Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 1529704, 19 pages
https://doi.org/10.1155/2020/1529704

Research Article

Infrared Small Target Detection with Total Variation and Reweighted ℓ_1 Regularization

Houzhang Fang,[1] **Min Chen**,[1] **Xiyang Liu**,[1] **and Shoukui Yao**[2]

[1]*Software Engineering Institute, School of Computer Science and Technology, Xidian University, Xi'an, Shaanxi 710071, China*
[2]*Huazhong Institute of Electro-Optics, Wuhan 430223, China*

Correspondence should be addressed to Xiyang Liu; xyliu@xidian.edu.cn

Received 1 September 2019; Revised 12 December 2019; Accepted 24 December 2019; Published 27 January 2020

Academic Editor: Bogdan Smolka

Copyright © 2020 Houzhang Fang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Infrared small target detection plays an important role in infrared search and tracking systems applications. It is difficult to perform target detection when only a single image with complex background clutters and noise is available, where the key is to suppress the complex background clutters and noise while enhancing the small target. In this paper, we propose a novel model for separating the background from the small target based on nonlocal self-similarity for infrared patch-image. A total variation-based regularization term for the small target image is incorporated into the model to suppress the residual background clutters and noise while enhancing the smoothness of the solution. Furthermore, a reweighted sparse constraint is imposed for the small target image to remove the nontarget points while better highlighting the small target. For higher computational efficiency, an adapted version of the alternating direction method of multipliers is employed to solve the resulting minimization problem. Comparative experiments with synthetic and real data demonstrate that the proposed method is superior in detection performance to the state-of-the-art methods in terms of both objective measure and visual quality.

1. Introduction

Infrared small target detection has been receiving considerable attention due to their wide applications in many fields, such as remote sensing [1], surveillance [2], environment monitoring [3], aerospace [4], and unmanned aerial vehicles [2]. The robust detection of infrared small target is still a challenging problem. The main difficulties of infrared small target detection lie in the following facts: (1) in realistic applications, the infrared imaging system is easily disturbed by such factors as the climate, atmospheric turbulence, and vibrating, as a result of which the observed small target often suffers complex background clutters and noise, thus a low signal-to-noise and signal-to-clutter ratio; (2) small target pixels occupy a smaller proportion of the whole observed image due to the long imaging distance; and (3) the infrared small target has no obvious texture and shape features.

A large number of methods have been developed to address the issues of small target detection. These methods can be roughly classified into two categories: single-frame detection [5–16] and sequential multiframe detection [17–19]. Recently, Gao et al. [17] employed the mixture of the Gaussians model [20] with the Markov random field to model the complex noise of which the target is assumed as a component. Then the small targets are separated from noise by using the spatiotemporal information. The sequential methods usually utilize the temporal information of the multiframe image sequence which requires some strong assumptions of consistent information of targets and background between frames. However, these assumptions and prior knowledge can hardly be attained in applications. As such, single-frame detection has been attracting the attention of many researchers. In this research, we consider single-frame detection. In the conventional single-frame detection methods, background clutter suppression-based

methods are the representative type for their implementation simplicity and computational efficiency, such as Top-hat filter [7], Max-mean and Max-median filter [5], and multilevel filter [6], which first suppressed the background clutters and noise using the filter and then extract the small target with an intensity threshold. Moon et al. [6] utilized a set of filters (multilevel filter) to enhance the target and suppress background clutters. However, these methods will lead to high false alarm rates and poor detection performance when the signal-to-clutter ratio is low. To better suppress the background clutters and noise while enhancing the small target, some methods [8, 9] based on local contrast enhancement have been proposed. Chen et al. [8] proposed a local contrast detection method by defining a new contrast measure inspired by the biological visual mechanism. Han et al. [9] proposed an improved difference of Gabor filter method based on human visual system to suppress the background edges and lower the false alarm rates. These methods try to improve the ability to suppress the clutters and noise as well as enhance the target. Shi et al. [16] proposed an enhanced high-boost-based multiscale local contrast measure (HB-MLCM) method to enhance the true target and suppress background clutters, and the experimental results verified the effectiveness of this method.

Different from the background clutter suppression-based methods, another representative method tackles the infrared small target detection based on the robust principal component analysis (RPCA) [21]. The core idea of this kind of approach is to seek the best sparse and low-rank approximation of the given observation matrix, where the target is regarded as the sparse component and the background is low-rank component. RPCA framework has been widely applied to infrared small target detection [22–26] for its generalization and computation efficiency and can be extended to cope with large-scale problems. In realistic applications, sparse and low-rank recovery normally cannot be directly applied to small target detection because the decomposed background image matrix usually does not satisfy the low-rank characteristics. To deal with this issue, Gao et al. [22] proposed a more generic method by constructing a new infrared patch-image (IPI) model using the repeated local patches based on the nonlocal self-correlation property of the infrared image. The small target detection task is converted into an optimization problem of recovering low-rank and sparse matrices, which can be solved efficiently using existing optimization solution techniques. Wang et al. [23] introduced a total variation (TV) constraint on the background into the RPCA framework for better preserving the edges and corners of background region so as to improve the background estimation and achieve good detection results. Yao et al. [25] tackled infrared small target detection using a coarse-to-fine two-phase method. In the coarse phase, the background is separated by exploiting the nonlocal self-similarity properties under the RPCA framework, and then the target is separated from the unstructured clutters and noise in the unstructured layer in the fine phase. In this method, it is still very difficult to separate the small target from the unstructured clutters and noise using only the local prior when the clutters and noise are heavy. Therefore, the detection performance for the coarse-to-fine method is still limited. Zhang et al. [26] proposed to constrain the small target with nonconvex ℓ_p norm so as to restore sparser target images, and the method achieved an impressive detection performance.

Recently, several convolution neural network- (CNN-) based infrared small target detection methods have been proposed [27–30]. Fan et al. [27] simulated the training data using the MNIST database (http://yann.lecun.com/exdb/mnist/) containing handwritten images and labels, and then the target and background are estimated based on the convolution kernels in the first layer of the CNN architecture, which are learned from the input data. The main disadvantage of applying the deep learning method to infrared small target detection is that feature learning will be very difficult for the CNN-based method as the infrared small target has no remarkable texture and shape features. In addition, because of insufficient training data, the visible datasets [27] or simulated infrared datasets [29, 30] are used to train the detection model. As a result, generalizability of the detection method is also limited.

In practical applications, the observed image acquired by the infrared thermal imaging often suffers from evolving background clutters and heavy noise, especially in poor imaging conditions. Such background clutters and random noise are the main interferences to target-background separation. To obtain a more stable detection, an appropriate formulation under the RPCA framework is employed to decompose the constructed infrared patch-image into the sparse target component, the low-rank background component, and the noise component [21]. With complex background clutters and noise, the simple sparse constraint cannot distinguish the small target from the heavy clutters and noise well. And the heavy residual clutters and noise in the small target component result in high false alarm rates, thus poor robustness of small target detection. To resolve this problem, we apply the TV regularization on the small target to suppress the residual clutters and noise in the target image. TV regularization is originally introduced by Rudin et al. [31] to deal with the image denoising problem. TV regularization and its variants are also successfully extended to handle other issues, such as clutter suppression in radar target detection [32], blind image deconvolution [33, 34], and image destriping [35]. Zhang et al. [32] proposed a radar extended target detection method by combining the sparse constraint and total variation for suppressing the clutters and noise. The rational of using total variation as a regularizer for removing the clutters and noise is that the total variation of the interfering image is significantly larger than that of the clean image, and the minimization of the total variation can remove the clutters and noise [31]. Existing research work [31–35] shows that TV-based regularization has good performances on removing the clutters and noise from images without rich textures and details, which is suitable for small target images. Thus, incorporation of TV regularization can prevent the clutters and noise from remaining in the target image. Additionally, we integrate a weighted ℓ_1 norm into the sparse constraint to better suppress the nontarget sparse points [24]. Since the size of the infrared object is small, the introduction of the weighting

factor into the ℓ_1 norm makes the target image sparser. Experiments show that the TV regularization and weighted mechanism constraints added in the RPCA formulation are suitable for small target detection.

The contributions of this paper are listed as follows:

(1) We propose a new variant of low-rank matrix decomposition based on the RPCA framework for the separation of the target from the background. The TV regularization and weighted ℓ_1 norm are combined to constrain the small target. The proposed model enforces the spatial smoothing constraint on the target image, thus effectively suppressing the residual clutters and noise with TV regularization. Furthermore, the model accurately characterizes the sparse small target while suppressing the nontarget points via weighted ℓ_1 norm constraint, which leads to a lower false alarm rate.

(2) The resulting constructed model contains the nuclear norm, TV norm, and weighted ℓ_1 norm, which is a challenging nonconvex optimization problem. To effectively solve this problem, an efficient numerical optimization algorithm based on the alternating direction method of multipliers is applied to solve the proposed minimization problem.

The remainder of this paper is organized as follows. Section 2 gives some notations and briefly introduces the IPI model. The proposed model is described in Section 3. Section 4 provides extensive experiments on synthetic and real infrared images to validate the effectiveness of the proposed method. We conclude this work in Section 5.

2. Notations and IPI Model

The mathematical notations we use in this paper are described in Table 1.

The original RPCA formulation decomposes the observation matrix into a low-rank and a sparse matrix [21]. Generally, the stable RPCA formulation for the infrared image model can be formulated as follows [21, 22]:

$$f_{\mathrm{D}} = f_{\mathrm{T}} + f_{\mathrm{B}} + f_{\mathrm{N}}, \tag{1}$$

where f_{D}, f_{T}, f_{B}, and f_{N} are the original infrared image, the target image, the background image, and the random noise image, respectively. Then, we need to formulate appropriate constraints for the target, the background, and the noise based on some prior knowledge.

In equation (1), the low-rank properties of the background image f_{B} typically do not hold true in practice. The seminal work of [22] first proposed to employ the nonlocal self-similarity of the background to construct a new infrared patch-image (IPI) model. To calculate three components, the IPI model is transformed into the following form:

$$\mathbf{D} = \mathbf{T} + \mathbf{B} + \mathbf{N}, \tag{2}$$

where $\mathbf{D}$, $\mathbf{T}$, $\mathbf{B}$, and $\mathbf{N}$ are the corresponding patch-images constructed from f_{D}, f_{T}, f_{B}, and f_{N} by sliding the window of size $a \times a$ from left to right and top to bottom with a fixed step size. The IPI model can be extended to practical application very well.

In practical infrared images, the small target usually has only several pixels or dozens of pixels compared with the whole image, so the target patch-image $\mathbf{T}$ can be described as a sparse matrix:

$$\|\mathbf{T}\|_0 < k, \tag{3}$$

where k is determined by the number of small targets and their sizes. $\|\cdot\|_0$ represents the ℓ_0 norm.

The background image f_{B} is usually considered to change slowly, so there are strong correlation properties between its local and nonlocal patches. And the low-rank property in $\mathbf{B}$ can be depicted as

$$\mathrm{rank}\,(\mathbf{B}) \le r, \tag{4}$$

where r is an integer determined by the rank of $\mathbf{B}$.

The noise is assumed to be Gaussian random noise, and $\mathbf{N}$ can be described as

$$\|\mathbf{D} - \mathbf{B} - \mathbf{T}\|_F \le \delta, \tag{5}$$

where δ is a very small positive constant.

According to the constructed IPI model, the infrared small target detection can be transformed into an optimization problem of recovering low-rank and sparse matrices, which can be formulated as

$$\begin{aligned} \min_{\mathbf{T},\mathbf{B}} \quad & \|\mathbf{B}\|_* + \lambda\|\mathbf{T}\|_1 \\ \text{s.t.} \quad & \|\mathbf{D} - \mathbf{T} - \mathbf{B}\|_F \le \delta, \\ & \mathbf{D} = \mathbf{T} + \mathbf{B} + \mathbf{N}, \end{aligned} \tag{6}$$

where $\|\cdot\|_*$ is introduced to constrain the low-rank property in $\mathbf{B}$. $\|\cdot\|_1$ is used to replace $\|\cdot\|_0$ to simplify the solution of this model. λ is a weighting parameter.

3. Proposed Method

In this section, we first present the proposed detection model and analyze the rational behind the model. Secondly, the solution of the proposed model is carefully designed. Finally, we give the target detection procedure.

