2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Fast Ship Detection of Synthetic Aperture Radar Images via Multi-view Features and Clustering

Shigang Wang, Shuyuan Yang, *IEEE member*, Zhixi Feng and Licheng Jiao
Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education,
International Research Center for Intelligent Perception and Computation, Xidian University,
Xi'an, Shaanxi Province 710071, China
sgwang@stu.xidian.edu.cn

***Abstract*—This paper proposes a novel ship detection scheme in coastal regions for high-resolution synthetic aperture radar (SAR) imagery based on prior knowledge of the different properties presented by target and clutter. To begin with, image segmentation and land masking are applied to eliminate the areas that are unlikely to contain targets and get the index image which indicates the likely target positions. Ship detection is conducted only on these likely target positions using power ring algorithm (PR), which can avoid unnecessary and exhaustive searches. In the discrimination stage, two new features named number of 8 connected regions and average power of target areas are proposed and used to form a discriminative feature group. Unlike most discriminators, which are based on supervised learning, we use an unsupervised method based on K-means clustering to deal with the situations where there are few or no labeled samples. Experimental results show that the proposed scheme is fast in speed and can detect most of the targets while few false alarms occur.**

Keywords—synthetic aperture radar (SAR); land masking; ship detection; K-Means clustering;

I. INTRODUCTION

Synthetic aperture radar (SAR) images have been collected by diverse platforms for various applications till now. The number of such imagery is growing rapidly, and along with that growth is the expanding need for computer-aided or automated exploitation of SAR imagery. One of the most important applications is automatic target recognition (ATR), which aims to find target like regions and attach a class to each region. Generally, the SAR ATR processing is split into three successive stages: prescreening stage, discrimination stage and classification stage as is shown in Fig.1 [1], [15]. The first two stages are commonly known as the focus of attention module, which has high data load and low computational complexity. This paper mainly deals with the focus of attention module and tries to find an efficient way to solve this problem.

The function of prescreening stage is to search through the entire scene for the regions of interest, based on the fact that the radar cross section (RCS) of a target region is typically higher than the surrounding background regions due to corner reflection. The commonly used detection methods are based on the contrast between target pixels and the surrounding areas. Among these methods, constant false alarm rate (CFAR) technique is widely used, especially in some real time

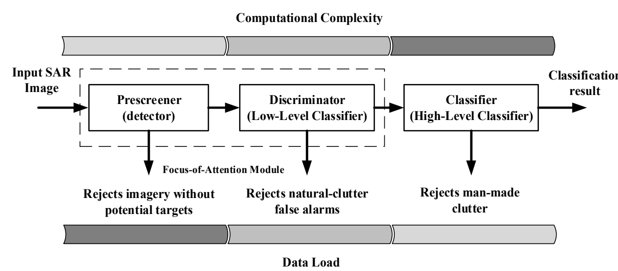


Fig.1. General structure of the SAR ATR system

application systems like SAIP [2]. Various CFAR algorithms have been proposed with regard to statistical modeling of background as well as detector design in order to cope with complex circumstances. A suitable model is of great importance to the performance of detection. However, it is always not easy to make a compromise between the fitness of model and its corresponding cost, which in some sense limits the further development of CFAR. Besides, non-statistical methods like generalized likelihood ratio test (GLRT) and power ring (PR) are feasible [3]. GLRT is based on the distributions of both target and background to get a theoretical Bayesian optimal solution. In real applications, statistical modeling of target is unknown, which means GLRT is not a practical way. PR is a simple but effective detector if the parameters are chosen properly. In this paper, we empirically discuss how to set the suitable parameters for PR to get an acceptable result. Besides, algorithms that use other image features, such as extended fractal (EF), multi-resolution and sub-aperture coherent characteristics, are also widely investigated [4]-[7]. Calculating these features generally costs a lot of time, thus they can't meet the real time requirement.

