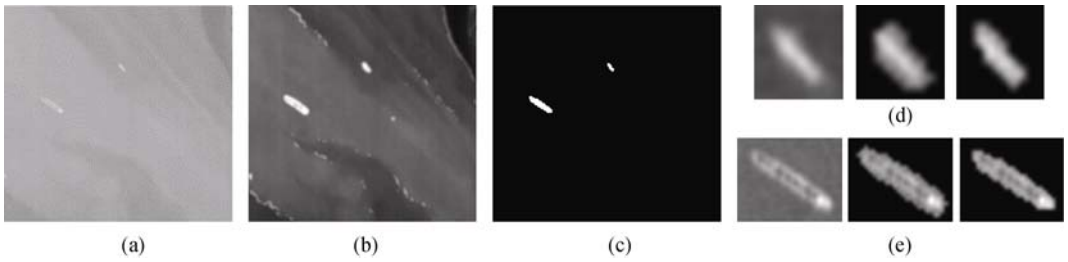


图 4　舰船候选区域的检测

3.2 舰船候选区识别

“舰船候选区检测”的目的是保证所有疑似目标被完整无误的定位，因此必然会产生很多虚警，所以在“舰船候选区识别”中，首先利用 PLSA 模型对候选目标进行建模，之后再利用 SVM 分类器对模型进行判别，从而剔除虚警完成最终的检测。考虑到 PLSA 模型是在一定的特征频率矩阵基础上生成的，所以在阐述具体的判别方法前，我们先说明一下该矩阵的计算方式。

3.2.1　特征频率矩阵计算

本文使用了形状和纹理这两大类特征用于特征频率矩阵的计算。其中，形状特征是基于图像精分割后的二值图像计算得到的，该特征主要反映了目标自身特性；而纹理特征是基于目标候选区域计算得到，除了目标特性之外，该特征还表述了一定的上下文信息。因此，二者结合可以有效提高检测的准确率。

形状频率矩阵

考虑到待检测图像往往来自于不同分辨率的遥感数据，我们选用如下具有旋转和尺度缩放不变性的形状特征：偏心率、矩形度，以及 Hu 矩中的前 3 个不变矩。同时，为了构造形状特征频率矩阵，将提取的形状特征 s 按式(9)所示最近邻原则量化为词集 V 中的某一个具体数值 v_i

$$\begin{aligned}&s\mapsto Q(s)=v_i\leftrightarrow\\&\quad\mathrm{d}(s,v_i)\leqslant\mathrm{d}(s,v_j),\forall j\in\{1,\cdots,N_v\}\end{aligned} \tag{9}$$

式中，d(*x*,*y*)为 x 和 y 之间的欧氏距离；N_v 为词集 V 的大小。具体来说，首先利用 K-means 算法对从训练集中提取的每一种形状特征单独进行聚类，以类心值作为该特征对应词集的 v_i 值；然后将特征按式(9)量化为相应词集中的一个 v_i；最后，根据计算得到的 v_i 组成形状特征频率矩阵 $H(s)$：

$$H(s)=\left(h^l(s)\right)_{l=1,\cdots L},\quad h^l(s)=\left(h_0^l(s),h_1^l(s),\cdots,h_{N_v^l}^l(s)\right) \tag{10}$$

其中，当 $Q(s)=v_i$ 时，$h_i(s)=1$，否则为 0；L 表示形状特征的种类数(对应于本文为 5)，N_v^l 表示第 l 种形状特征聚类后所得词集大小。

纹理频率矩阵

当前，人们提出了很多纹理描述子，局部二进制模式(Local-Binary-Pattern, LBP)纹理描述子(Ojala 等，2002)作为一种基于局部结构意义下的统计性纹理描述方法，取得比常规方法更好的纹理描述性能。但是，它将局部结构中的非一致局部模式简单地合并为一类，使得相当一部分结构信息没有得到有效利用，因此，本文选用先期工作中提出的一种扩展的 LBP 描述子(Zhou 等，2008)提取纹理特征，它根据相似性度量和结构性度量标准，将不同的非一致模式区别对待，从而进一步提高了 LBP 描述子的表述能力。限于篇幅，具体的定义式可参考文献(Ojala 等, 2002; Zhou 等, 2008)。同时，基于 LBP 描述子的纹理特征本身就是离散化的直方图，因此我们将该直方图归一化后即可作为纹理特征频率矩阵。