3.1. Proposed Model. The small target detection based on the IPI model in (6) does work well in suppressing the clutters, noise, and nontarget points only using the ℓ_1 norm constraint for the target image. In order to better separate the small target from the constructed observation image, the TV regularization term and the adaptive weighted ℓ_1 norm on the target image are introduced into equation (6). The proposed model can be written as

$$\begin{aligned} \min_{\mathrm{T},\mathrm{B}} \quad & \|\mathbf{B}\|_{\mathrm{W_B},*} + \lambda_1\|\mathbf{T}\|_{\mathrm{W_T},1} + \lambda_2\|\nabla\mathbf{T}\|_1 \\ \text{s.t.} \quad & \|\mathbf{D} - \mathbf{T} - \mathbf{B}\|_F \le \delta, \\ & \mathbf{D} = \mathbf{T} + \mathbf{B} + \mathbf{N}, \end{aligned} \tag{7}$$

Table 1: Mathematical notations.

Notations	Explanation
$\mathbf{X}$	Matrix
x	Scalar
$\mathrm{TV}(\cdot)$	TV norm, which is defined as $=\sum_{i=1}^{m-1}\sum_{j=1}^{n-1}\sqrt{(x_{i,j}-x_{i+1,j})^2+(x_{i,j}-x_{i,j+1})^2}=\sum_{i=1}^{m-1}\lvert x_{i,n}-x_{i+1,n}\rvert+\sum_{j=1}^{n-1}\lvert x_{m,j}-x_{m,j+1}\rvert$
$\nabla\mathbf{X}$	$[\nabla_x\mathbf{X},\nabla_y\mathbf{X}]^T$
$\langle\mathbf{X},\mathbf{Y}\rangle$	Inner product between matrix $\mathbf{X}$ and $\mathbf{Y}$, which is defined as $\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbf{X}_{ij}\mathbf{Y}_{ij}$
$\lVert\mathbf{X}\rVert_0$	ℓ_0 norm of $\mathbf{X}$, which is the number of nonzero elements
$\lVert\mathbf{X}\rVert_1$	ℓ_1 norm of $\mathbf{X}$, which is the sum of all $\lvert\mathbf{X}_{ij}\rvert$, i.e., $\sum_{ij}\lvert\mathbf{X}_{ij}\rvert$
$\lVert\mathbf{X}\rVert_F$	The Frobenius norm of $\mathbf{X}$, which is defined as $\sqrt{\sum_{ij}\lvert\mathbf{X}_{ij}\rvert^2}$
$\lVert\mathbf{X}\rVert_*$	Nuclear norm of $\mathbf{X}$, which is the sum of singular values

where λ_1 and λ_2 are compromising parameters. The vector $\mathbf{w}_\mathrm{B}$ and the matrix $\mathbf{W}_\mathrm{T}$ are the weights for the background $\mathbf{B}$ and the target $\mathbf{T}$, respectively. The weight $\mathbf{w}_\mathrm{B}$ is incorporated into the nuclear norm term for better preserving the strong edges of the background image [24], which is defined as

$$w_{\mathrm{B},j}=\frac{1}{\sigma_{\mathrm{B},j}+\varepsilon_\mathrm{B}}, \tag{8}$$

where $\sigma_{\mathrm{B},j}$ is the jth singular value of $\mathbf{B}$ and $\mathbf{w}_\mathrm{B}=[w_{\mathrm{B},1},w_{\mathrm{B},2},\ldots,w_{\mathrm{B},n}]^T$ is a weight for $\sigma_{\mathrm{B},j}$. ε_B is a positive value. $\lVert\mathrm{B}\rVert_{\mathrm{w_B},*}=\sum_j w_{\mathrm{B},j}\sigma_{\mathrm{B},j}$ is the reweighted nuclear norm of $\mathbf{B}$.

The weight $\mathbf{W}_\mathrm{T}$ is employed to suppress the nontarget points while enhancing the small target, and each element of $\mathbf{W}_\mathrm{T}$ is defined as

$$\mathbf{W}_{\mathrm{T},ij}=\frac{C}{\left\lvert\mathbf{T}_{ij}\right\rvert+\varepsilon_\mathrm{T}}, \tag{9}$$

where $\mathbf{T}_{ij}$ is the pixel value of position (i,j) in the target image $\mathbf{T}$. C is a compromising constant and ε_T is a small positive constant to avoid division by zero. The advantage of the weighted ℓ_1 norm in equation (7) is to give every single element a different weight, thus better depicting the small target component. Because the small targets are a little brighter than other nontarget parts, this allows the small target to be better separated as the iteration proceeds. In equation (7), the TV regularization has good performance in removing the background clutters and the noise. The two main interferences can be effectively suppressed, and thus, the small target can be further highlighted. By balancing the sparsity prior $\lVert\mathbf{T}\rVert_{\mathbf{W}_\mathrm{T},1}$ and the TV prior $\lVert\nabla\mathbf{T}\rVert_1$ using parameters λ_1 and λ_2, the proposed regularization constraints on the small target yield a sound target detection result. Therefore, the combination of the TV regularization and the weighted ℓ_1 norm is a reasonable design for small target detection.

3.2. Solution of the Proposed Model. The minimization problem (7) with the equality and inequality constraints can be solved using the fast optimization algorithm. The difficulties in finding $\mathbf{T}$ are that $\lVert\mathbf{T}\rVert_{\mathrm{W_T},1}$ is nonquadratic and nonseparable and that the ℓ_1 norm term $\lVert\nabla\mathbf{T}\rVert_1$ is nonsmooth and nonseparable. The nuclear norm $\lVert\mathbf{B}\rVert_{\mathrm{w_B},*}$ is nonquadratic and nonseparable. We can see from the above analysis that the objective function and constraints in equation (7) have different structures. Therefore, it is useful and necessary to split and solve them separately, which is exactly the forte of the alternating direction method of multipliers (ADMM) [36]. The idea is to convert the minimization problem on $\mathbf{T},\mathbf{B}$ in (7) into a constrained one by introducing three auxiliary variables $\mathbf{Z}_1$, $\mathbf{Z}_2$, and $\mathbf{Z}_3$, and four Lagrangian multiplier variables $\mathbf{Y}_1$, $\mathbf{Y}_2$, $\mathbf{Y}_3$, and $\mathbf{Y}_4$.

In each iteration, the ADMM updates the splitting variables separately and alternatively by solving the partial augmented Lagrangian of (7). Problem (7) can be further converted into an unconstrained problem:

$$\begin{aligned}\min_{\mathrm{T,B}}\ &\lVert\mathbf{Z}_1\rVert_{\mathrm{w_B},*}+\lambda_1\lVert\mathbf{Z}_2\rVert_{\mathrm{W_T},1}+\lambda_2\lVert\mathbf{Z}_3\rVert_1+\langle\mathbf{Y}_1,\mathbf{Z}_1-\mathbf{B}\rangle\\&+\langle\mathbf{Y}_2,\mathbf{Z}_2-\mathbf{T}\rangle+\langle\mathbf{Y}_3,\mathbf{Z}_3-\nabla\mathbf{T}\rangle+\langle\mathbf{Y}_4,\mathbf{D}-\mathbf{T}-\mathbf{B}\rangle\\&+\frac{\beta}{2}\left(\lVert\mathbf{Z}_1-\mathbf{B}\rVert_F^2+\lVert\mathbf{Z}_2-\mathbf{T}\rVert_F^2+\lVert\mathbf{Z}_3-\nabla\mathbf{T}\rVert_F^2+\lVert\mathbf{D}-\mathbf{T}-\mathbf{B}\rVert_F^2\right),\end{aligned} \tag{10}$$

where $\langle\cdot,\cdot\rangle$ denotes the inner product operation. β is the positive penalty parameter. To solve the above minimization problem about several variables, we utilize the alternating minimization method. Its idea is to solve only one subminimization problem while fixing other variables in each step of the iterative procedure.

(1) For $\mathbf{Z}_1$ subproblem, it can be transformed to

$$\min_{\mathbf{Z}_1}\lVert\mathbf{Z}_1\rVert_{\mathrm{w_B},*}+\frac{\beta}{2}\left\lVert\mathbf{Z}_1-\mathbf{B}+\frac{1}{\beta}\mathbf{Y}_1\right\rVert_F^2. \tag{11}$$

The above problem is a nuclear norm minimization problem, which can be readily solved by singular value thresholding [37], so the solution of $\mathbf{Z}_1$ can be written as

$$\mathbf{Z}_1=\mathrm{SVT}_{\mathrm{w_B}/\beta}\left(\mathbf{B}-\frac{1}{\beta}\mathbf{Y}_1\right), \tag{12}$$

where $\mathrm{SVT}_\mu(\cdot)$ is the singular thresholding operator.

$$\mathrm{SVT}_{\mu}(\mathbf{Y}) = \mathbf{U}\mathrm{diag}\left[(\sigma-\mu)_{+}\right]\mathbf{V}^{T},$$
$$(\sigma-\mu)_{+} = \begin{cases} \sigma-\mu, & \sigma > \mu, \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

Here, $\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^{T}$ is the singular value decomposition of matrix $\mathbf{X}$, and $\Sigma = \mathrm{diag}(\{\sigma_i\}_{1\le i\le r})$ are the first r largest singular values of matrix $\mathbf{X}$.

(2) For $\mathbf{Z}_2$ subproblem, it can be translated to

$$\min_{\mathrm{Z}_2} \lambda_1 \left\|\mathbf{Z}_2\right\|_{\mathrm{W}_{\mathrm{T}},1} + \frac{\beta}{2}\left\|\mathbf{Z}_2 - \mathbf{T} + \frac{1}{\beta}\mathbf{Y}_2\right\|_F^2. \tag{14}$$

The optimization problem (14) is convex and can be solved by using the well-known soft thresholding operator:

$$\mathbf{Z}_2 = \mathrm{softTh}_{\lambda_1 \mathrm{W}_{\mathrm{T}}/\beta}\left(\mathbf{T} - \frac{1}{\beta}\mathbf{Y}_2\right), \tag{15}$$

where

$$\mathrm{softTh}_{\varepsilon}(\mathbf{X}) = \mathrm{sign}(X)\max\{0, |X| - \varepsilon\}, \tag{16}$$

where $|\cdot|$ denotes the absolute value, and all operations are implemented by componentwise. For $t \in \mathbb{R}$, the signum function $\mathrm{sign}(t)$ is defined as

$$\mathrm{sign}(t) = \begin{cases} +1, & t > 0, \\ 0, & t = 0, \\ -1, & t < 0. \end{cases} \tag{17}$$

For matrix $\mathbf{X}$, we let $\mathrm{sign}(\mathbf{X})$ be defined by componentwise.

(3) For $\mathbf{Z}_3$ subproblem, it can be translated to

$$\min_{\mathrm{Z}_3} \lambda_2 \left\|\mathbf{Z}_3\right\|_1 + \frac{\beta}{2}\left\|\mathbf{Z}_3 - \nabla\mathbf{T} + \frac{1}{\beta}\mathbf{Y}_3\right\|_F^2. \tag{18}$$

Optimization problem (18) can also be solved by using the soft thresholding operator. Because $\nabla\mathbf{T} = [\nabla_x\mathbf{X}, \nabla_y\mathbf{X}]^{T}$, equation (18) can be solved by two equations:

$$\mathbf{Z}_{3x} = \mathrm{softTh}_{\lambda_2/\beta}\left(\nabla_x\mathbf{T} - \frac{1}{\beta}\mathbf{Y}_3\right), \tag{19}$$

$$\mathbf{Z}_{3y} = \mathrm{softTh}_{\lambda_2/\beta}\left(\nabla_y\mathbf{T} - \frac{1}{\beta}\mathbf{Y}_3\right), \tag{20}$$

where ∇_x and ∇_y are the horizontal and vertical first-order forward finite difference operators, respectively.