After the detection stage, most of the false alarms are eliminated and ship candidates are obtained. However, a large number of target like false alarms still exist. To further reduce these false alarms, we need to explore features and classifiers to distinguish targets from clutter. Actually, this is the function of the discrimination stage, which intends to accept target candidates and reject the others. Features are crucial as they provide information from different perspectives that can be integrated to give a judgment of whether the candidate is a

This work was supported by the National Basic Research Program of China (973 Program) under Grant no. 2013CB329402.

target or not. However, not too much effort has been devoted to develop effective and robust features [6]. Current features can be generally divided into four categories: textural features, size features, contrast features and polarimetric features [8]. Among these features, Lincoln Laboratory Discrimination Features and ERIM Discrimination Features are the most widely used [9]. The three texture based features developed by Lincoln Laboratory are extracted from the original image and reflect the intensity fluctuation, spatial dimensionality and percentage of power contained in the brightest scatters respectively. ERIM Discrimination Features are extracted from both the binary and CFAR image in a target-shaped blob to reflect the size and contrast properties. Each of these features can represent a certain characteristic of ship candidates and only by combing some of them to form a more discriminative feature group can we hope to obtain an acceptable outcome. Moreover, two features though computed in different ways, may be relevant to each other, namely they describe the similar property. Thus, a good feature group is the one whose elements are more likely to be independent of each other and have complete description over the candidates. In this paper, we develop two new features: number of 8 connected regions and average power of target areas based both on the binary and original images, which are proved by experiments to be quite discriminative.

By far, most of the effort has been devoted to developing effective discriminators based on supervised learning methods such as QD, SVM, BNN, QPD and KNN [9]-[11], which usually include two steps: training and testing. If enough representative training data is available, these methods will usually have satisfying performance. While, according to [12], it's nearly an exhausting work to get the training data in real world applications. Even though we build a target database after great endeavor, we may still not get the expected results due to the reason that the training and testing data are collected under different background, by different sensors, in different weather conditions. What's more, target itself is also a large category where different types of targets are included. A commonly used strategy is to divide the candidate chips (image patches we extracted after the detection stage) into two groups and choose one as the training set and the other as the testing set to verify the performance of features as well as classifiers [6]. But it is limited to theoretical argument and not enough data can be used to learn the classifier well. Inspired by this situation, we try to solve the discrimination problem in another way. After visualizing the samples collected from SAR images in the proposed description manner, we find that targets will group together and clutters will scatter but far away from targets in the feature space. It is more suitable to use a clustering method as K-means to divide the candidates into target and clutter groups. Thus, we can use the rules inside the candidates instead of learning it from other candidates to distinguish between target and clutter. The diagram of the proposed scheme is shown in Fig.2.

The rest of this paper is organized as follows. Section II introduces our PR based ship detection method. The newly developed feature group and unsupervised discrimination strategy is illustrated in Section III. Experimental results are provided in Section IV and finally conclusions are given in Section V.

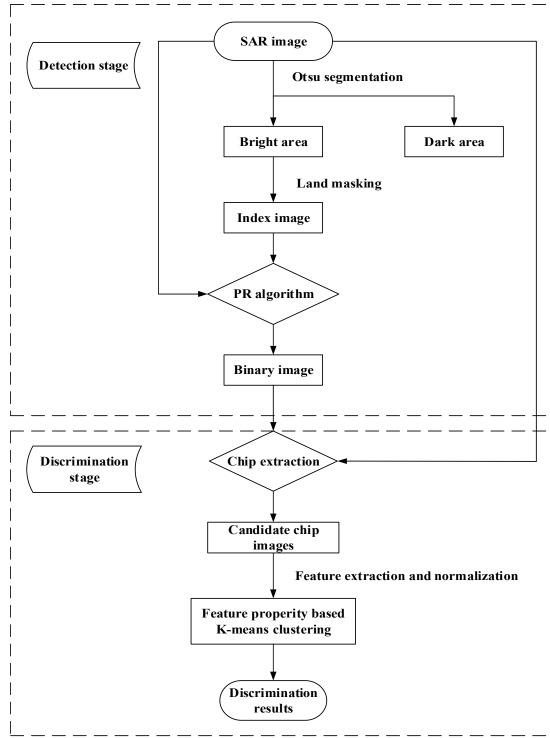


Fig.2. Diagram of the proposed ship detection scheme

II. CANDIDATE DETECTION

In this section, preprocessing based on Otsu segmentation is first introduced in Section II-A to locate the potential target areas. Then, in Section II-B, PR is discussed to form a suitable detector. Ship candidate chip extraction is finally described in Section II-C.

