A Kernel-Based Nonparametric Regression Method for Clutter Removal in Infrared Small-Target Detection Applications

Yanfeng Gu, *Member, IEEE*, Chen Wang, BaoXue Liu, and Ye Zhang

***Abstract*—Small-target detection in infrared imagery with a complex background is always an important task in remote-sensing fields. Complex clutter background usually results in serious false alarm in target detection for low contrast of infrared imagery. In this letter, a kernel-based nonparametric regression method is proposed for background prediction and clutter removal, furthermore applied in target detection. First, a linear mixture model is used to represent each pixel of the observed infrared imagery. Second, adaptive detection is performed on local regions in the infrared image by means of kernel-based nonparametric regression and two-parameter constant false alarm rate (CFAR) detector. Kernel regression, which is one of the nonparametric regression approaches, is adopted to estimate complex clutter background. Then, CFAR detection is performed on "pure" target-like region after estimation and removal of clutter background. Experimental results prove that the proposed algorithm is effective and adaptable to small-target detection under a complex background.**

***Index Terms*—Clutter removal, constant false alarm rate (CFAR), infrared image, kernel regression, target detection.**

I. Introduction

SMALL-TARGET detection is one of the most important applications of thermal infrared imagery. Infrared small-target detection technology has developed rapidly these years, and lots of effective methods were proposed. In an infrared image with complex background, i.e., low signal-to-clutter ratio (SCR), the contrast between targets and background is very low, and targets have no concrete shape and texture because of long imaging distance, so target detection in single-frame infrared image with low SCR has been considered as a difficult and challenging problem. Generally, the high gray area of background clutters can blur the small targets, and a strong background fluctuation may lead to a high rate of false alarm in detection. Moreover, background constitutes a large proportion of an infrared image. Therefore, the detection method based on background prediction and suppression is available. Existing background suppression methods for single-frame infrared image are mainly classified into the following two categories: One is the filtering methods [1]–[3], and the other one is statistical regression [4], [5]. The filtering methods include processing in space domain, which uses filter templates [1], morphological operators [2], etc., and processing in frequency domain, which relies on eliminating the low-frequency component [3]. The filtering methods can suppress most part of the correlative background but may be easily interfered because of strong fluctuation of complex background clutters. Regression methods are classified as parametric [4] and nonparametric [5] methods. Classical parametric regression methods rely on a specific model of background clutters and seek to estimate the parameters of this assumed model. In comparison with the parametric methods, nonparametric methods rely on the data itself to estimate the regression function. In practice, nonparametric methods are more suitable and adaptive for complex background because of lack of *a priori* knowledge about background clutters. As a result of the recent development of machine learning theory, kernel methods have been used widely in pattern analysis and statistical regression problems [6].

In this letter, a small-target detection algorithm in infrared image is proposed, which predicts and eliminates the complex background clutter by a kernel-based nonparametric regression model and obtains residual "pure" target-like image which only consists of noise and possible targets on local regions in infrared images. Then, a two-parameter constant false alarm rate (CFAR) [7] detecting algorithm is performed to extract the small target from the "pure" target-like image.

II. Methodology

A. Mathematical Model of an Infrared Image

Generally, an infrared image consists of three components: target, background, and noise. It can be described by an additive model as follows:

$$f(x,y) = f_T(x,y) + f_B(x,y) + n(x,y) \tag{1}$$

where $f(x,y)$ is the gray scale of a pixel (x,y) at the original infrared image, $f_T(x,y)$ is the gray scale of a pixel at the region of targets, $f_B(x,y)$ is the gray scale of a pixel at the background, and $n(x,y)$ is the intensity of the random noise.

Therefore, the conventional idea of target detection is to eliminate the background component $f_B(x,y)$ first. The proposed method is based on background prediction with the purpose of obtaining an exact estimation of the background clutter $\hat{f}_B(x,y)$. Then, the predicted background clutter can be eliminated from the original image to obtain the "pure" target-like image containing potential target candidates.

Manuscript received May 29, 2009; revised September 2, 2009 and November 18, 2009. Date of publication February 17, 2010; date of current version April 29, 2010. This work was supported in part by the National Natural Science Foundation of China under Project No. 60972144 and 60972143 and in part by the Heilongjiang Province Postdoctoral Scientific and Research Foundation and Research Fund for the Doctorial Program of Higher Education of China (No. 20092302110033).