3.2.2 目标判别

进行目标判别时包含两个步骤：训练过程与预测过程，如图 3 下半部分所示。在训练时，一般涉及两个子步骤。其一，利用 EM 算法拟合舰船候选区训练集的 PLSA 模型，算法收敛后得到模型参数 $P(f_i|z_k)$ 和 $P(z_k|o_i^{\text{train}})$，然后每个训练样本可以用一个 K 维向量 $P(z|o_i^{\text{train}})$ 来表示。其二，根据得到的向量矩阵 $P(z|o^{\text{train}})$ 及其对应的标注结果训练 SVM 分类器。其中，SVM 分类器使用 LIBSVM 软件包(Chang and Lin, 2001)来实现, SVM 的模型参数通过交叉验证的方法训练得到。

在预测时，同样包含两个子步骤。其一，保持由训练集得到的 $P(f_i|z_k)$ 不变，利用 EM 算法拟合舰船候选区测试集的 PLSA 模型，从而得到测试样本集的 $P(z|o^{\text{test}})$。其二，使用 SVM 分类器对每一组 $P(z|o_i^{\text{test}})$ 进行预测得到最终的识别结果。

需要指出，为了避免标准 EM 算法所导致的过度逆合(overfitting)，本文使用回火期望极大算法(Tempered Expectation Maximization, TEM)(Ueda and Nakano, 1988)用于拟合模型。TEM 算法保持标准 EM 算法中的 M-步不变，而将 E-步变为：

$$P(z_k|o_i,f_j)=\frac{\left[P(f_j|z_k)P(z_k|o_i)\right]^{\beta}}{\sum_{l=1}^{K}\left[P(f_j|z_l)P(z_l|o_i)\right]^{\beta}} \tag{11}$$

其中，$\beta(0<\beta<1)$ 为温度控制参数。

然而，实际测试表明，并不是在任意一个 β 值下，TEM 算法可以获得比 EM 更好的 PLSA 模型拟合结果。为此，本文参考文献(Hofmann, 2001)，根据如下步骤选择最优 β 值：

(1) 令 $\beta=1$ 进行 EM 迭代，直到式(12)定义的性能指标“perplexity”不再降低;

(2) 减小 β，令 $\beta=\eta\beta$，其中，$\eta=0.95$；

(3) 在当前 β 值下进行 TEM 迭代，同样直到 perplexity 不再降低为止，之后进行步骤(2);

(4) 当 β 的减小无法继续降低性能指标 perplexity 时，最优 β 值选择结束。

上述迭代过程中，性能指标 perplexity 是模型推广能力的一个定量评价(Hofmann, 2001)，其定义为：

$$\text{perplexity}=\exp\left[-\frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\tilde{n}(o_i,f_j)P(o_j|f_j)}{\sum_{i=1}^{N}\sum_{j=1}^{M}\tilde{n}(o_i,f_j)}\right] \tag{12}$$

式中，$\tilde{n}(o_i,f_j)$ 表示辅助测试数据的特征频率矩阵。本文是从测试集中选择与训练集数量相当的样本来组成 $\tilde{n}(o_i,f_j)$。

4 实验结果分析

为证明本文所提舰船检测方法的有效性，本节利用该方法在大量光学遥感数据上进行了实验。

4.1 实验建立

在实验中选用 316 幅光学遥感图像，涉及的数据类型包括 CBERS, SPOT 4 等，它们的分辨率从 10m 到 5m 不等，其中一些典型的测试样本如图 5，包括云、岛屿、海浪等干扰情况。通过舰船候选区域检测共计得到 1653 个切片图像，经目视检查所有舰船目标均包含在候选区域的检测结果中。随机选取其中的 495 个切片图像组成训练样本集，剩余的 1158 个切片图像组成测试样本集。本文涉及实验均在 Intel Pentium4 2.66GHz 工作平台上进行，而提出的检测算法是在 VC++6.0 环境下开发完成的，其中，实现细节在第 3 章算法描述中进行了说明，主要参数的设置则在 4.2 节中给予阐述。