A. Preprocessing

For SAR images from coastal regions, there exist ocean, land, buildings and ships in the scene. Usually, the image can be roughly considered to be composed of two parts based on contrast: bright area and dark area. Due to the different scatter properties, targets will be in the bright area and the ocean in the dark area. Land and buildings also belong to the bright area but they have much larger areas than that of targets.

We use Otsu algorithm [13], [14] to find the optimal threshold so as to divide the whole scene into bright and dark areas. Otsu is based on statistical decision theory and tries to find the maximum between-class variance by choosing a proper threshold. The between-class variance is given by [13]

$$\sigma_B^2(k) = \frac{[m_G P_1(k) - m(k)]^2}{P_1(k)[1 - P_1(k)]}, \tag{1}$$

where, k is the threshold variable, $P_1(k)$ is the CDF of the image, $m(k)$ is the value of the *k-th* level and m_G is the mean value of image. The optimal threshold is obtained by [13]

$$\sigma_B^2(k^*) = \max \sigma_B^2(k). \tag{2}$$

The original image and image after segmentation are shown in Figs.3(a) and (b). For this image, the optimal threshold computed by Otsu is 60. We apply this threshold to divide the whole scene into bright and dark areas. As expected before, ships are included in the bright area due to the strong echoes.

After segmentation, we mainly focus our attention on the bright area where the targets lie in. It's necessary and possible to further investigate into the bright area to eliminate the obvious non-target areas. Each target area is limited to a certain size while land and buildings are often large connected areas. Based on this, morphological operation is applied to fill the holes in the binary image. Then we remove the regions whose areas are much larger than that of the target to obtain the index image. In fact, the function of the index image is to indicate whether the corresponding pixel under test is a potential target pixel or not. The following detection stage is conducted only to the potential target pixels according to the index image which can help avoiding unnecessary and meaningless search. It's predictable that this preprocessing step can reduce the data size a great deal and the speed will surely be improved.

B. *Ship Candidate Detector*

Contrast based detection methods usually adopt a sliding window to make decision as is shown in Fig.4. The test cell is at the center of the defined local region and the cells in the boundary stencil are used to estimate the characteristics of the local clutter such as mean, standard deviation and probability density function [13]. The guard area ensures that no target cells are included in the estimation of clutter characteristics. For PR algorithm we used here, we choose a target support area to estimate the mean of the region of interest (ROI) and a clutter support area to estimate the mean of local background. The decision is made according to the following formula [3]:

$$\begin{aligned} \frac{\mu_{ROI}}{\mu_C} &\geq \lambda \Rightarrow \text{target pixel} \\ \frac{\mu_{ROI}}{\mu_C} &< \lambda \Rightarrow \text{nontarget pixel}, \end{aligned} \tag{3}$$

where μ_{ROI} is the average power in ROI around the test pixel, μ_C is the average power of surrounding clutter and λ is chosen empirically so as to get a satisfying detection result. Here, we set the number of target support cells to be 9, the number of clutter support cells to be 400 and λ to be 3.

Before detection, image pad is first performed such that locations near the image borders can also be examined. The sli-

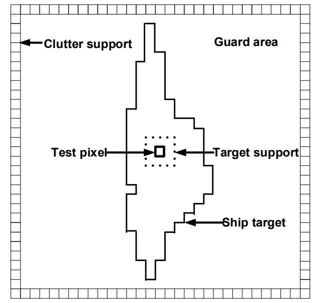


Fig.4. Sliding window used for the power ratio (PR) algorithm

ding window then slips through all the potential target pixels according to the index image to search for the ship candidates. The index image and binary image after detection are shown in Figs.5(a) and (b) respectively. As we can see from the result that there is no significant missing detection in our proposed method and the number of false alarms is acceptable. We will discuss this further in Section IV.