The authors are with the School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China (e-mail: guyf@hit.edu.cn; wangchen300@163.com; baoxue_liu@sina.com; zhye@hit.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2009.2039192

B. Elimination of Background Clutter Based on Nonparametric Statistical Regression

The nonparametric statistical regression model of 2-D sample data is written as follows:

$$g_i = m(\mathbf{z}_i) + \varepsilon_i, \qquad i = 1, 2, \ldots, N \tag{2}$$

where $\mathbf{z}_i$ is a 2×1 vector like $\mathbf{z}_i = [x_i, y_i]^T$, with (x_i, y_i) as the spatial coordinates of a sample pixel; g_i is the image gray of a sample pixel; $m(\mathbf{z}_i)$ is the regression function whose form is unspecified, which should be a nonlinear function to approximate the complex background better; ε_i is a random error or interference that is uncorrelated with $\mathbf{z}_i$, where, generally, ε_i is supposed to be a Gaussian random variable with zero mean; and N is the total number of sample pixels which are used to estimate the regression model.

The parametric methods generally assume a signal model of background clutter such as a Gaussian model or Gaussian mixture model to predict the background. Under the complex background having no priori, the parametric regression model for estimation is not adaptive and precise enough. Considering the complexity of background fluctuation, nonparametric regression is more reasonable in fitting the background component.

To estimate the value of regression function at any point $\mathbf{z}$, if $\mathbf{z}$ is near the sample $\mathbf{z}_i$, the local expansion of the regression function $m(\mathbf{z}_i)$ can be utilized, which is given as follows:

$$m(\mathbf{z}_i) = m(\mathbf{z}) + \{\nabla m(\mathbf{z})\}^T (\mathbf{z}_i - \mathbf{z}) + \frac{1}{2}(\mathbf{z}_i - \mathbf{z})^T \{\mathcal{H} m(\mathbf{z})\} (\mathbf{z}_i - \mathbf{z}) + \cdots \tag{3}$$

where ∇ is the gradient (2×1) operator and $\mathcal{H}$ is the Hessian (2×2) operator, which is a symmetrical matrix. Then, (3) can be written as follows:

$$m(\mathbf{z}_i) = \beta_0 + \boldsymbol{\beta}_1^T (\mathbf{z}_i - \mathbf{z}) + \boldsymbol{\beta}_2^T vech\left\{(\mathbf{z}_i - \mathbf{z})(\mathbf{z}_i - \mathbf{z})^T\right\} + \cdots \tag{4}$$

where $vech(\cdot)$ acts as a half-vectorization operator that transforms the lower triangular portion of a symmetrical matrix into a vector. For example, for a 2×2 symmetrical matrix $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ with $a_{12} = a_{21}$

$$vech\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = [a_{11} a_{12} a_{22}]^T \tag{5}$$

and contrasting (3) with (4), we can conclude that $\beta_0 = m(\mathbf{z})$, which is the value of the pixel to be estimated. Then, we can also know that the vectors $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are

$$\boldsymbol{\beta}_1 = \nabla m(\mathbf{z}) = \left[\frac{\partial m(\mathbf{z})}{\partial x}, \frac{\partial m(\mathbf{z})}{\partial y}\right]^T \tag{6}$$

$$\boldsymbol{\beta}_2 = \frac{1}{2}\left[\frac{\partial^2 m(\mathbf{z})}{\partial x^2}, 2\frac{\partial^2 m(\mathbf{z})}{\partial xy}, \frac{\partial^2 m(\mathbf{z})}{\partial y^2}\right]^T. \tag{7}$$