本文采用舰船目标检测的查全率(Recall)和查

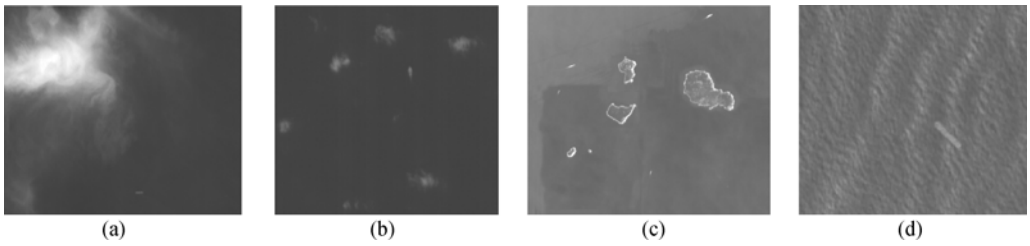
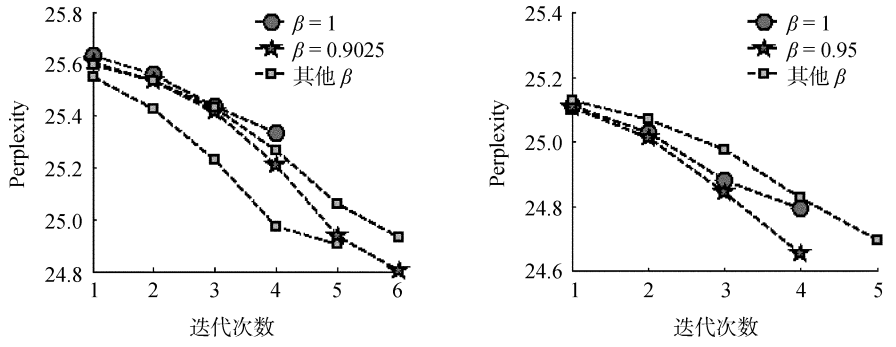


图 5　部分典型测试样本图像



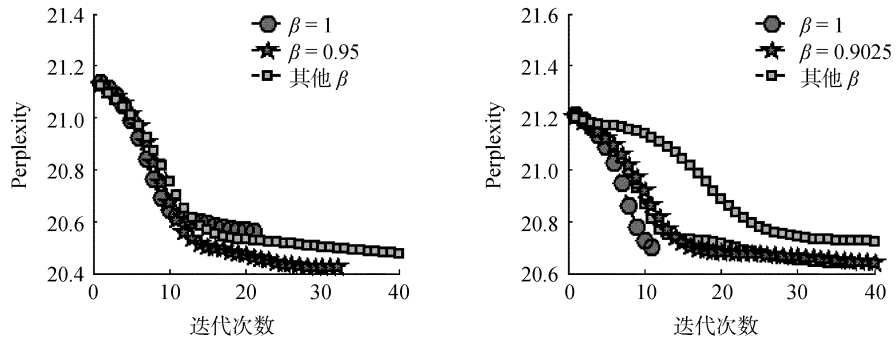


图 6　基于不同特征频率矩阵的 TEM 与 EM 的模型拟合结果

(a) 形状频率矩阵拟合结果; (b) 纹理频率矩阵拟合结果

准率(Precision)来度量不同方法性能的好坏。对于给定的测试图像集，其中手工标注的舰船目标数目为 $|W_G|$，使用一定检测方法得到的舰船目标数目为 $|W_M|$，其中 $|W_C|$ 个是正确的，则，

$$\text{Recall} = |W_C| / |W_G|, \text{Precision} = |W_C| / |W_M| \quad (13)$$

其中，查全率度量了对舰船目标检测的完整性，查准率则度量了舰船目标检测的精度。

4.2　算法参数分析

主要就算法中涉及的关键参数为温度控制参数 β和潜在成分个数 K 进行分析，通过实验观察不同参数设置条件下对识别率的影响，从而确定出最优参数设置方式。

4.2.1　β值确定

我们在此进行了 4 组实验来验证最优β值选择方法的可行性，其中形状和纹理各两组，如图 6(a)和(b)所示。各条曲线代表了不同 β 值下 perplexity 的收敛情况。为了清晰表达，我们在此只标出了 $\beta=1$(也就是标准 EM 算法的结果)和最优 β 值对应曲线，它们分别为圆形和五角星标出的曲线。根据实验结果可以发现，通过上述方法进行 TEM 模型拟合时，其 perplexity 指标明显低于标准 EM 算法所获得的值，也就是说 TEM 算法对模型拟合的效果更好。

4.2.2　K 值确定

在 PLSA 方法中，需要人工指定潜在成分 Z 的个数 K，同时文献(Quelhas 等, 2007; Wang 等, 2009)也指出 K 值的大小影响着算法性能的优劣。所以，我们需要分析不同数目的潜在成分对算法检测结果的影响。在此，将形状特征和纹理特征分别进行实验，并对每一个 K 值进行 20 组测试，然后统计同一 K 值下检测结果的均值和方差。同时，为了进一步说明