C. *Ship Candidate Chip Extraction*

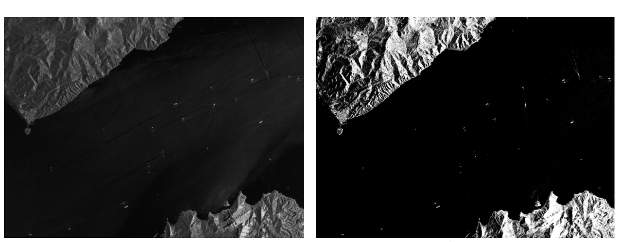
After detection stage, we get a binary image and the ship candidate chips are extracted based on this image. First, three morphological operations: hole filling, erosion and dilation are used to regulate the shape of the detected regions [13], [14]. After that, too big or too small regions are removed in order to eliminate the influence of nontarget chips. Finally, ship candidate chips are cut from both the original and binary images according to each detected region in the binary image. Note that we will combine the information from original chips with that from binary chips in the discrimination stage to get a better description of the extracted candidate chips. Based on the prior knowledge of the image resolution as well as the general scale of the ships, we set the size of the chip to be 64× 64. Fig.6 shows some of the extracted chips containing both targets and clutters.

III. Ship Discrimination

In this section, the three features used for discrimination are firstly described and analyzed in Section III-A. Section III-B develops a new discriminator based on K-Means clustering method.

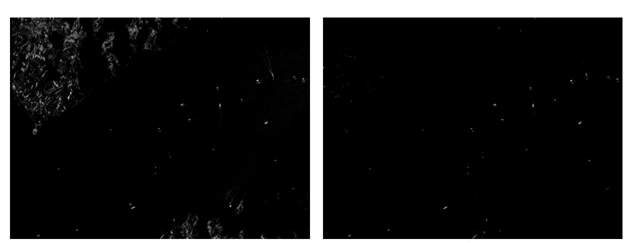
A. *Discrimination Features*

The three features we use here are log standard deviation (proposed by Lincoln Laboratory), number of 8 connected regions and average power of target areas (proposed by us). Log standard deviation provides information about the fluctua-



(a) (b)

Fig.3. (a) Original SAR image. (b) Image after Otsu segmentation.



(a) (b)

Fig.5. (a) Index image. (b) Binary image after detection with PR.

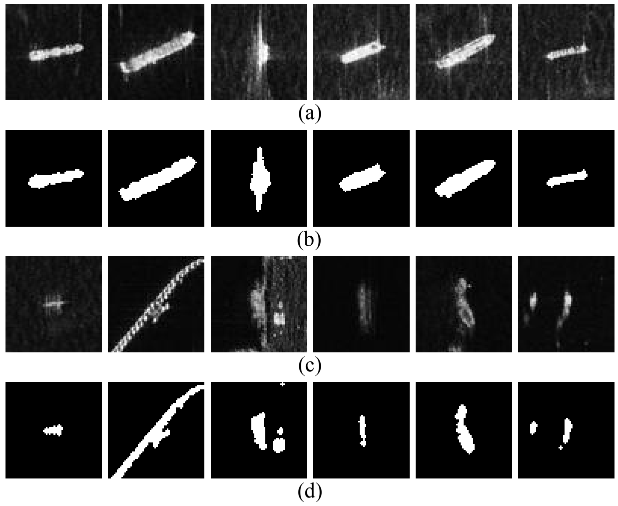


Fig.6. Extracted ship candidate chips. (a) Six target chips extracted from the original image. (b) The binary chips corresponding to (a). (c) Six clutter chips extracted from the original image. (d) The binary chips corresponding to (c).

tions in intensity across the region. Number of 8 connected regions is a measure of spatial distribution of the brightest scatters in the region. Average power of target areas reflects the average intensity of the target blob. The three features are extracted from the original image, binary image and both images respectively, which provide information from texture, region and contrast.