According to the least-square theory, $\{\beta_n\}$ is solved from the following optimization problem:

$$\min_{\{\beta_n\}} \sum_{i=1}^{N} \Big[g_i - \beta_0 - \boldsymbol{\beta}_1^T (\mathbf{z}_i - \mathbf{z}) - \boldsymbol{\beta}_2^T \times vech\left\{(\mathbf{z}_i - \mathbf{z})(\mathbf{z}_i - \mathbf{z})^T\right\} - \cdots\Big]^2 \cdot K_H(\mathbf{z}_i - \mathbf{z}) \tag{8}$$

with

$$K_H(\mathbf{u}) = \frac{K(\mathbf{H}^{-1}\mathbf{u})}{\det(\mathbf{H})} \tag{9}$$

where $K_H(\cdot)$ is the kernel weight function that supplies sample data with different weight values related to distance away from the pixel estimated; $K(\cdot)$ is the 2-D probability density function, which is defined as a kernel function; and $\mathbf{H}$ is a 2×2 smoothing matrix, which decides the quantity of samples and affects the smoothness and precision of estimation. To simply calculate, $\mathbf{H}$ can be obtained from a model as $\mathbf{H} = h\mu\mathbf{I}$, where $\mathbf{I}$ is an identity matrix, μ is a scalar capturing local density of samples and is nominally set to one, and h is the global smoothing parameter [5].

In fact, this optimization problem can be described as a weighted least-square estimation model. Because we just focus on the gray scale of the pixels, it is only necessary to calculate β_0 from the estimation model, and the local gradient $\boldsymbol{\beta}_1$ is not our concerned problem. The matrix form of the estimation model is written as follows:

$$\hat{m}(\mathbf{z}) = \hat{\beta}_0 = \mathbf{e}_1^T \left(\mathbf{Z}_\mathbf{z}^T \mathbf{W}_\mathbf{z} \mathbf{Z}_\mathbf{z}\right)^{-1} \mathbf{Z}_\mathbf{z}^T \mathbf{W}_\mathbf{z} \mathbf{g} \tag{10}$$

where

$$\mathbf{e}_1 = (1, 0)^T \quad \mathbf{g} = [g_1, g_2, \ldots, g_N]^T$$

$$\mathbf{W}_\mathbf{z} = \mathrm{diag}\left[K_H(\mathbf{z}_1 - \mathbf{z}), K_H(\mathbf{z}_2 - \mathbf{z}), \ldots, K_H(\mathbf{z}_N - \mathbf{z})\right]$$

$$\mathbf{Z}_\mathbf{z} = \begin{bmatrix} 1 & (\mathbf{z}_1 - \mathbf{z})^T & vech^T\left\{(\mathbf{z}_1 - \mathbf{z})(\mathbf{z}_1 - \mathbf{z})^T\right\} & \cdots \\ 1 & (\mathbf{z}_2 - \mathbf{z})^T & vech^T\left\{(\mathbf{z}_2 - \mathbf{z})(\mathbf{z}_2 - \mathbf{z})^T\right\} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (\mathbf{z}_N - \mathbf{z})^T & vech^T\left\{(\mathbf{z}_N - \mathbf{z})(\mathbf{z}_N - \mathbf{z})^T\right\} & \cdots \end{bmatrix}.$$

After the aforementioned matrix operation, the value of gray at any pixel can be estimated precisely from the statistical regression model, and the prediction of the entire image can be obtained. Let $\hat{f}_B(x, y)$ be the estimation of the background clutter; then $\hat{f}_B(x, y) = \hat{m}(\mathbf{z})$ according to (10). Now, we can rewrite (1) as follows:

$$f(x, y) = f_T(x, y) + \hat{f}_B(x, y) + n'(x, y) \tag{11}$$

where $\hat{f}_B$ is the predicted background from regression estimation and n' ($n' = n + \varepsilon$) is the residual noise that contains noise n existing in the original infrared image and the bias ε of the kernel regression prediction. Thus, we can represent the "pure" target-like image as

$$f'(x, y) = f_T(x, y) + n'(x, y). \tag{12}$$

C. CFAR Detector

After the background component is eliminated, a "pure" target-like image, including only target and noise, is obtained. To extract the real targets and suppress the residual noise, a CFAR detector is used in this letter.