本文所采用的 TEM 算法的优势，给出 EM 算法的检测结果。实验结果如图 7，其中方框点和菱形点分别对应着 TEM 算法和 EM 算法在对应 K 值下的平均检测结果，每个点上两个叉号之间的线段则对应着该方法检测结果的两倍方差，虚线则表示了检测结果随 K 值变化的趋势。根据图 7，TEM 算法得到平均查全率和查准率要明显优于 EM 算法所得到的结果，而且前者检测结果的方差绝大部分也小于后者的。也就是说，TEM 算法的检测性能要优于 EM 算法且更加稳定。同时，根据整体变化趋势，K 值对 TEM 算法检测结果的影响也不是很大。

在图 8 中，我们还给出了两种拟合算法在 K 取不同值时分别用于拟合训练集和测试集的平均时间。该图说明，TEM 算法的模型拟合时间一般要大于 EM 算法的时间，这主要是由于在 TEM 算法中首先要估计最优的温度控制参数 β，另外在 E-步中的指数运算也导致运算量的提升。综合考虑检测性能与计算代价，在随后的对比实验中，将 K 值设为 32。其中，形状潜在成分的个数为 19 个，纹理潜在成分的个数为 13 个。

4.3 检测性能分析

为了进一步阐述本文所提算法的性能，我们将它与两种现有算法进行了比较。其中，第 1 种是传

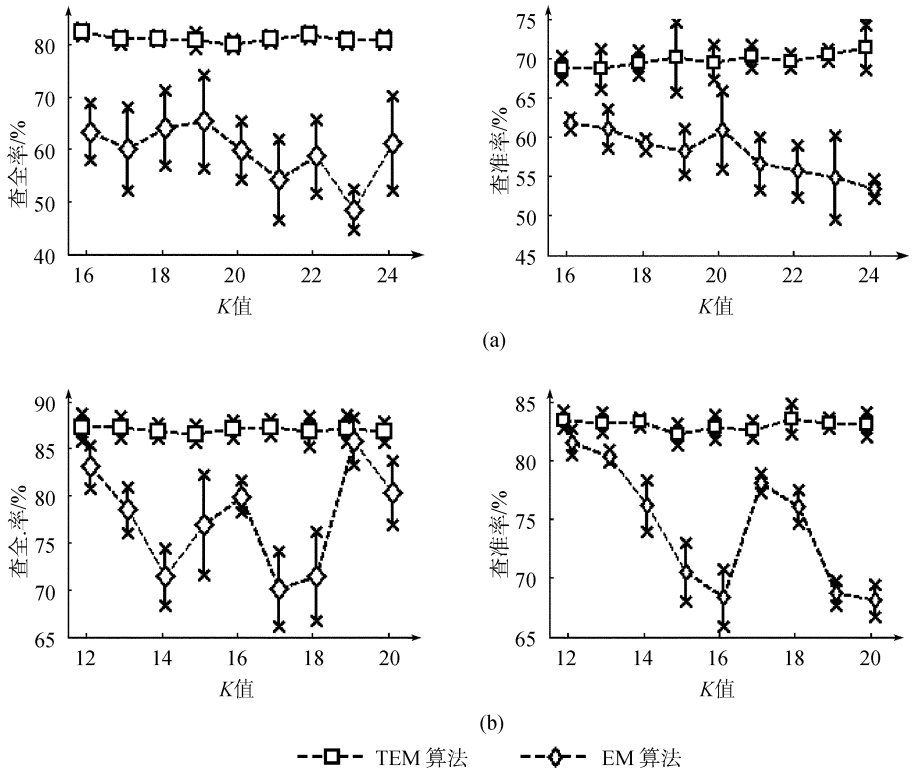


图 7 TEM 算法和 EM 算法检测结果随 K 值变化示意图

(a) 形状特征频率矩阵检测结果; (b) 纹理特征频率矩阵检测结果

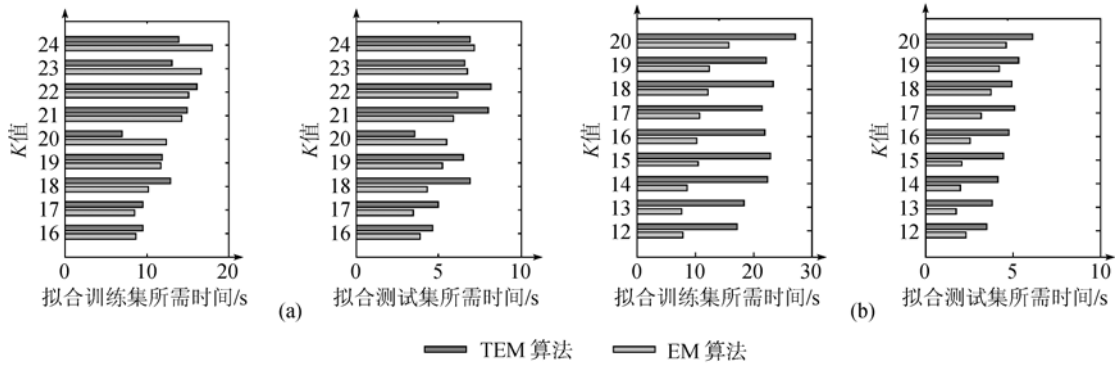


图 8 TEM 算法和 EM 算法用于模型拟合时间随 K 值变化示意图

(a) 形状特征频率矩阵拟合时间; (b) 纹理形状特征频率矩阵拟合时间

统的 PCA+KNN 方法，该方法首先对提取的特征做主成分分析，然后再利用 KNN 分类器对分析结果进行分类识别，其中最近邻的数量取值为 7；第 2 种是当前常用的 Feature+SVM 方法，该方法将提取的特征直接交给 SVM 进行识别从而获取最终检测结果。同时，我们还测试了两种不同策略的检测方法：第 1 种是二元检测，也就是在训练时只标注舰船目标，对于其他虚警干扰标注为非舰船目标；第 2 种是多元分类，也就是训练时除了标注舰船目标外，对其他虚警也按不同类型进行标注。表 1 和表 2 分别给出了策略 1 和策略 2 下 3 种方法在使用不同特征集时检测性能的评价结果。