(4) For $\mathbf{B}$ subproblem, it can be translated to

$$\min_{\mathrm{B}} \left\|\mathbf{Z}_1 - \mathbf{B} + \frac{1}{\beta}\mathbf{Y}_1\right\|_F^2 + \left\|\mathbf{D} - \mathbf{T} - \mathbf{B} + \frac{1}{\beta}\mathbf{Y}_4\right\|_F^2. \tag{21}$$

It is a quadratic programming problem and has a analytical solution:

$$\mathbf{B} = \frac{1}{2}\left(\mathbf{Z}_1 + \mathbf{D} - \mathbf{T} + \frac{1}{\beta}\left(\mathbf{Y}_1 + \mathbf{Y}_4\right)\right). \tag{22}$$

(5) For $\mathbf{T}$ subproblem, it can be translated to

$$\min_{\mathrm{T}} \left\|\mathbf{Z}_2 - \mathbf{T} + \frac{1}{\beta}\mathbf{Y}_2\right\|_F^2 + \left\|\nabla\mathbf{T} - \left(\mathbf{Z}_3 + \frac{1}{\beta}\mathbf{Y}_3\right)\right\|_F^2 + \left\|\mathbf{T} - \left(\mathbf{D} - \mathbf{B} + \frac{1}{\beta}\mathbf{Y}_4\right)\right\|_F^2. \tag{23}$$

Optimization problem (23) is also a quadratic programming problem. The equivalent liner system of the model (23) can be expressed as

$$\left(2\mathbf{I} + \nabla^{T}\nabla\right)\mathbf{T} = \mathbf{Z}_2 + \nabla^{T}\mathbf{Z}_3 + \frac{1}{\beta}\left(\mathbf{Y}_2 + \nabla^{T}\mathbf{Y}_3 + \mathbf{Y}_4\right) + \mathbf{D} - \mathbf{B}, \tag{24}$$

where the superscript T denotes the transpose operation.

Problem (23) has a closed solution and can be solved readily using the fast Fourier transform (FFT):

$$\mathbf{T} = \mathscr{F}^{-1}\left\{\frac{\mathscr{F}\left(\beta\zeta_1 + \nabla^{T}\zeta_2 + \zeta_3\right)}{2\beta + \beta\mathscr{F}\left(\nabla^{T}\nabla\right)}\right\}, \tag{25}$$

where $\mathscr{F}^{-1}$ and $\mathscr{F}$ denote the inverse FFT and FFT operators, respectively, and $\boldsymbol{\zeta}_1 = \mathbf{Z}_2 + \mathbf{D} + \mathbf{B}$, $\boldsymbol{\zeta}_2 = \beta\mathbf{Z}_3 + \mathbf{Y}_3$, and $\boldsymbol{\zeta}_3 = \mathbf{Y}_2 + \mathbf{Y}_4$.

(6) Finally, we update the Lagrangian multiplier variables $\mathbf{Y}_1$, $\mathbf{Y}_2$, $\mathbf{Y}_{3x}$, $\mathbf{Y}_{3y}$, and $\mathbf{Y}_4$ in the following manner:

$$\begin{cases} \mathbf{Y}_1 = \mathbf{Y}_1 + \beta(\mathbf{Z}_1 - \mathbf{B}), \\ \mathbf{Y}_2 = \mathbf{Y}_2 + \beta(\mathbf{Z}_2 - \mathbf{T}), \\ \mathbf{Y}_{3x} = \mathbf{Y}_{3x} + \beta(\mathbf{Z}_{3x} - \nabla_x \mathbf{T}), \\ \mathbf{Y}_{3y} = \mathbf{Y}_{3y} + \beta(\mathbf{Z}_{3y} - \nabla_y \mathbf{T}), \\ \mathbf{Y}_4 = \mathbf{Y}_4 + \beta(\mathbf{D} - \mathbf{T} - \mathbf{B}). \end{cases} \tag{26}$$

Taking into account the above equations, we obtain the complete iteration used in our algorithm for the target and background separation, then the optimization algorithm is shown in Algorithm 1.

3.3. Detection Procedure. The diagram of the proposed detection approach is given in Figure 1. The detailed detection procedure is described as follows:

(1) The patch-image $\mathbf{D} \in \mathbb{R}^{a^2 \times t}$ (we set $m = a^2, n = t$) is constructed from the original image $f_{\mathrm{D}} \in \mathbb{R}^{M \times N}$ by sliding the window of size $a \times a$ from left to right and top to bottom with a fixed step size, where t is the number of window slips.

(2) The target patch-image $\mathbf{T}$ and the background patch-image $\mathbf{B}$ are obtained from the patch-image $\mathbf{D}$ by applying Algorithm 1.

(3) The target image f_{T} and background image f_{B} are reconstructed from the patch-image $\mathbf{T}$ and $\mathbf{B}$, respectively.

(4) The final small target detection result is obtained from the target image f_{T} by the adaptive threshold t_{up}, which is defined as

$$t_{\mathrm{up}} = \mu + k\sigma, \tag{27}$$

where μ and σ are the mean and standard deviation of the target image f_{T}, respectively. k is constant determined experimentally. In our experiments, we empirically set the parameter ranging as $[1, 10]$. A pixel at (x, y) can be segmented as the target pixel if $f_{\mathrm{T}}(x, y) > t_{\mathrm{up}}$; otherwise, it is a background pixel.

4. Experimental Results and Analysis

In this section, we perform the experiments on simulated and real infrared images to validate the performance of the proposed method. We firstly introduce the experimental setup, including the dataset, the baseline methods, and the evaluation metric indexes. Secondly, we analyze how the key parameters influence the performance of the proposed method. Then, we compare the proposed method with the baseline methods. Finally, an analysis is made of the computational complexity of the proposed method.

4.1. Experimental Settings

4.1.1. Datasets

(1) Real Datasets. Firstly, we exploit six real infrared images, as shown in Figure 2, to validate the background suppression and target detection ability of the proposed method. The backgrounds in these images are varying from uniform to complex. The size of the targets is small, and some of the targets are also very dim. The specific description of the selected images is shown in Table 2.

(2) Simulated Datasets. The synthesized images f_{D} are obtained by embedding four simulated targets f_{T} in the background image f_{B}. The synthesized image can be generated by

$$f_{\mathrm{D}}(x, y) = \begin{cases} \max\left(r f_{\mathrm{T}}(x - x_0, y - y_0), f_{\mathrm{B}}(x, y)\right), & x \in (1 + x_0, n + x_0), y \in (1 + y_0, m + y_0), \\ f_{\mathrm{B}}(x, y), & \text{otherwise.} \end{cases} \tag{28}$$

The simulated target f_{T} is obtained by resizing the target with size of $M \times N$ by using the bicubic interpolation. (x_0, y_0) is the left upper corner coordinates of the target embedding, which is randomly generated. r is a randomly generated value from $[h, 255]$, where h is the maximum pixel value of the background image. Finally, a Gaussian filter is used to blur the synthesized image. We generate 100 frames of simulation sequences for each of the six background images. In each frame, four simulated small targets are randomly embedded. Then, we select the representative image of each sequence as shown in Figure 3.

4.1.2. Evaluation Metrics. To qualitatively evaluate the performance of the proposed method, we introduce several numerical metrics. Next, we define these metrics.

The signal-to-clutter ration (SCR) indicates the difficulty of target detection in a local area, and the computation formula can be written as

$$\mathrm{SCR} = \frac{\left|\mu_{\mathrm{t}} - \mu_{\mathrm{b}}\right|}{\sigma_{\mathrm{b}}}, \tag{29}$$

where μ_{t} and μ_{b} represent the average pixel value of the target region and the neighboring area around the target, respectively. σ_{b} is the standard deviation of the pixel value in the neighboring area around the target. As shown in Figure 4, we assume the size of the small target is $a \times b$, and then the size of its background region is $(a + 2d) \times (b + 2d)$, where d is the pixel width of neighboring area. In our experiment, we set $d = 20$ pixels.

(1) **Input**: infrared patch-image matrix $\mathbf{D} \in \mathbb{R}^{m\times n}$, λ_1, λ_2, C, β.
(2) **Output**: infrared patch-image matrix $\mathbf{B}, \mathbf{T}$.
(3) **Initialization:** $\mathbf{B}^0 = \text{zeros}(m,n), \mathbf{T}^0 = \text{zeros}(m,n), \mathbf{Y}_1^0 = \mathbf{Y}_2 = \mathbf{Y}_{3x}^0 = \mathbf{Y}_{3y}^0 = \mathbf{Y}_4^0 = \text{zeros}(m,n), \beta_{\max} = 10^7, k = 0.$
(4) **while** not converged do **do**
(5) **% Updating procedure for** $\mathbf{Z}_1^{k+1}$
(6) $\mathbf{Z}_1^{k+1} = \text{SVT}_{\mathbf{w}_B^k/\beta^k}(\mathbf{B}^k - (1/\beta^k)\mathbf{Y}_1^k)$;
(7) **% Updating procedure for** $\mathbf{Z}_2^{k+1}$
(8) $\mathbf{Z}_2^{k+1} = \text{softTh}_{(\lambda_1 \mathbf{w}_T^k/\beta^k)}(\mathbf{T}^k - (1/\beta^k)\mathbf{Y}_2^k)$;
(9) **% Updating procedure for** $\mathbf{Z}_3^{k+1}$
(10) $\mathbf{Z}_{3x}^{k+1} = \text{softTh}_{\lambda_2/\beta^k}(\nabla_x \mathbf{T}^k - (1/\beta^k)\mathbf{Y}_3^k)$; $\mathbf{Z}_{3y}^{k+1} = \text{softTh}_{\lambda_2/\beta^k}(\nabla_y \mathbf{T}^k - (1/\beta^k)\mathbf{Y}_3^k)$;
(11) **% Updating procedure for** $\mathbf{B}^{k+1}$
(12) $\mathbf{B}^{k+1} = (1/2)(\mathbf{Z}_1^{k+1} + \mathbf{D} - \mathbf{T}^k + (1/\beta^k)(\mathbf{Y}_1^k + \mathbf{Y}_4^k))$;
(13) **% Updating procedure for** $\mathbf{T}^{k+1}$
(14) $\mathbf{T}^{k+1} = \mathcal{F}^{-1}\left\{(\mathcal{F}(\beta\boldsymbol{\zeta}_1^{k+1} + \nabla^T\boldsymbol{\zeta}_2^{k+1} + \boldsymbol{\zeta}_3^k))/(2\beta + \beta\mathcal{F}(\nabla^T\nabla))\right\}$; $\boldsymbol{\zeta}_1^{k+1} = \mathbf{Z}_2^{k+1} + \mathbf{D} + \mathbf{B}^{k+1}$, $\boldsymbol{\zeta}_2^{k+1} = \beta\mathbf{Z}_3^{k+1} + \mathbf{Y}_3^k$, $\boldsymbol{\zeta}_3^k = \mathbf{Y}_2^k + \mathbf{Y}_4^k$;
(15) Compute $\mathbf{w}_B^{k+1}$ and $\mathbf{W}_T^{k+1}$ using (8) and (9), respectively;
(16) **% Update Lagrangian multipliers** $\mathbf{Y}_1^{k+1}$, $\mathbf{Y}_2^{k+1}$, $\mathbf{Y}_{3x}^{k+1}$, $\mathbf{Y}_{3y}^{k+1}$, $\mathbf{Y}_4^{k+1}$
(17) $\mathbf{Y}_1^{k+1} = \mathbf{Y}_1^k + \beta^k(\mathbf{Z}_1^{k+1} - \mathbf{B}^{k+1})$; $\mathbf{Y}_2^{k+1} = \mathbf{Y}_2^k + \beta^k(\mathbf{Z}_2^{k+1} - \mathbf{T}^{k+1})$; $\mathbf{Y}_{3x}^{k+1} = \mathbf{Y}_{3x}^k + \beta^k(\mathbf{Z}_{3x}^{k+1} - \nabla_x\mathbf{T}^{k+1})$; $\mathbf{Y}_{3y}^{k+1} = \mathbf{Y}_{3y}^k + \beta^k(\mathbf{Z}_{3y}^{k+1} - \nabla_y\mathbf{T}^{k+1})$; $\mathbf{Y}_4^{k+1} = \mathbf{Y}_4^k + \beta^k(\mathbf{D} - \mathbf{T}^{k+1} - \mathbf{B}^{k+1})$;
(18) **% Updating procedure for** β^{k+1}
(19) $\beta^{k+1} = \min(\rho * \beta^k, \beta_{\max})$;
(20) **% Check the convergence conditions**
(21) $\|\mathbf{D} - \mathbf{T}^{k+1} - \mathbf{B}^{k+1}\|_F / \mathbf{D}_F < \varepsilon$;
(22) Update $k := k + 1$.
(23) **end While**

Algorithm 1: Target and background separation solver via ADMM.