CFAR detector is a test based on pixel, and it obeys the Neyman–Pearson criterion. It can confirm whether the pixel belongs to the target region by means of a contrast between the pixel value and a specific threshold. The threshold changes with the surrounding clutter power in order to retain a constant false alarm ratio in the global image. This important threshold

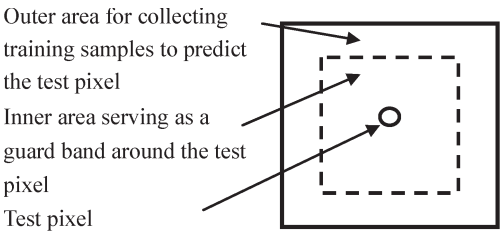

Fig. 1. Two-window background prediction and CFAR detection in the local region of the infrared image.

is calculated from an assumed probability distribution of the clutter model. In this letter, a two-parameter CFAR detector based on an assumption of the Gaussian clutter model is adopted. A local clutter in the infrared image can be reasonably considered as Gaussian distributed, and the bias of regression estimation is also supposed as a Gaussian random variable, which is mentioned in Section II-B. Then, the residual noise can be supposed that it obeys Gaussian distribution. Therefore, this two-parameter CFAR detector is reasonable.

Whether the pixel tested belongs to the target region or not is judged by the rule defined in the following. The two-parameter CFAR detector calculates the detection likelihood ratio

$$\frac{f'(x,y)-\hat{\mu}_c}{\hat{\sigma}_c} > \eta_{\mathrm{CFAR}} \tag{13}$$

where $\hat{\mu}_c$ and $\hat{\sigma}_c$ are the estimated mean and standard deviation of the clutter on the local regions of the residual background-eliminated image and $f'(x,y)$ is the amplitude of the pixel under test.

If (13) is satisfied, the test pixel will be considered as a pixel of targets; otherwise, it will be judged as a pixel of clutter. Moreover, the threshold η_{CFAR} is calculated by the parameters of the assumed Gaussian distribution model and the CFAR P_{fa}. We assume that the model of noise in the residual image obeys a Gaussian distribution with a zero mean and one variance. Then, the value of η_{CFAR} can be given as follows:

$$\eta_{\mathrm{CFAR}} = \sqrt{-2\ln\left(\sqrt{2\pi}P_{fa}\right)}. \tag{14}$$

III. Target Detection Using Kernel Regression

The detailed process steps of the proposed algorithm for small-target detection in infrared image are as follows.

1) Set the size of the hollow background window [8] used to pick samples for gray estimation. The window consists of two regions (the inner and outer window regions), as shown in Fig. 1. The inner window should be larger than the small target in order to prevent the target pixels from leaking out and interfering with the precision of background estimation. The outer window contains sample pixels used to estimate the value of the regression function at the central pixel of the dual window.
2) Select the form of the kernel function; set the parameters and the smoothing matrix. It is known from the regression theory that the choice of the kernel function is various, such as the Gaussian kernel, exponential kernel, etc.
3) Predict the gray scale of the test pixel by means of the kernel regression estimation described in Section II-B. We are interested in estimating the background component, so we use a dual window to avoid the interference of the target component. Move the window to finish the estimation pixel by pixel, and the predicted background is obtained. Then, subtract the predicted image from the original image, and get a residual image with possible targets and noise, namely, the "pure" target-like image.
4) Set the CFAR, and use the two-parameter CFAR detector to process the residual image. The dual window is also used here to estimate the parameters of the local clutter. After an operation pixel by pixel according to (13), most of the noise and false alarms can be eliminated in this step.
5) Output the detected results.

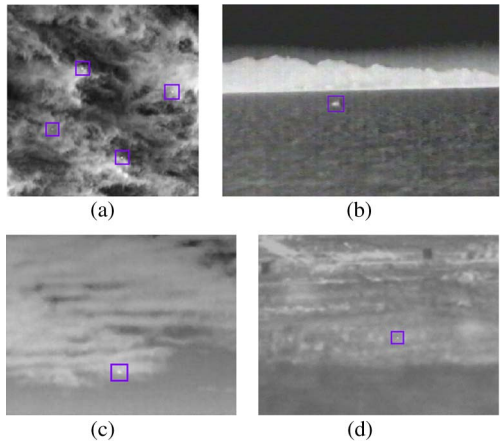

Fig. 2. Original infrared images for test.

IV. Experiments and Results

In order to validate the proposed algorithm, experiments are performed on both of the simulated and real infrared images with sky or sea background, which are shown in Fig. 2. Fig. 2(a) shows the simulated image, which are composed of real background image and simulated targets. The approach that is used to generate the simulated targets can be found in [9]. Fig. 2(b)–(d) shows all images with real targets.