The log standard deviation of a region is defined as the standard deviation of the radar returns (in dB) from the region. If the radar intensity from range r and azimuth a is denoted by $P(r,a)$, then the log standard deviation σ can be estimated as follows [16]:

$$\sigma = \sqrt{(S_2 - S_1^2 / N) / (N-1)}, \tag{4}$$

$$S_1 = \sum 10\log_{10} P(r,a) \quad (r,a) \in \text{region}, \tag{5}$$

$$S_2 = \sum [10\log_{10} P(r,a)]^2 \quad (r,a) \in \text{region}, \tag{6}$$

where, N is the number of units in the region. Usually, there are large fluctuations in the target regions and as a result, the value of standard deviation for target region is much larger than non-target region. Number of 8 connected regions is the number of independent or unconnected regions in the binary chips. We can determine this feature by using morphological methods. The brightest scatters in target regions often gather together and around the centroid of target blob, while for nontarget regions, this is just the opposite. Therefore, target regions will have smaller number of 8 connected regions than nontarget regions. For example, the first region in Fig.4(b) has one 8 connected regions while the third region in Fig.4(d) has four. Average power of target areas is calculated from original chips according to the blobs of binary chips. The blob of the binary image can be viewed as an index used to point out the pixel positions of the target. We use the pixel values of the target positions from the original image to calculate this feature by:

$$Apt = (1/N)\sum P(r,a) \quad (r,a) \in \text{target blob}, \tag{7}$$

where, Apt represents the average power of target areas and N is the number of target pixels. This feature is discriminative due to the reason that it is robust to different size of targets and can give a contrast based measurement. For target chip, Apt is much larger than that of clutter chip.

The other two features developed by Lincoln Laboratory are fractal dimension and weighted-rank fill ratio [16], [17]. They are effective in most of the cases but not always. The two new features described above are in fact the improved version of the fractal dimension and weighted-rank fill ratio. As can be seen later, our new features are more discriminative and robust than the classical version.

B. *K-means Discriminator*

As we have stated earlier, current discriminators are based on supervised learning method, which require a large amount of samples and usually a lot of time to train a classifier. Is it possible to find another way to solve this problem? By reviewing this problem, we may find that what we do is to tell whether the chip under test contains a target or not regardless of what kind of information we use or what kind of method we adopt. Current discriminators train some labeled data to form decision rules and use the rules to make decisions for unlabeled data. After observing the feature space formed by the three features, we may find that targets will group together and clutters will scatter but far away from targets as can be seen in Fig.7. If we normalize the three feature values in interval [0,1], the ideal situation may be that target points will gather around position [1,0,1] and clutter points will gather around position [0,1,0]. This leads to the use of K-means for discrimination between target and clutter.

The K-means discriminator makes use of both the target and clutter feature information to guide its clustering process. To eliminate modality-dependent amplitude differences, we normalize the features into interval [0,1]. After normalization, target group will gather around [1,0,1] and clutter group around [0,1,0] in the feature space. The number of clusters is set to be 2 and their initial centers are predefined as [1,0,1] and [0,1,0] respectively. The motivation behind this lies in the fact that we hope to assign the target points to the first cluster and clutter points to the second one by prior information based clustering process. After a small number of iterations, this algorithm will converge to a stable state, which is not time consuming [18].

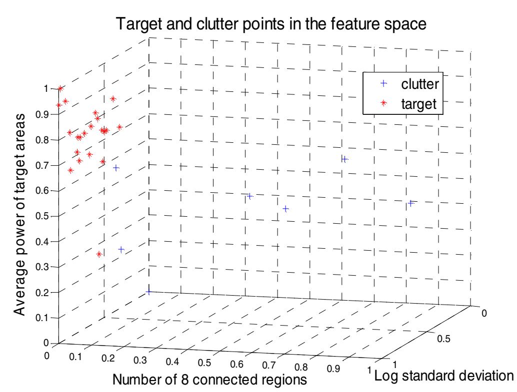


Fig.7. Distribution of target and clutter points in the feature space

After chip extraction, 28 candidate chips pass to the discrimination stage, including 21 ship targets and 7 clutters. Through the above described discrimination stage, the ship targets are well kept and one false alarm is misjudged to be target. The result before and after discrimination are shown in Figs.8(a) and (b) respectively.