总体上看，尽管策略 1 可以获得较高的查全率，但其查准率非常差，一般不超过 70%；若增加虚警的训练样本，则会导致查全率的迅速降低。我们认为，这主要是因为在策略 1 下检测算法无法对舰船与虚警目标进行有效建模，因此也就无法兼顾查全率与查准率。所以在后续讨论中，我们主要就策略 2 进行分析。根据表 2 看出，纹理特征相对于形状特征可以获得更好的检测结果，这是因为很多朵云和海浪样本与舰船目标的形状极为相似，也就是说在形状特征中，存在较多的“同特征异类别”现象；另一方面，组合使用形状特征和纹理特征可以有效提升检测的查全率，这是因为它们提供的信息互为补充，但也出于上述“同特征异类别”现象的影响，使得查准率不升反降。

表 2 还说明，本文所提算法要优于另外两种方法。其中，与 PCA+KNN 方法相比，本文方法的查全率在 3 个特征集上均提升超过 7 个百分点，同时，查准率也高于该方法。而与 Feature+SVM 方法相比，在保证更优查准率的情况下，本文方法的查全率可以提升 2—3 个百分点。尽管提升幅度不高，但考虑到 SVM 方法突出的优越性，以及在模型判别时用于训练和测试时间均明显少于该方法(表 3)，本文方法从整体性能上还是带来了相当大的提升。另外，由于 PCA+KNN 属于非参数化方法，基本不需离线训练，所以在表 3 中未列出训练时间。但它用于测试的时间最长，且检测性能在三者之中也是最差的。

5　结　论

本文提出了一种基于模型生成与判别的光学遥感图像舰船检测算法，该算法以快速的舰船候选区

表 1　策略 **1** 下不同方法的检测性能评价

特征类型	查全率/%			查准率/%		
	PCA+KNN	Feature+SVM	本文方法	PCA+KNN	Feature+SVM	本文方法
形状	79.39	79.15	**81.27**	55.27	**62.39**	60.52
纹理	86.52	90.10	**91.52**	60.09	67.46	**68.27**
形状 + 纹理	88.23	**89.39**	88.51	60.28	70.08	**71.55**

表 2　策略 **2** 下不同方法的检测性能评价

特征类型	查全率 /%			查准率 /%		
	PCA+KNN	Feature+SVM	本文方法	PCA+KNN	Feature+SVM	本文方法
形状	75.78	79.14	**82.38**	68.65	70.29	**71.24**
纹理	76.26	84.89	**87.41**	81.80	83.98	**84.97**
形状 + 纹理	77.34	86.48	**88.77**	82.98	81.45	**83.03**