Two kinds of functions are adopted as the kernel function in this letter. One is the popular Gaussian radial basis function (RBF) which is defined as

$$K(x,y) = \exp\left(-\frac{\|x-y\|^2}{\sigma^2}\right) \tag{15}$$

where σ is a scale parameter controlling the capacity of local samples. The choice of the parameter σ is very critical. It should be carefully chosen so that the overall data variations can be fully exploited by the Gaussian RBF kernel. The Gaussian RBF kernel is translation invariant, and its Fourier transform is also Gaussian.

The other function adopted as a kernel is a multiscale wavelet basis function [1]. The multiscale wavelet kernel function based on the Morlet wavelet [10] is defined as follows:

$$K(x,y) = \sum_{l=1}^{L} \cos\left(\frac{1.75x}{a_l}\right)\exp\left(-\frac{x^2}{2a_l^2}\right) \times \cos\left(\frac{1.75y}{a_l}\right)\exp\left(-\frac{y^2}{2a_l^2}\right) \tag{16}$$

TABLE I
PARAMETERS SET IN THE PROPOSED KERNEL REGRESSION DETECTION METHOD

Parameters / Images	Inner size	Outer size	σ	h (Gauss)	a	h (Wavelet)
a	5×5	11×11	1.2	2	0.6	0.9
b	21×21	27×27	1.3	14	1.1	1
c	7×7	11×11	1.5	10	1.5	0.9
d	5×5	9×9	0.5	11.1	0.5	1

where l is used to denote different scales, L is the upper limit of scales, a_l is related to l, and $a_l = a^l$, with a defined as a dilation factor of the wavelet function. The multiscale wavelet kernel has both the wavelet decomposition property and the multiscale property that can help approximate nonlinear signals.

The sizes of the inner and outer windows are set, respectively, in different images according to the rules described in Section III point 1). The scale factors σ and a and the global smoothing parameter h are chosen by fivefold cross validation in which the first several rows (20%) of the image are portioned into five subsamples. The cross validation is repeated five times, and each subsample is used once as the validation data. The mean-square error function is used as the minimized optimal rule in cross validation. In the process of CFAR detecting, the CFAR is set as 10^{-5}. The parameters related to the proposed algorithm for each image are listed in Table I.

The residual image after background suppression by kernel regression is assumed as obeying Gaussian distribution. Fig. 2(d) is chosen as an instance to be analyzed. The residual image after clutter removal by wavelet kernel regression is shown at the top of Fig. 3. The probability density function of the gray intensity in the residual image is also displayed here to produce evidence in support of the Gaussian assumption.

As observed from the PDF graphic earlier, the distribution of intensity in residual images approximately obeys the Gaussian distribution. Furthermore, a Gaussian fitting is processed with maximum likelihood estimation. The fitted results are also shown in Fig. 3. The fitted Gaussian distribution has a mean of 0.0159 and a standard deviation of 5.6790 with the root-mean-square error 1.6×10^{-3}(Gauss kernel) and has a mean of 0.0054 and a standard deviation of 4.7557 with the root-mean-square error 4.8×10^{-4}(multiscale wavelet kernel).

Another three algorithms are adopted to make a comparison with the proposed algorithm. These algorithms are morphological detection algorithm (a conventional top-hat transformation adopted here [11]), adaptive Butterworth high-pass filter method [3], and directional filters based on the least square support vector machine (LS-SVM) algorithm [1].

To compare the performance of the algorithms more clearly, two common evaluation indicators [12], i.e., SCR gain and background suppression factor (BSF), are adopted. They are defined as follows:

$$SCR\ gain = \frac{(S/C)_{\text{out}}}{(S/C)_{\text{in}}} \qquad BSF = \frac{C_{\text{in}}}{C_{\text{out}}} \tag{17}$$

where S is the signal amplitude and C is the clutter standard deviation. The subscripts in and out express the images before and after the detection.