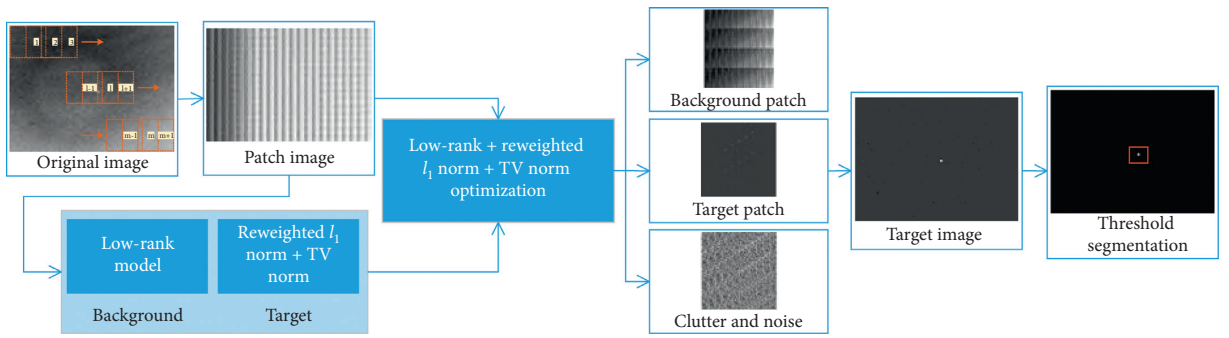


Figure 1: The various steps of the proposed algorithm.

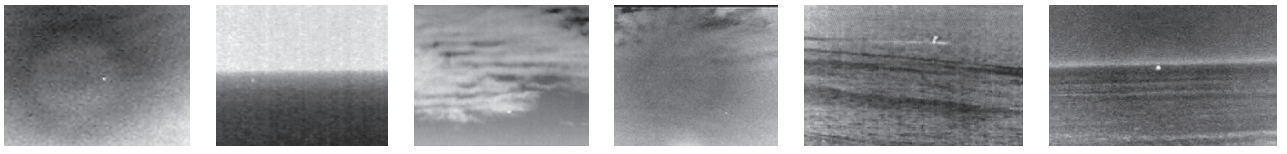


Figure 2: Six real infrared images.

The signal-to-clutter ratio gain (SCRG) is a widely used numerical metric to evaluate the performance of small target detection, which can be defined as the ratio of SCR in the original image to that in the processed image:

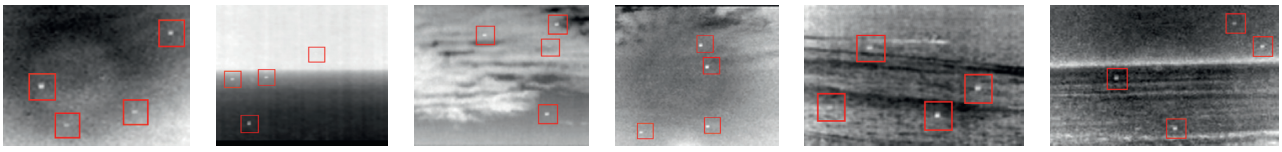
$$\text{SCRG} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}}}. \tag{30}$$

The background suppression factor (BSF) can also be used to describe how difficult the small target detection is. It can be defined as follows:

$$\text{BSF} = \frac{\sigma_{\text{in}}}{\sigma_{\text{out}}}, \tag{31}$$

Table 2: The specific description of six real images.

Image	Target type	Image	Target characteristic	Background characteristic
Image (a)	A helicopter	200×150	Very small	Uniform backgrounds to heavy noise
Image (b)	A ship	280×228	Low SCR value, small and dim buried in background	Blurred sea-sky background and almost unchanged
Image (c)	An airplane	250×200	Low SCR value	Heavy cloudy-sky background clutters
Image (d)	A helicopter	281×240	A small size, low SCR value	Heavy cloudy clutters
Image (e)	A ship	220×140	A small size, low SCR value	Heavy sea clutters background almost unchanged
Image (f)	A ship	320×196	Low SCR value, close to the sea-sky-line	Heavy sea clutters of sea-sky-line



(a) (b) (c) (d) (e) (f)

Figure 3: Synthesized multiple target images of six sequences.

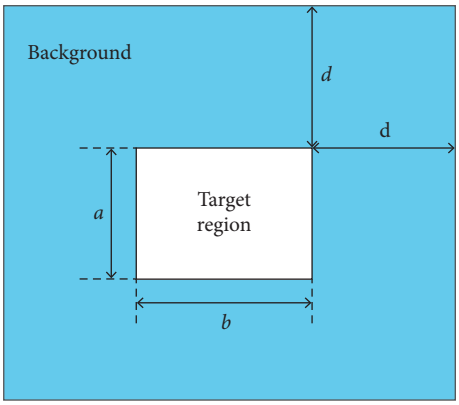


Figure 4: The external rectangle and neighboring background rectangle of a small target.

where σ_{in} and σ_{out} denote the standard deviations of pixel value in the original and processed images except the target region, respectively. In general, the higher the SCR and BSF values of a small target image are, the easier it is for the target to be detected.

Since the calculation of SCRG and BSF involves the standard deviation of the background area, both indexes can obtain a zero value if the background is too clean. To handle this problem, we adopt the contrast gain (CG) to assess the ability of clutters and noise suppression and small target enhancement. CG is defined as follows:

$$\mathrm{CG} = \frac{\mathrm{CON}_{\mathrm{out}}}{\mathrm{CON}_{\mathrm{in}}}. \tag{32}$$

In the above equation, $\mathrm{CON}_{\mathrm{in}}$ and $\mathrm{CON}_{\mathrm{out}}$ denote the contrasts (CON) of the original and processed images, respectively, and CONs can be calculated as follows:

$$\mathrm{CON} = \left|\mu_t - \mu_b\right|, \tag{33}$$

where μ_t and μ_b are the same as those in equation (29).

The detection probability (P_d) and false alarm rate (F_a) are the two frequently used metrics for evaluating the detection performance, which are defined as follows:

$$P_d = \frac{\#\text{true targets detected}}{\#\text{total bmtrue targets}},$$
$$F_a = \frac{\#\text{false pixels detected}}{\#\text{total pixels in images}}. \tag{34}$$

If a method has a good detection performance, then it has high (P_d) and low (F_a) values. The detected result is correct if it simultaneously meets two conditions: (i) the detected result and a ground truth have overlap pixels; (ii) the pixel distance between centers of the detection result and the ground truth is less than a threshold (3 pixels in our research). The receiver operating characteristic (ROC) curve shows the dynamic relationship between the detection probability (P_d) and false alarm rate (F_a) quantitatively.

4.1.3. Baseline Methods and Parameter Setting. We compare the proposed method with six baseline methods: Top-hat [7], Max-mean [5], Max-median [5], IPI [22], TV-PCP [23], and MLMC [16]. Table 3 presents the specific parameter setting for each method.

The proposed model has several parameters which influence the performance of the proposed method. The regularization parameters λ_1 and λ_2 control the trade-off among three components in the proposed model (7). Specifically, the two parameters control the sensitivity level of the model to the nuclear norm, reweighted ℓ_1 norm, and TV term. We set the parameter λ_1 to be $1/\sqrt{\max(m,n)}$ the same as in PCP problem [38]. For λ_2, a large value of λ_2 means that the influence of TV regularization becomes more significant and that the clutters and noise can be suppressed remarkably. A small value of λ_2 may cause more clutters and noise residual in the target image. In our implementation, we

Table 3: Detailed parameter setting for the test methods.

Methods	Acronyms	Parameter setting
Top-hat method	Top-hat	Structure size: 12×12
Max-mean filter	Max-mean	Filter size: 15×15
Max-median filter	Max-median	Filter size: 15×15
Infrared patch-image model	IPI	Patch size: 50×50, sliding step: 10, $\lambda = (1/\sqrt{\max(m,n)})$, $\varepsilon = 10^{-6}$
Total variation regularization and principal component pursuit	TV-PCP	Patch size: 50×50, sliding step: 14 maxIter = 250, $\varepsilon = 5 \times 10^{-6}$
MLMC		Scale number $K = 4$
Ours	Ours	Patch size: 50×50, sliding step: 10, $\varepsilon = 10^{-7}$, $\lambda_1 = (1/\sqrt{(\max(m,n))})$, $\lambda_2 = 0.006$, maxIter = 500

empirically set the parameter range as $\lambda_2 \in [1/\sqrt{\max(m,n)}, 1.5/\sqrt{\max(m,n)}]$. We find that the proposed method can work very well in such a range. β is a penalization parameter and is updated by $\beta^{k+1} = \rho\beta^k$, where k is the number of iterations. $\rho = 1.25$ and the initial $\beta^0 = 1.25/\|\mathbf{D}\|$, where $\|\mathbf{D}\|$ denotes the spectral norm of matrix $\mathbf{D}$. The stopping criterion of the proposed method for the iteration is "normalized step difference energy" (NSDE): NSDE $= \|\mathbf{D} - \mathbf{T}^{k+1} - \mathbf{B}^{k+1}\|_F / \|\mathbf{D}\|_F < \varepsilon$ for each iteration, where $\varepsilon = 10^{-7}$. We set the maximum number of iterations to maxIter = 500.

Both the proposed method and the methods for comparison are implemented using MATLAB 2018a running on a computer server with an Intel Xeon E5-2620 2.10 GHz CPU, 142 GB of RAM, and Linux system.

4.2. Parameter Analysis. The performance of the proposed method is influenced by several important parameters, such as the patch size, sliding step, and trade-off parameter C. In this subsection, we analyze the influence of these parameters on the performance of our method and discuss the appropriate selection of these parameters. The ROC curves corresponding to the patch size and sliding step on sequences 1–6 are shown in Figures 5 and 6, respectively.

4.2.1. Patch Size. Patch size is a very important parameter in our experiments. A small patch size usually increases the correlation in a local area and prevents the background and target parts from being well separated from the original image, thus a lower detection probability. A large patch size reduces the correlation between the nonlocal areas and causes a large background residue in the target, thus a larger false alarm rate. Large patch sizes also increase the computational complexity. In order to choose the appropriate patch size, we fix the sliding step as 10 and vary the patch size from 20 to 80 with an interval of 10. And then our method is tested in six simulated image sequences and the ROC curves are shown in Figure 5. We can observe from the results that the proposed method has a similar performance when the patch size is 50, 60, and 70. From Figures 5(a), 5(b), 5(d), and 5(f), we can see that a small patch size yields a low detection probability. In Figures 5(c) and 5(e), it is observed that a large patch size results in a high false alarm rate. In order to balance the detection probability and false alarm rate, we choose the patch size as 50.

4.2.2. Sliding Step. Sliding step is also a parameter that affects the size of the patch-image matrix. A large sliding step reduces the correlation between the nonlocal areas, leaving a large background component in the target, which causes higher false alarm rates. In order to get a suitable sliding step, we fix the patch size as 50×50 and vary the patch size from 2 to 16 with an interval of 2 and present the ROC curves in Figure 6. In Figures 6(b) and 6(f), it can be seen that a large sliding step brings about a high false alarm rate. In order to balance the detection probability and false alarm rate, we choose the sliding step of 10.

4.2.3. Trade-Off Parameter. C is a trade-off parameter introduced in the adaptive weight factor. A large value of C tends to yield better background suppression effect. A small value of C results in more background and clutters and noise residual in the small target image. For a better balance between the preservation of the small target pixels and the suppression of the background clutters and noise, the parameter C needs to be chosen carefully. Empirically, a reasonable C for an infrared image typically lies in the range [0.01, 0.3]. Figure 7 shows the processed results by using different values of C and we chose 0.1 as the final C value.

4.3. Comparison with the Baseline Methods. Table 4 presents three index values of seven methods for six real images. As shown in Table 4, our method consistently achieves the highest values for the three evaluating metrics as compared with other five methods. For each measure of SCR, BSF, and CG, a higher score implies better performance. High SCR, BSF, and CG values mean that the background clutters and noise in the processed images are suppressed better, thus easier detection of the small targets. In this table, it is shown that the value of the metric SCRG approaches infinity in some cases. It happens when the computation of the standard deviation in equation (30) is performed based on a clean local background area.