表 3　不同方法用于目标判别的训练和测试时间

特征类型	训练时间 /s			测试时间 /s		
	PCA+KNN	Feature+SVM	本文方法	PCA+KNN	Feature+SVM	本文方法
形状	-	10.297	**7.266**	0.515	0.250	**0.203**
纹理	-	7.328	**4.875**	0.453	0.375	**0.324**
形状 + 纹理	-	14.187	**7.437**	0.647	0.421	**0.391**

域检测以及多种有效的描述特征为基础，首先利用 PLSA 方法生成舰船候选区域的潜在成分描述模型，然后利用 SVM 构建舰船候选区域的判别模型。基于上述策略的检测方法，一方面能够克服目标与特征之间不对应现象所造成的目标描述困难，另一方面，能够降低 SVM 在进行舰船候选区域确认过程中的识别难度，缩短了判别模型的训练和预测时间。同时，本文使用的自适应参数选择方法可以得到 TEM 算法的最优温度控制参数，从而大大提升模型拟合的性能及稳健程度。实验结果表明，针对多种复杂情况下的光学图像舰船检测问题，本文所提算法均能获得令人满意的检测结果。

当然，在检测过程中算法还存在不少问题，如何更好的确定最优潜在成分个数，以及如何进行特征量化以有效的统计特征频率矩阵。另外，如何将合理的先验约束引入到 PLSA 模型中，也是下一步继续提高检测性能的可行途径。

REFERENCES

Chan F T and Vese A L. 2001. Active contours without edges. *IEEE Transactions on Image Processing*, **10**(2): 266—277

Chang C C and Lin C J. 2001. LIBSVM: A Library for Support Vector Machines. available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

Hofmann T. 2001. Unsupervised learning by probabilistic latent semantic analysis. *Machine Learning*, **42**(2): 177—196

Li X W and Chong J S. 2007. A ship detectionMethod of SAR images based on wavelets and K-distribution. *Journal of Test and Measurement Technology*, **21**(4): 350—354

Ojala T, Pietikäinen M and Mäenpää T. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(7): 971—987

Otsu N. 1979. A threshold selection method from gray-level histogram. *IEEE Trans. on Systems, Man and Cybernetics*, **9**(1): 62—66

Quelhas P, Monay F, Odobez J M, Gatica Perez D and Tuytelaars T. 2007. A thousand words in a scene. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**(9): 1575—1589

Tello M, Martínez C L and Mallorqui J J. 2005. A novel algorithm for ship detection in SAR imagery based on the wavelet transform. *IEEE Transactions on Geoscience and Remote Sensing Letters*, **2**(2): 201—205

Tian S R, Sun G Y, Wang C and Zhang H. 2007. A ship detection method in SAR image based on gravity enhancement. *Journal of Remote Sensing*, **11**(4): 452—459

Ueda N and Nakano R. 1988. Deterministic annealing EM algorithm. *Neural Networks*, **11**(2): 271—282

Vapnik V. 1998. Statistical Leaning Theory. New York, Wiley

Wang Y, Mei T, Gong S and Hua X S. 2009. Combining global, regional and contextual features for automatic image annotation. *Pattern Recognition*, **42**: 259—266

Wanielik G and Stock D J R. 1989. Use of radar polarimetric information in CFAR and classification algorithms. Proceeding of International Radar Conference

Yang W D, Zhang T X and Song C J. 2008. A novel algorithm for ship detection in low-resolution SAR images. *J. Huazhong Univ. of Sci. & Tech.* (*Natural Science Edition*), **36**(2): 78—81

Zhao Y H, Wu X Q, Wen L Y and Xu S S. 2008. Ship target detection scheme for optical remote sensing images. *Opto-Electronic Engineering*, **35**(8): 102—106

Zhou H, Wang R and Wang C. 2008. A novel extended local-binary-pattern operator for texture analysis. *Information Sciences*, **178**(22): 4314—4325

附中文参考文献

李晓玮，种劲松. 2007. 基于小波分解的 K-分布 SAR 图像舰船检测.测试技术学报, **21**(4)：350—354

田巳睿，孙根云，王超，张红. 2007. 基于引力场增强的 SAR 图像舰船检测方法研究.遥感学报, **11**(4)：452—459

杨卫东，张天序，宋成军. 2008. 低分辨率 SAR 图像舰船目标检测.华中科技大学学报(自然科学版), **36**(2)：78—81

赵英海，吴秀清，闻凌云，徐守时. 2008. 可见光遥感图像中舰船目标检测方法.光电工程, **35**(8)：102—106