IV. EXPERIMENTAL RESULTS

The proposed approach is verified using an X-band TerraSAR data collected near the strait of Gibraltar. The polarization mode is HH and the azimuth and range resolutions are 1m respectively. The data covers a scene of about 4134m ×2987m, which corresponds to an image size of 4134×2987. The experimental results after detection and discrimination stage will be presented and analyzed in the following.

Large amount of redundant information exists in the original image, which is a main source of computational burden. Preprocessing stage aims at eliminating the nontarget points and thus to alleviate the data load for detection stage. After Otsu segmentation and land masking, we remove dark area and land area. Table I shows the number of pixels that are considered to be target pixels in each step.

In coastal regions, there are different types of clutter and it's difficult to use a model to describe the clutter distribution properly. Non-model based methods as PR will outperform model based methods as 2-para CFAR [1], [9]. Fig.9 shows the detection results by using PR and 2-para CFAR respectively. Table II shows the computation time and number of candidates for each detector after detection.

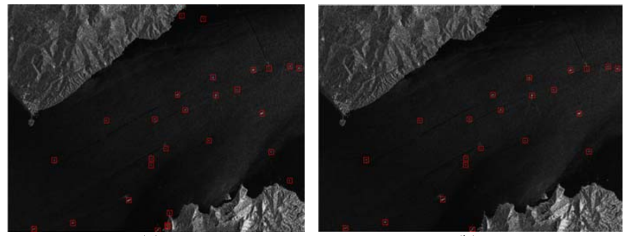
To verify the validity of the proposed features, we display the feature values of the 28 candidates in Fig.10. Due to the similarity between target and clutter as well as the diversity existing within targets, each feature is discriminative to certain candidates and only by choosing a proper feature group can we hope to achieve a satisfactory performance. The three features proposed by Lincoln Laboratory will be used to form a feature group and the two new features along with log standard deviation will be used to form another feature group. We use the above proposed discriminator to verify the ability of the two feature groups. The discrimination results are shown in Fig.11. As can be seen from the result, one target is lost in three Lincoln Laboratory features while the new feature group finds all the targets. Besides, one false alarm occurs in both two feature groups.

TABLE I. NUMBER OF PIXELS IN EACH STEP

Image	Original image	Otsu image	Index image
Pixel number	12348258	1372412	176917

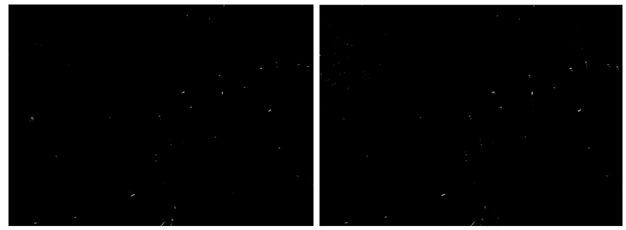
TABLE II. COMPUTATION TIME AND NUMBER OF CANDIDATES

	Computation time (s)	Number of candidates
PR	40.6518	200
2-para CFAR	55.6705	608



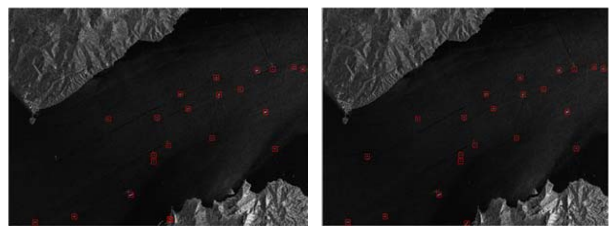
(a) (b)

Fig.8. (a) Candidates before discrimination. (b) Discrimination result.



(a) (b)

Fig.9. (a) Detection result with PR. (b) Detection result with 2-para CFAR.



(a) (b)

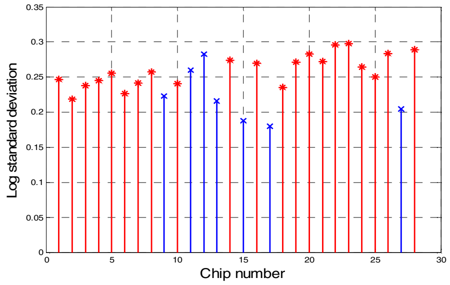
Fig.11. Discrimination results. (a) Lincoln Laboratory feature group. (b) New feature group.