Figures 8 and 9 show the detected results of the seven methods for six real infrared images. We also give the close-up of the small target for each detected result. As shown in Figure 8 (a1)–(a8), the Top-hat method has the enhancement effect to some extent; however, the background is not suppressed well and thus this method suffers a high false alarm rate. It can be seen from Figure 8 (b1)–(b8) and

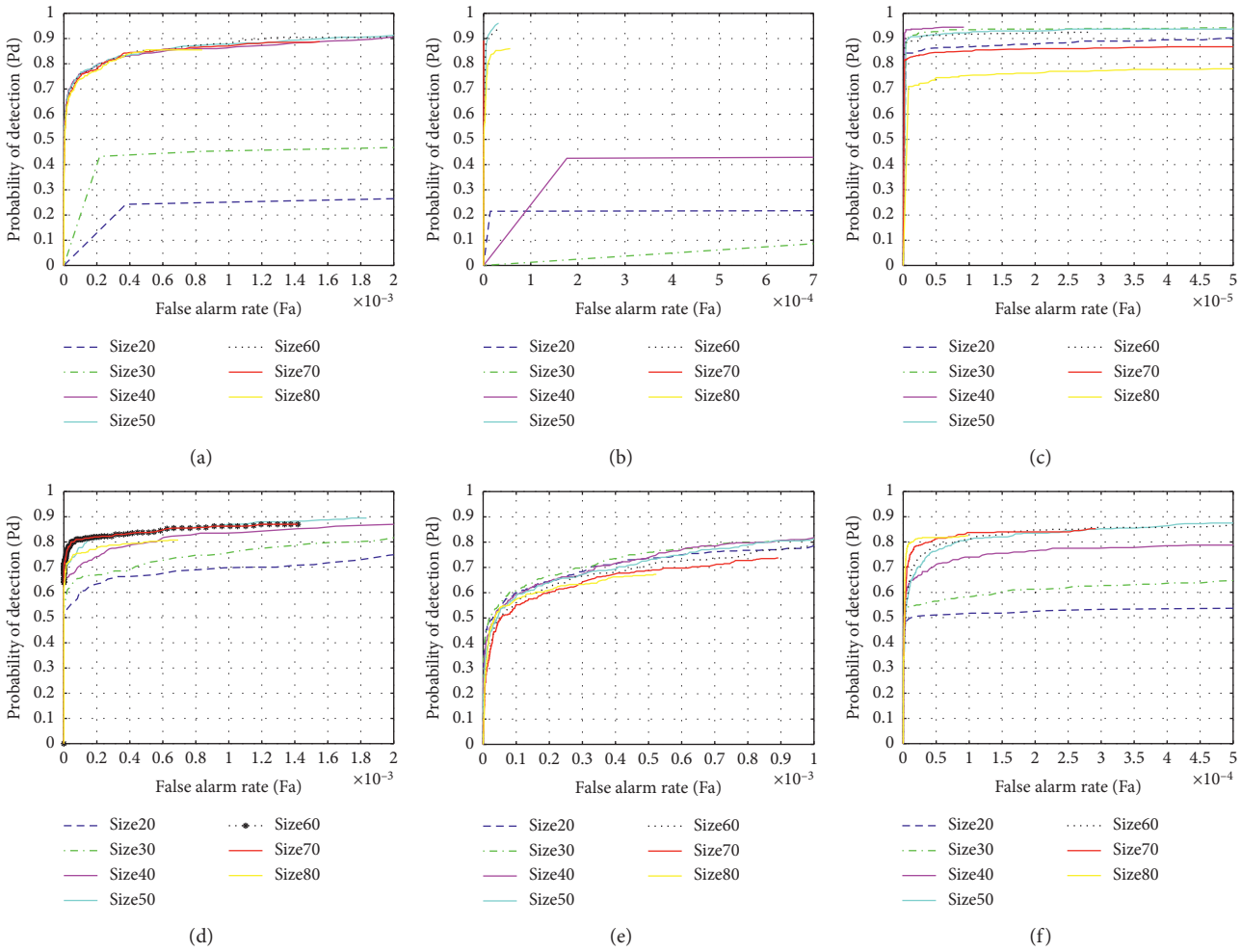


Figure 5: ROC curves for sequences 1–6 with respect to different patch sizes: (a) sequence 1; (b) sequence 2; (c) sequence 3; (d) sequence 4; (e) sequence 5; (f) sequence 6.

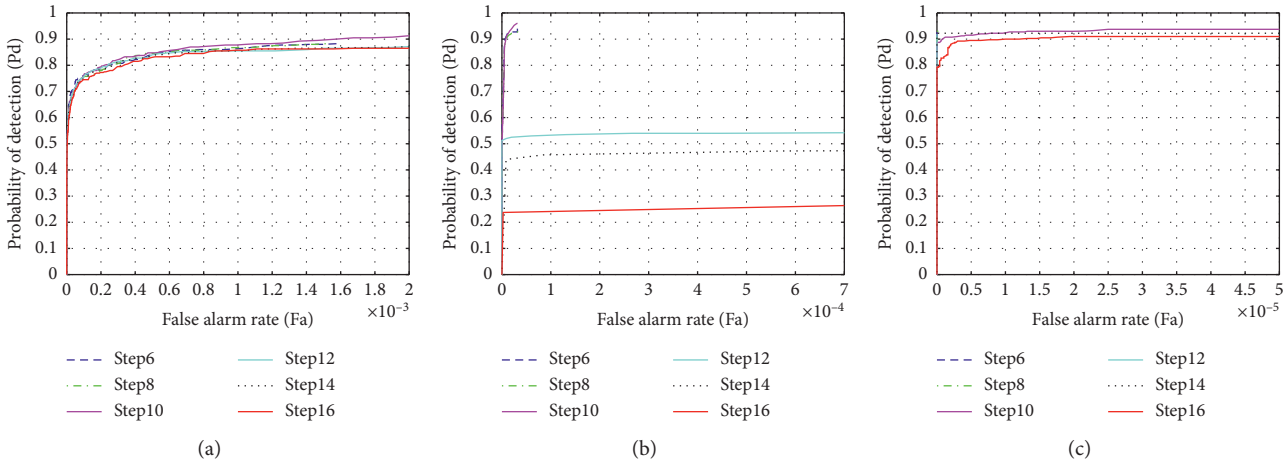


Figure 6: Continued.

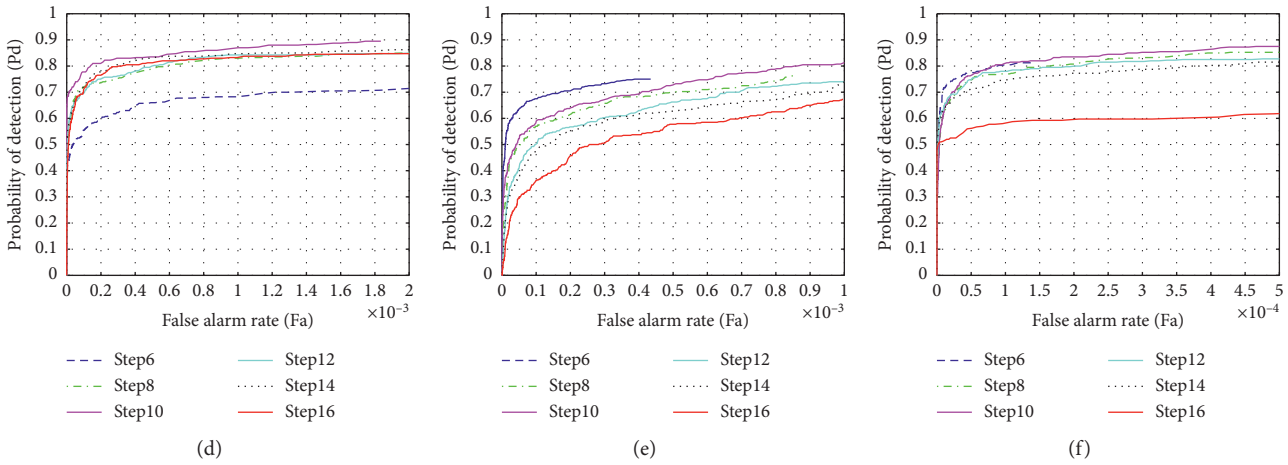


Figure 6: ROC curves for sequences 1–6 with respect to different sliding steps: (a) sequence 1; (b) sequence 2; (c) sequence 3; (d) sequence 4; (e) sequence 5; (f) sequence 6.

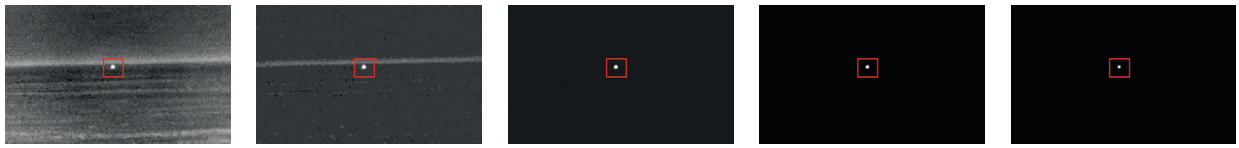


Figure 7: The detection results using different choices C: (a) $C = 0.001$; (b) $C = 0.01$; (c) $C = 0.1$; (d) $C = 0.3$; (e) $C = 0.5$.

Table 4: SCRG, BSF, and CG comparisons of seven methods using six test images.

Methods	Evaluation indexes	Figure 7(a)	Figure 7(b)	Figure 7(c)	Figure 7(d)	Figure 7(e)	Figure 7(f)
Top-hat	SCRG	0.7293	2.8136	1.2570	1.2127	1.0951	1.2586
	BSF	4.3850	12.4920	2.3549	2.5420	1.6498	1.5829
	CG	0.9145	1.0303	0.6520	0.9672	1.0954	0.8683
Max-mean	SCRG	3.1266	8.7743	9.1730	2.3488	2.0967	4.5015
	BSF	7.5985	5.5402	3.9490	4.3879	2.4770	2.4768
	CG	1.5546	3.1316	2.5962	1.6843	1.1479	1.1179
Max-median	SCRG	2.9499	9.7824	7.6414	3.2046	2.1385	4.4410
	BSF	9.0372	6.2069	4.2723	6.3041	2.7539	2.5966
	CG	1.5748	2.5862	1.6389	1.8157	1.2600	1.2297
IPI	SCRG	6.2432	**Inf**	36.9539	**Inf**	12.4616	41.3657
	BSF	44.0414	79.7198	14.9939	31.3549	32.6376	27.5367
	CG	1.6242	4.4338	2.3021	1.1791	2.8150	1.7741
TV-PCP	SCRG	10.7821	**Inf**	61.4484	2270.5115	16.0146	48.1482
	BSF	96.4571	64.2095	28.7416	57.4668	57.0914	30.2031
	CG	1.4729	3.6194	2.0528	1.0971	2.5456	1.7436
MLMC	SCRG	4.0888	25.9482	17.5429	8.1452	2.0800	11.6160
	BSF	97.8868	27.6753	10.6644	5.2934	12.3820	39.9455
	CG	1.0410	1.7478	1.7112	1.2009	1.1074	0.5745
Ours	SCRG	**15.3347**	**Inf**	109.9528	**Inf**	284.3204	168.5884
	BSF	133.3089	130.7386	31.8744	57.5505	61.5704	77.4739
	CG	1.6911	5.2929	2.5061	1.2011	3.7795	1.9793

Figure 8 (c1)–(c8) that the Max-mean and Max-median methods have a better ability to enhance the small target than the Top-hat method, but the background suppression for these two methods is not satisfactory. The background clutters and noise residuals are still obvious, and thus, it is still difficult for the dim targets to be detected. Figure 9 (d1)–(d8) shows that the IPI method achieves an obvious improvement in enhancing the small target and suppressing the background clutters and noise. But the background suppression is not enough: the target images