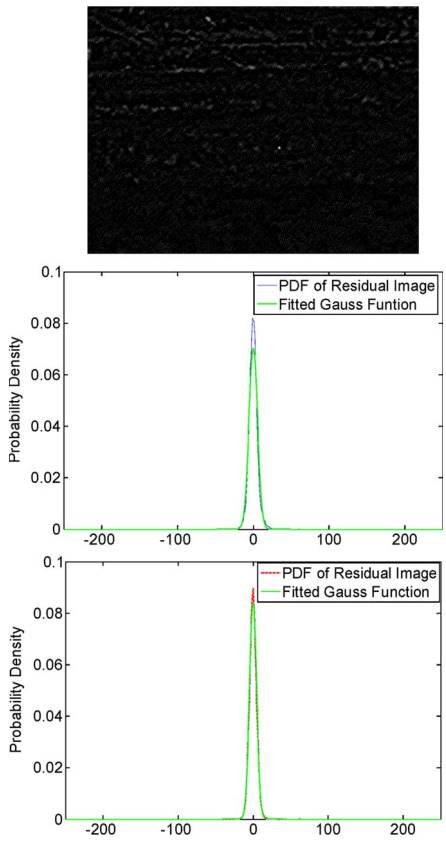

Fig. 3. Residual image after background suppression and the fitted result of Gaussian probability distribution function (PDF). From top to bottom, the residual image after clutter removal for Fig. 2(d), the fitted Gaussian function for the residual of Gaussian kernel regression, and the fitted Gaussian function for the residual of multiscale wavelet kernel regression.

TABLE II
EVALUATION INDICATORS OF DIFFERENT DETECTION METHODS

Methods	Indicators	Image a	Image b	Image c	Image d
Top-hat	SCR Gain	6.3668	8.0223	2.9982	4.1059
	BSF	12.8662	9.3287	14.1252	7.688
BHPF	SCR Gain	3.1322	6.7634	2.2255	4.4809
	BSF	11.1042	9.4095	9.8954	9.7885
LS-SVM	SCR Gain	9.0615	1.5182	12.0609	27.5835
	BSF	**43.5968**	**108.3437**	19.8779	41.0807
KR-CFAR (Gauss)	SCR Gain	19.9041	**29.3041**	20.4739	22.4774
	BSF	32.6881	27.6424	33.2384	34.2563
KR-CFAR (Wavelet)	SCR Gain	**25.3335**	29.0120	**21.6572**	**31.7351**
	BSF	35.4818	27.7253	**38.5484**	**42.2435**

The experimental data of the aforementioned indicators are listed in Table II. The receiver operation characteristic (ROC) curves of the detection results for each image in Fig. 2 are shown in Fig. 4. The ROC represents the varying relationship of the detection probability $\boldsymbol{P}_d$ and false alarm rate $\boldsymbol{P}_f$, and it can provide a quantitative comparison of the detection performance. $\boldsymbol{P}_d$ is defined as the ratio of the number of detected pixels to the number of real target pixels. $\boldsymbol{P}_f$ is defined as the ratio of the number of false alarms to the total number of pixels in the whole image.

In Table II, it is shown that the proposed detection algorithm based on kernel regression has great performance in both SCR

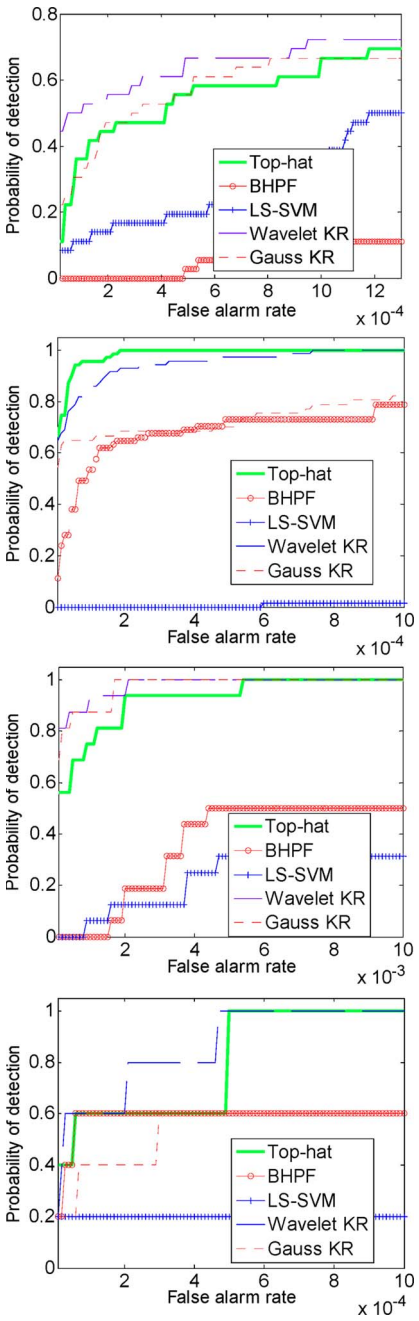

Fig. 4. Comparison of ROC curves obtained by four methods (where the proposed algorithm includes Gauss kernel and multiscale wavelet kernel regressions).