V. CONCLUSION

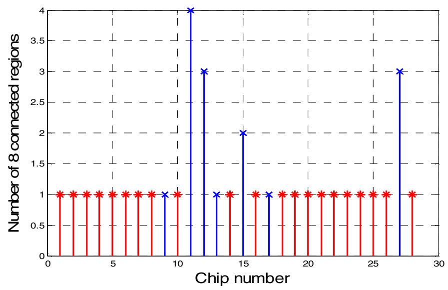
For ship detection in coastal regions of high resolution SAR imagery, a new preprocessing method based on Otsu is first applied. Ship targets are usually small bright regions scattering in the whole scenario and searching the large areas of land and ocean for targets is exhausting. By focusing only on the potential target area, we filter out a lot of false alarms while targets areas are well preserved. Simulation results show that data load for the following stages are greatly alleviated by the proposed method.

Different types of ground objects have different scatter properties and thus it is a quite challenging work to model the clutter. Experimental Results indicate that PR outperforms 2-para CFAR in both speed and accuracy. For coastal regions, it's preferred to use a detection method without modeling the clutter. Note that we are focusing on the general SAR ATR process and aiming at improving the performance as a whole. Sometimes, if too much effort is put into one stage, say finding a suitable but complex model for the clutter in the detection stage, the improvements made in this stage will not necessary be kept.

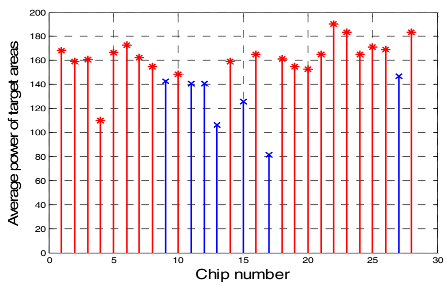
Prior knowledge is of great significance to ship detection and discrimination. To some extent, the whole process is based on prior knowledge, such as the difference in target and clutter



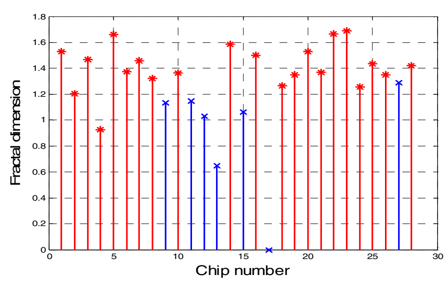
(a)



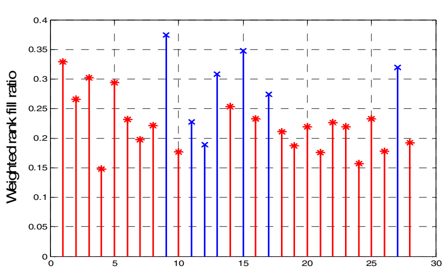
(b)



(c)



(d)



(e)

Fig.10. The five discrimination features extracted from the 28 candidates. (a) Log standard deviation. (b) Number of 8 connected regions. (c) Average power of target areas. (d) Fractal dimension. (e) Weighted rank fill ratio. Asterisk represents targets and cross represents clutters.

echoes, difference in land and target areas, high contrast between target point and its surrounding areas. Current discriminators such as QD, SVM and BNN also take advantage of the prior knowledge from similar training samples. However, this kind of knowledge is not easy to access and at the current time is not reliable. K-means discriminator makes use of the feature information of target and clutter in feature space to discriminate target from clutter. If the feature group is chosen properly, K-means will behave well. Experimental results present that the proposed feature group is more discriminative than the Lincoln Laboratory feature group.

Currently, the proposed scheme has been verified using TerraSAR-X data and some preliminary results are obtained. Further testing is needed to verify the effectiveness of this scheme in diversified high resolution SAR imagery as long as the data is available.

ACKNOWLEDGMENT

The TerraSAR-X data used in this paper was offered to us by our laboratory for experiment purpose. The authors would like to thank the editors and anonymous reviewers for their helpful comments and suggestions that will greatly improve this paper.