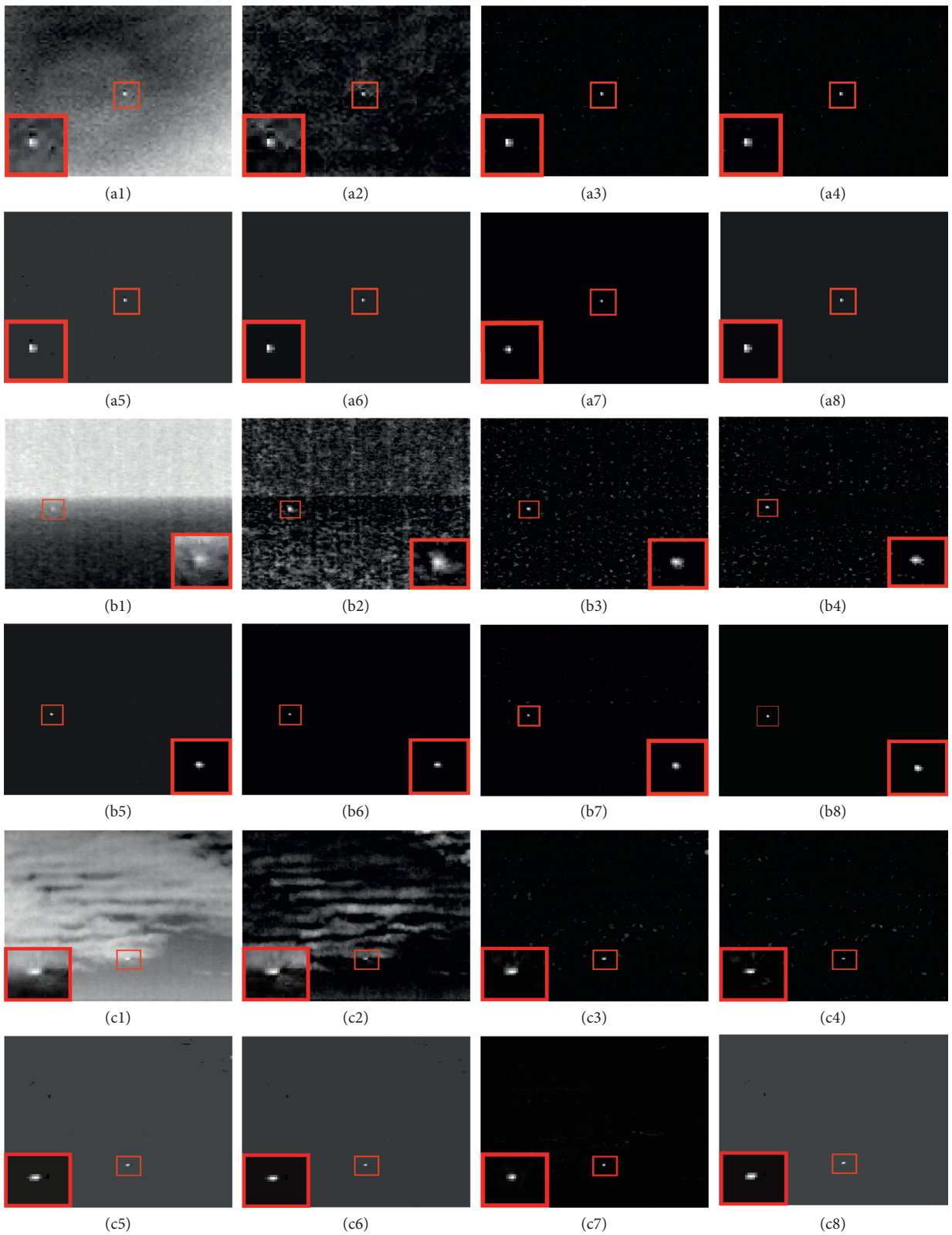


FIGURE 8: The detected results of the seven methods for three real infrared images and the close-up of the small target for each detected result. (a1)-(c1) are the real infrared images corresponding to Figures 2(a)–2(c). (a2)-(c2), (a3)-(c3), (a4)-(c4), (a5)-(c5), (a6)-(c6), (a7)-(c7), and (a8)-(c8) are the detected results obtained by using the Top-hat, Max-mean, Max-median, IPI, TV-PCP, MLMC, and proposed methods, respectively.

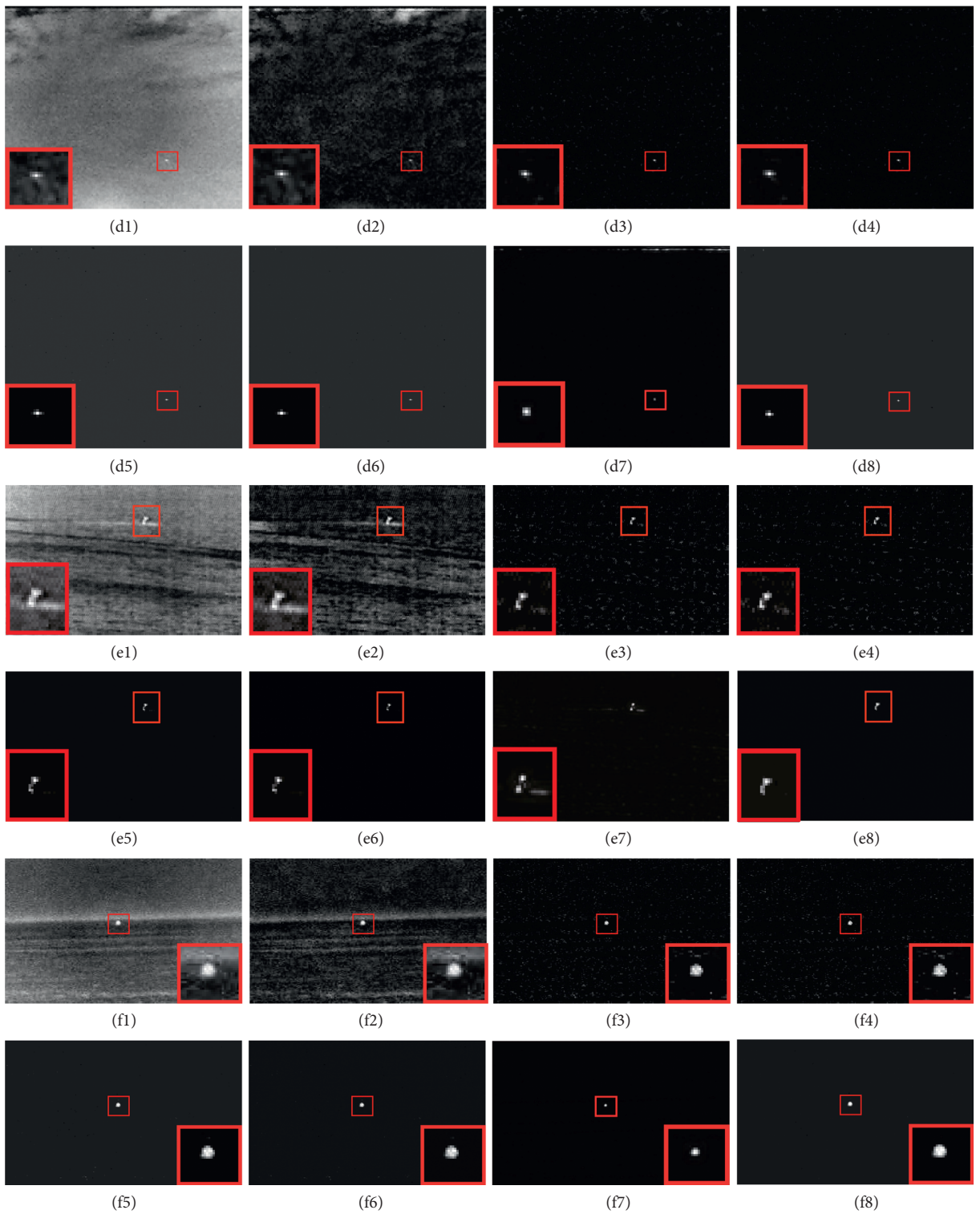


Figure 9: The detected results of the seven methods for three real infrared images and the close-up of the small target for each detected result. (d1)-(f1) are the real infrared images corresponding to Figures 2(d)–2(f). (d2)-(f2), (d3)-(f3), (d4)-(f4), (d5)-(f5), (d6)-(f6), (d7)-(f7), and (d8)-(f8) are the detected results obtained by using the Top-hat, Max-mean, Max-median, IPI, TV-PCP, MLMC, and proposed methods, respectively.

still have a bit of clutters and noise residuals. As shown in Figure 9 (e1)–(e8), the TV-PCP method gets a good background suppression and target enhancement performance under the relatively weak background clutters. However, there is still room for improvement in the strong background clutters. As shown in Figure 9 (f7), the detection result obtained with the TV-PCP method still has the sea clutters and noise residual. From Figure 9 (f1)–(f8), it can be seen that compared with the IPI method and the TV-PCP method, the proposed method has better

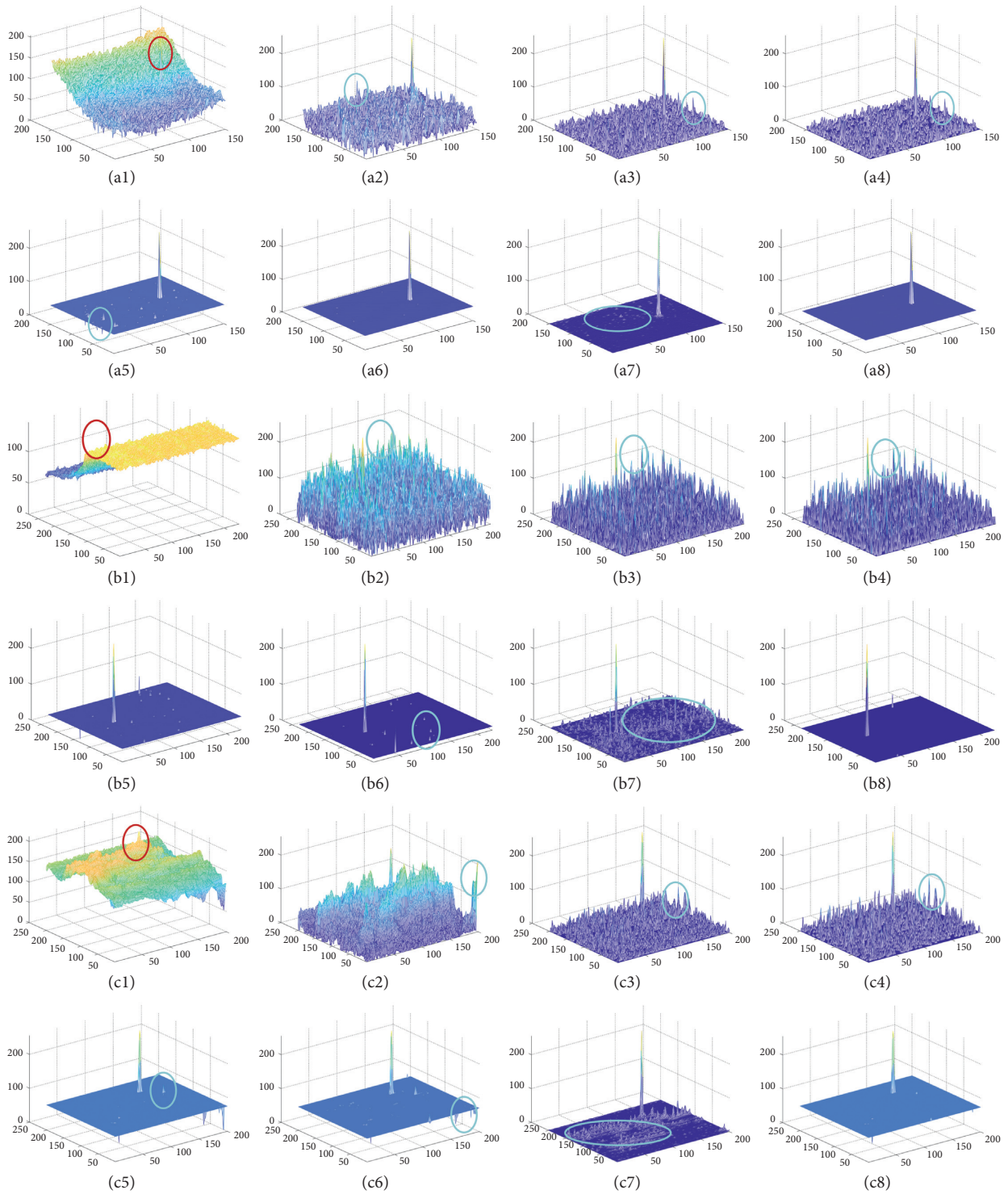


FIGURE 10: 3D gray maps of the detected results of seven methods for three real infrared images. (a1)-(c1) are the 3D gray maps of real infrared images corresponding to Figures 7(a)–7(c). (a2)-(c2), (a3)-(c3), (a4)-(c4), (a5)-(c5), (a6)-(c6), (a7)-(c7), and (a8)-(c8) are the 3D gray maps obtained by using the Top-hat, Max-mean, Max-median, IPI, TV-PCP, MLMC, and proposed methods, respectively.