gain and BSF (bold type shows which method has the most efficient performance for each image). In most cases, the algorithm based on wavelet kernel regression performs better than that based on Gauss kernel regression. The multiscale wavelet kernel really shows some merits on signal analysis. The novel method based on LS-SVM has a little better performance in several conditions. It is known from the related theory that a means of segmentation is used to suppress the background and eliminate the bias of prediction in this method. Moreover, the feature of the second-order gradient is enhanced by an operation of directional fusion. That is why the LS-SVM method has a little higher performance sometimes. However, this method may lead to a low probability of detection, which is caused by the effect of directionality in this filter method. Therefore, it usually misses pixels when detecting a little bigger or blurry target. For example, it can hardly extract the target from image (b), and the ROC curve also shows this shortcoming.

In Fig. 4, it can be found that the proposed algorithm can control the false alarm rate and improve the probability of detection. Combining the indicators with ROC curves, the proposed method based on nonparametric kernel regression has a better detection performance than others.

V. Conclusion

In this letter, a kernel regression-based background prediction algorithm has been proposed for clutter removal and small-target detection in infrared images. The SCR gain and BSF are combined with the ROC curve to evaluate the performance of both clutter removal and target detection. A united evaluation like this is more reasonable and effective than that relying on only one aspect of them. Considering both clutter removal and target detection, the proposed algorithm has a better performance than other methods. The experimental results also indicate that the kernel regression utilizing multiscale wavelet kernel is superior to that utilizing Gaussian kernel in the work because of the wavelet decomposition property and the multiscale property of the multiscale wavelet kernel.

References

[1] P. Wang, J. W. Tian, and C. Q. Gao, "Infrared small target detection using directional highpass filters based on LS-SVM," *Electron. Lett.*, vol. 45, no. 3, pp. 156–158, Jan. 2009.

[2] J. F. Guo and G. L. Chen, "Analysis of selection of structural element in mathematical morphology with application to infrared point target detection," *Proc. SPIE*, vol. 6835, pp. 683 50P-1–683 50P-8, Nov. 2007.

[3] L. Yang, J. Yang, and K. Yang, "Adaptive detection for infrared small target under sea–sky complex background," *Electron. Lett.*, vol. 40, no. 17, pp. 1083–1085, Aug. 2004.

[4] Y. Xiong, J. X. Peng, M. Y. Ding, and D. H. Xue, "An extended track-before-detect algorithm for infrared target detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 3, pp. 1087–1092, Jul. 1997.

[5] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, Feb. 2007.

[6] P. Yee and S. Haykin, "Pattern classification as an ill-posed, inverse problem: A regularization approach," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 1993, vol. 1, pp. 597–600.

[7] L. M. Novak, S. D. Halversen, G. J. Owirka, and M. Hiett, "Effects of polarization and resolution on SAR ATR," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 1, pp. 102–115, Jan. 1997.

[8] A. Banerjee, P. Burlina, and C. Diehl, "A support vector method for anomaly detection in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2282–2290, Aug. 2006.

[9] T. Soni, J. R. Zeidler, and W. H. Ku, "Performance evaluation of 2-D adaptive prediction filters for detection of small objects in image data," *IEEE Trans. Image Process.*, vol. 2, no. 3, pp. 327–340, Jul. 1997.

[10] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

[11] F. Meyer, "Iterative image transformations for an automatic screening of cervical smears," *J. Histochemistry Cytochemistry*, vol. 27, no. 1, pp. 128–135, 1979.

[12] C. I. Hilliard, "Selection of a clutter rejection algorithm for real-time target detection from an airborne platform," *Proc. SPIE*, vol. 4048, pp. 74–84, Apr. 2000.