REFERENCES

[1] L.M. Novak, S.D. Helversen, and C.M. Netishen, "Performance of a high resolution polarimetric SAR automatic target recognition system," *The Lincoln Laboratory Journal*, vo1.6, no.1, 1993.

[2] L.M. Novak, G.J. Owirka, W.S. Brower, and A.L. Weaver, "The automatic target-recognition system in SAIP," *The Lincoln Laboratory Journal*, vo1.10, no.2, 1997.

[3] P. Lombardo, M. Sciotti, and L.M. Kaplan, "SAR prescreening using both target and shadow information," in *Proc. IEEE Radar Conf., 2001*, pp. 147–152.

[4] C. Zhang, T. Zou, and Z.Z. Wang, "A Target Discrimination Algorithm for High Resolution SAR imagery," *International Conference on Robotics, Intelligent Systems and Signal Processing, IEEE, 2003*, pp. 863-867.

[5] M. Tello, C. López-Martínez, and J. J. Mallorqui, "A novel algorithm for ship detection in SAR imagery based on the wavelet transform," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, no. 2, pp. 201–206, Apr. 2005.

[6] Y.H. Wang and H.W. Liu, "A hierarchical ship detection scheme for high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no.10, pp. 4173-4184, Oct. 2012.

[7] M. Tello, C. Lopez-Martinez, J. Mallorqui, and A. Aguasca, "Use of the multiresolution capability of wavelets for ship detection in SAR imagery," in *Proc. IGARSS*, vol. 6, 2004, pp. 4247–4250.

[8] L.M. Novak and S.D. Halversen, "Effects of polarization and resolution on SAR ATR," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 1, pp. 102–116, Jan. 1997.

[9] D.E. Kreithen, S.D. Halversen, and G.J. Owirka, "Discriminating targets from clutter," *Lincoln Lab. J.*, vol. 6, no. 1, pp. 25–52, 1993.

[10] Saur G., Teutsch M., "SAR signature analysis for TerraSAR-X-based ship monitoring". *Proc. SPIE* Vol. 7830, *Image and Signal Processing for Remote Sensing*, Toulouse, France, September 20–22, 2010.

[11] K.I. Ranney, H. Khatri and L.H. Nguyen, "Comparison of three different detectors applied to synthetic aperture radar data," *International Society for Optics and Photonics*, 2002: 91-99.

[12] T.D. Ross, J.J. Bradley, L.J. Hudson, and M.P.O'Connor, "SAR ATR: So what's the problem? An MSTAR perspective," in *Proc. SPIE Conf. Algorithms SAR Imagery IV, Orlando, FL, 1999*, vol. 3721, pp. 662–672.

[13] R.C. Gonzalez and R.E. Woods, Digital Image Processing, 3nd ed. Beijing, China: Publ. House Electron. Ind., 2011, translated by Q. Ruan, Y. Ruan, and etc., (in Chinese).

[14] R.C. Gonzalez, R.E. Woods, and S.L. Eddins,"Digital image processing using MATLAB." Vol. 2. Tennessee: Gatesmark Publishing, 2009.

[15] K.E. Darymli, P.M. Guire, D. Power, and C. Moloney,"Target detection in synthetic aperture radar imagery: a state-of-the-art survey," *Journal of Applied Remote Sensing* 7.1 (2013): 071598-071598.

[16] M.C. Burl, G.J. Owirka, and L.M. Novak, "Texture discrimination in synthetic aperture radar imagery," in *Proc. IEEE 23rd Asilomar Conf. Signals, Syst., Comput.*, 1989, pp. 399–404.

[17] N. Sarkar and B. B. Chaudhuri, "An efficient differential box-counting approach to compute fractal dimension of image," *IEEE Trans. Syst. Man Cybernet.*, vol. 24, pp. 115–120, Jan. 1994.

[18] A.K. Jain, "Data clustering: 50 years beyond k-means," *Pattern Recognit. Lett.*, vol. 31, no. 8, pp. 651–666, Jun. 2010.