background and noise suppression ability when the target pixels are similar in the detected results of the three methods, and thus, the target can be easily detected in the subsequent segmentation. The reason may be that the proposed method introduces the total variation constraint to small target component and that the background clutters and noise in the small target part are suppressed better. Due to the introduction of the weighted ℓ_1 norm, some non-target points are also suppressed. In a word, the combination of the TV regularization and the weighted ℓ_1 norm is

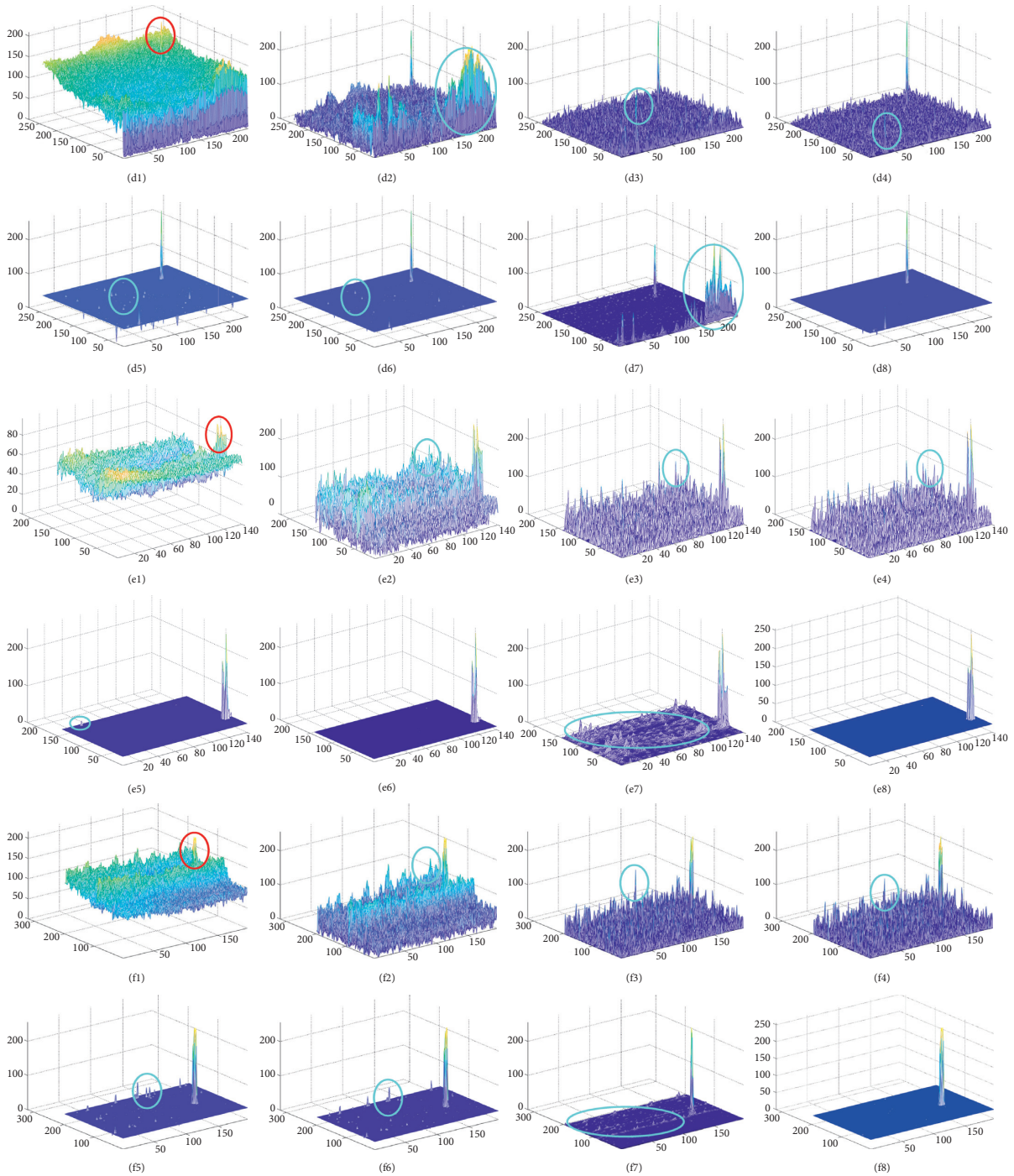


FIGURE 11: 3D gray maps of the detected results of seven methods for three real infrared images. (d1)-(f1) are the 3D gray maps of real infrared images corresponding to Figures 7(d)–7(f). (d2)-(f2), (d3)-(f3), (d4)-(f4), (d5)-(f5), (d6)-(f6), (d7)-(f7), and (e8)-(f8) are the 3D gray maps obtained by using the Top-hat, Max-mean, Max-median, IPI, TV-PCP, MLMC, and proposed methods, respectively.

an effective approach in suppressing the clutters and noise while enhancing the small target.

In order to intuitively present the ability of the clutter suppression and target enhancement, the 3D gray maps of the detected results of the seven methods for six real infrared images are also given in Figures 10 and 11. We use the blue-green circle to mark the obvious background clutters residual generated by other methods. We can see from Figure 10 (a1)-(c1) and Figure 11 (d1)–(f1) that there exist the background components whose gray levels are close to or higher than those of the targets, which means that it is very challenging to detect dim and weak target. The detection

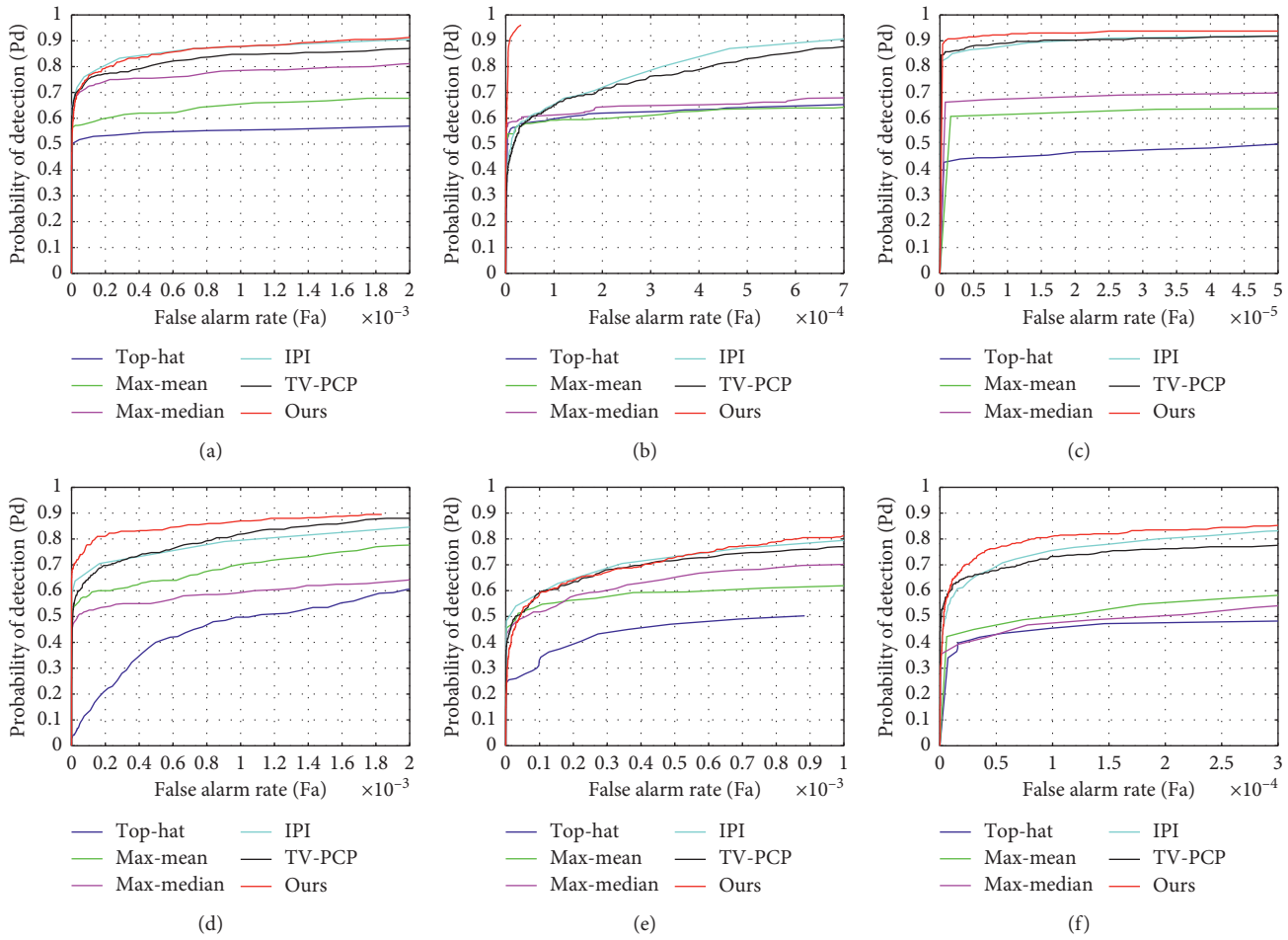


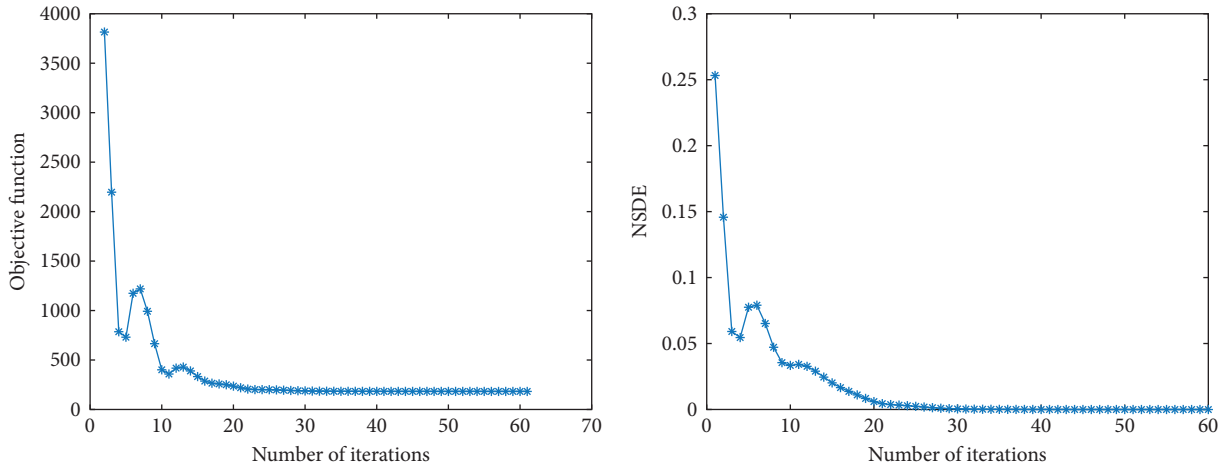
Figure 12: ROC curves of the detection results of six simulated sequences: (a) sequence 1; (b) sequence 2; (c) sequence 3; (d) sequence 4; (e) sequence 5; (f) sequence 6.

effect with the Top-hat, Max-mean, and Max-median methods as shown in Figures 10 and 11 is consistent with those in Figures 8 and 9. The IPI and TV-PCP methods present a relatively good ability in background suppression and target enhancement in most cases, but in the detection results of both methods many clutters and noise residuals still exist in the scene with strong clutters and noise. In general, the background clutters and random noise are the main obstructions to the separation of target and background. From the Figure 10 (a7), (b7), and (c7) and Figure 11 (d7), (e7), and (f7), we can observe that the MLMC method still exists many clutters and noise residuals. The detected results produced by our method in Figures 10 and 11 display a flatter background and more prominent target than those by other methods, which implies that our method is capable of removing the interferences with a consistent robustness in background suppression and target enhancement in different scenes by combining the TV regularization and reweighted ℓ_1 norm sparse constraint.

To further verify the performance of our method, we make a comparison of the ROC curves produced by six methods for the six simulated sequences in Figure 12. It is found that the Top-hat method is not satisfactory, for it has a low detection probability. Both the Max-mean and Max-median methods possess higher detection probability than the Top-hat method for the sequences except sequence 2. The IPI and TV-PCP methods have better performance than the above three baseline methods for all sequences. In general, the proposed method consistently achieves the highest detection probability in almost all cases, implying that the proposed method has a better detection performance than the baseline methods.

4.4. Convergence Rate of the Algorithm. We investigated the convergence behavior of the proposed algorithm. Figure 13 illustrates the evolutional curve of the objective function and NSDE versus the iterations for the real infrared image object detection in Figure 7(a). It can be observed clearly that the proposed method achieves good convergence and converges after about 20 iterations.

4.5. Computational Complexity and Computation Time of the Algorithm. In this subsection, we compare the computational costs of seven methods. The computational complexity and computation time corresponding to the six images in Figure 7 by the seven methods are shown in Table 5. Suppose the size of the original image is $M \times N$. The



(a) (b)

Figure 13: Evolution of the objective function and NSDE versus number of iterations for the proposed method.

Table 5: Comparison of computational complexity and computation time (in seconds) of seven methods.

Method	Complexity	Figure 7(a)	Figure 7(b)	Figure 7(c)	Figure 7(d)	Figure 7(e)	Figure 7(f)
Top-hat	$O(MNL^2 \log L^2)$	0.273	0.280	0.262	0.275	0.264	0.239
Max-mean	$O(MNL^2)$	1.290	2.211	2.186	2.273	1.319	11.553
Max-median	$O(MNL^2 \log L^2)$	1.251	2.408	1.934	2.713	1.264	11.948
IPI	$O(mn^2)$	1.548	5.825	4.138	5.311	1.673	22.848
TV-PCP	$O(mn^2)$	15.154	40.589	31.039	45.586	12.617	201.407
MLMC	$O(MNL^3)$	0.034	0.028	0.026	0.029	0.025	0.028
Ours	$O(mn^2)$	6.319	23.242	11.695	18.743	6.498	88.440

size of the patch-image constructed from the original image is $m \times n$ (for simplicity, we assume $m > n$). The computational complexity of the Top-hat method is $O(MNL^2 \log L^2)$, where L^2 is the size of the structure element. For the Max-median method, the complexity is dominated by the median filtering, its computational cost being $O(L^2 \log L^2)$. Considering the image size, the final cost of Max-median is $O(MNL^2 \log L^2)$. As for the Max-mean method, the total cost is $O(MNL^2)$. For the TV-PCP method, the total cost of each iteration is $O(mn^2)$ operations. The proposed model contains an optimization problem: the objective function includes the nuclear norm, ℓ_1 norm, and total variation norm and is solved by ADMM. For the proposed method, the computational complexities of updating the variables $\mathbf{Z}_1$, $\mathbf{Z}_2$, $\mathbf{Z}_{3x}$, $\mathbf{Z}_{3y}$, $\mathbf{B}$, and $\mathbf{T}$ in (12), (15), (19), (20), (22), and (24) for each iteration are $O(mn^2)$, $O(mn)$, $O(mn)$, $O(mn)$, $O(mn)$, and $O(mn)$, respectively. Therefore, the total cost of each iteration for the proposed method is $O(mn^2)$ operations.

5. Conclusion

For better detection of infrared small targets with complex backgrounds clutters and noise, a novel method integrating the TV regularization with reweighted ℓ_1 norm constraints is proposed to describe the sparse small target component more accurately. Furthermore, a low-rank constraint item is adopted to depict the background component. The TV regularization is incorporated into the model for suppressing the residual background clutters and noise and a sparsity reweighted ℓ_1 norm on the target is employed to enhance the small target. Finally, the task of small target detection is transformed into an optimization problem, which can be efficiently solved by invoking ADMM. The experiments on both synthesized and real small target images demonstrate that our method is more efficient and effective in detecting small targets under complex background clutters and noise than the conventional baseline methods. The computational time of the proposed method can be improved with efficient C/C++ implementation.

The proposed method also has some limitations. The regularization parameter λ_2 is not updated adaptively and need to be manually tuned. Although the parameter setting in the proposed method can be empirically determined in a range, it is still more preferred for some applications to have a completely automatic parameter estimation method.

Data Availability

The data used to support the findings of this study available from the first author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The authors are grateful to Dr. Ming Zhang for improving the grammar. This work was supported by the National Natural Science Foundation of China under Grant 41501371, the Fundamental Research Funds for the Central Universities under Grants JB170210 and XJS190302, and the Research Project on Teaching Reform of Higher Education (20901190020).

References

[1] X. Jin, J. Li, T. J. Schmit, and M. D. Goldberg, "Evaluation of radiative transfer models in atmospheric profiling with broadband infrared radiance measurements," *International Journal of Remote Sensing*, vol. 32, no. 3, pp. 863–874, 2011.

[2] M. T. Eismann, "Emerging research directions in air-to-ground target detection and discrimination," in *Proceedings of the Infrared Technology and Applications XXXI*, vol. 5783, pp. 674–683, Orlando, FL, USA, May 2005.

[3] L. Dong, D. Ma, G. Qin, T. Zhang, and W. Xu, "Infrared target detection in backlighting maritime environment based on visual attention model," *Infrared Physics & Technology*, vol. 99, pp. 193–200, 2019.

[4] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Infrared small-target detection using multiscale gray difference weighted image entropy," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 1, pp. 60–72, 2016.

[5] S. D. Deshpande, M. Er, V. Ronda, and P. Chan, "Max-mean and max-median filters for detection of small-targets," in *Proceedings of the SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, vol. 3809, pp. 74–83, San Diego, CA, USA, 1999.

[6] Y.-S. Moon, T. Zhang, Z. Zuo, and Z. Zuo, "Detection of sea surface small targets in infrared images based on multilevel filter and minimum risk bayes test," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 14, no. 7, pp. 907–918, 2000.

[7] X. Bai and F. Zhou, "Analysis of new top-hat transformation and the application for infrared dim small target detection," *Pattern Recognition*, vol. 43, no. 6, pp. 2145–2156, 2010.

[8] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 574–581, 2014.

[9] J. Han, Y. Ma, J. Huang, X. Mei, and J. Ma, "An infrared small target detecting algorithm based on human visual system," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 3, pp. 452–456, 2016.

[10] L. Zhang, L. Peng, T. Zhang, S. Cao, and Z. Peng, "Infrared small target detection via non-convex rank approximation minimization joint l2,1 norm," *Remote Sensing*, vol. 10, no. 11, pp. 1–26, 2018.

[11] Y. Sun, J. Yang, Y. Long, Z. Shang, and W. An, "Infrared patch-tensor model with weighted tensor nuclear norm for small target detection in a single frame," *IEEE Access*, vol. 6, pp. 76140–76152, 2018.

[12] M. Wan, G. Gu, W. Qian et al., "Total variation regularization term-based low-rank and sparse matrix representation model for infrared moving target tracking," *Remote Sensing*, vol. 10, no. 4, p. 510, 2018.

[13] X. Fan, Z. Xu, J. Zhang, Y. Huang, and Z. Peng, "Infrared dim and small targets detection method based on local energy center of sequential image," *Mathematical Problems in Engineering*, vol. 2017, Article ID 4572147, 16 pages, 2017.

[14] Y. Hu, M. Xiao, K. Zhang, and X. Wang, "Aerial infrared target tracking in complex background based on combined tracking and detecting," *Mathematical Problems in Engineering*, vol. 2019, Article ID 2419579, 17 pages, 2019.

[15] Z.-J. Feng, X.-L. Zhang, L.-Y. Yuan, and J.-N. Wang, "Infrared target detection and location for visual surveillance using fusion scheme of visible and infrared images," *Mathematical Problems in Engineering*, vol. 2013, Article ID 720979, 7 pages, 2013.

[16] Y. Shi, Y. Wei, H. Yao, D. Pan, and G. Xiao, "High-boost-based multiscale local contrast measure for infrared small target detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 1, pp. 33–37, 2018.

[17] C. Gao, L. Wang, Y. Xiao, Q. Zhao, and D. Meng, "Infrared small-dim target detection based on Markov random field guided noise modeling," *Pattern Recognition*, vol. 76, pp. 463–475, 2018.

[18] L. Dong, B. Wang, M. Zhao, and W. Xu, "Robust infrared maritime target detection based on visual attention and spatiotemporal filtering," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 3037–3050, 2017.

[19] P.-L. Shui, D.-C. Li, and S.-W. Xu, "Tri-feature-based detection of floating small targets in sea clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 2, pp. 1416–1430, 2014.

[20] Q. Zhao, D. Meng, Z. Xu, W. Zuo, and L. Zhang, "Robust principal component analysis with complex noise," in *Proceedings of the 31st International Conference on International Conference on Machine Learning-ICML'14*, vol. 32, pp. II–55–II–63, Bejing, China, June 2014.

[21] T. Bouwmans, S. Javed, H. Zhang, Z. Lin, and R. Otazo, "On the applications of robust PCA in image and video processing," *Proceedings of the IEEE*, vol. 106, no. 8, pp. 1427–1457, 2018.

[22] C. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 4996–5009, 2013.

[23] X. Wang, Z. Peng, D. Kong, P. Zhang, and Y. He, "Infrared dim target detection based on total variation regularization and principal component pursuit," *Image and Vision Computing*, vol. 63, pp. 1–9, 2017.

[24] J. Guo, Y. Wu, and Y. Dai, "Small target detection based on reweighted infrared patch-image model," *IET Image Processing*, vol. 12, no. 1, pp. 70–79, 2018.

[25] S. Yao, Y. Chang, and X. Qin, "A coarse-to-fine method for infrared small target detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 2, pp. 256–260, 2019.

[26] T. Zhang, H. Wu, Y. Liu, L. Peng, C. Yang, and Z. Peng, "Infrared small target detection based on non-convex optimization with Lp-norm constraint," *Remote Sensing*, vol. 11, no. 5, 2019.

[27] Z. Fan, D. Bi, L. Xiong, S. Ma, L. He, and W. Ding, "Dim infrared image enhancement based on convolutional neural network," *Neurocomputing*, vol. 272, pp. 396–404, 2018.

[28] J. Ryu and S. Kim, "Small infrared target detection by data-driven proposal and deep learning-based classification," Edited by B. F. Andresen, G. F. Fulop, C. M. Hanson, J. L. Miller, and P. R. Norton, Eds., in *Proceedings of SPIE on Infrared Technology and Applications XLIV*, vol. 10624, Orlando, FL, USA, May 2018.

[29] W. Wang, H. Qin, W. Cheng, C. Wang, H. Leng, and H. Zhou, "Small target detection in infrared image using convolutional neural networks," Edited by Y. Jiang, H. Gong, W. Chen, and J. Li, Eds., in *Proceedings of SPIE-AOPC 2017: Optical Sensing and Imaging Technology and Applications*, vol. 10462, Beijing, China, 2017.

[30] S. Kim, "Infrared variation reduction by simultaneous background suppression and target contrast enhancement for deep convolutional neural network-based automatic target recognition," *Optical Engineering*, vol. 56, no. 6, pp. 1–11, 2017.

[31] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992.

[32] Q. Zhang, X. Lu, L. Sun, and W. Chen, "Extended target detection amid clutter suppression using the combination of the sparsity and total variation," in *Proceedings of the IEEE CIE International Conference on Radar (RADAR)*, Guangzhou, China, October 2016.

[33] L. Yan, H. Fang, and S. Zhong, "Blind image deconvolution with spatially adaptive total variation regularization," *Optics Letters*, vol. 37, no. 14, pp. 2778–2780, 2012.

[34] H. Fang, L. Yan, H. Liu, and Y. Chang, "Blind Poissonian images deconvolution with framelet regularization," *Optics Letters*, vol. 38, no. 4, pp. 389–391, 2013.

[35] Y. Chang, H. Fang, L. Yan, and H. Liu, "Robust destriping method with unidirectional total variation and framelet regularization," *Optics Express*, vol. 21, no. 20, pp. 23307–23323, 2013.

[36] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[37] J. Cai, E. Candes, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[38] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, pp. 1–37, 2011.