Research Article

Infrared Small Target Detection with Total Variation and Reweighted $\ell_1$ 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 measure inspired by the biological visual mechanism. Shi et al. [16] proposed the detection performance for the coarse-to-fine method is still limited. Zhang et al. [26] proposed to constrain the small target with nonconvex \( \ell_p \) 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 \( \ell_1 \) 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 $\ell_1$ 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 $\ell_1$ 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 $\ell_1$ norm constraint, which leads to a lower false alarm rate.

2. The resulting constructed model contains the nuclear norm, TV norm, and weighted $\ell_1$ 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]:

$$f_D = f_T + f_B + f_N,$$

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:

$$D = T + B + N,$$

where $D$, $T$, $B$, and $N$ 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 $T$ can be described as a sparse matrix:

$$\|T\|_0 < k,$$

where $k$ is determined by the number of small targets and their sizes. $\|\cdot\|_0$ represents the $\ell_0$ 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 $B$ can be depicted as

$$\text{rank}(B) \leq r,$$

where $r$ is an integer determined by the rank of $B$.

The noise is assumed to be Gaussian random noise, and $N$ can be described as

$$\|D - B - T\|_F \leq \delta,$$

where $\delta$ 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

$$\min_{T,B} \|B\|_* + \lambda \|T\|_1$$

s.t. $\|D - T - B\|_F \leq \delta,$

where $\|\cdot\|_*$ is introduced to constrain the low-rank property in $B$. $\|\cdot\|_1$ is used to replace $\|\cdot\|_0$ to simplify the solution of this model. $\lambda$ 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 $\ell_1$ 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 $\ell_1$ norm on the target image are introduced into equation (6). The proposed model can be written as

$$\min_{T,B} \|B\|_{nu,a} + \lambda_1 \|T\|_{W_1} + \lambda_2 \|VT\|_1$$

s.t. $\|D - T - B\|_F \leq \delta,$

$$D = T + B + N,$$
where $\lambda_1$ and $\lambda_2$ are compromising parameters. The vector $w_B$ and the matrix $W_T$ are the weights for the background $B$ and the target $T$, respectively. The weight $w_B$ is incorporated into the nuclear norm term for better preserving the strong edges of the background image [24], which is defined as

$$w_{B,j} = \frac{1}{\sigma_{B,j} + \epsilon_B},$$

(8)

where $\sigma_{B,j}$ is the $j$th singular value of $B$ and $w_B = [w_{B,1}, w_{B,2}, \ldots, w_{B,n}]^T$ is a weight for $\sigma_{B,j}$; $\epsilon_B$ is a positive value. $\|B\|_{w_B,*} = \sum_j w_{B,j}\sigma_{B,j}$ is the reweighted nuclear norm of $B$.

The weight $W_T$ is employed to suppress the nontarget points while enhancing the small target, and each element of $W_T$ is defined as

$$W_{T,ij} = \frac{C}{T_{ij} + \epsilon_T},$$

(9)

where $T_{ij}$ is the pixel value of position $(i, j)$ in the target image $T$. $C$ is a compromising constant and $\epsilon_T$ is a small positive constant to avoid division by zero. The advantage of the weighted $\ell_1$ 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 $\|T\|_{W_{\ell_1}}$ and the TV prior $\|\nabla T\|_1$ using parameters $\lambda_1$ and $\lambda_2$, the proposed regularization constraints on the small target yield a sound target detection result. Therefore, the combination of the TV regularization and the weighted $\ell_1$ 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 $T$ are that $\|T\|_{W_{\ell_1}}$ is nonquadratic and nonseparable and that the $\ell_1$ norm term $\|\nabla T\|_1$ is nonsmooth and nonseparable. The nuclear norm $\|B\|_{w_B,*}$ 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 $T, B$ in (7) into a constrained one by introducing three auxiliary variables $Z_1, Z_2,$ and $Z_3$, and four Lagrangian multiplier variables $Y_1, Y_2, Y_3,$ and $Y_4$.

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:

$$\min_{T, B} \|Z_1\|_{w_B,*} + \lambda_1 \|Z_2\|_{W_{\ell_1}} + \lambda_2 \|Z_3\|_{1} + \langle Y_1, Z_1 - B \rangle$$

$$+ \langle Y_2, Z_2 - T \rangle + \langle Y_3, Z_3 - \nabla T \rangle + \langle Y_4, D - T - B \rangle$$

$$+ \frac{\beta}{2} \left( \|Z_1 - B\|_F^2 + \|Z_2 - T\|_F^2 + \|Z_3 - \nabla T\|_F^2 + \|D - T - B\|_F^2 \right).$$

(10)

where $\langle \cdot, \cdot \rangle$ denotes the inner product operation. $\beta$ 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 sub-minimization problem while fixing other variables in each step of the iterative procedure.

(1) For $Z_1$ subproblem, it can be transformed to

$$\min_{Z_1} \|Z_1\|_{w_B,*} + \frac{\beta}{2} \|Z_1 - B + \frac{1}{\beta}Y_1\|_F^2.$$  

(11)

The above problem is a nuclear norm minimization problem, which can be readily solved by singular value thresholding [37], so the solution of $Z_1$ can be written as

$$Z_1 = SVT_{\frac{w_B}{\beta}}\left( B - \frac{1}{\beta}Y_1 \right).$$  

(12)

where $SVT_{\beta} (\cdot)$ is the singular thresholding operator.
SVT_{\mu}(Y) = \text{Udiag}(\sigma - \mu)_{\beta}V^T,
\sigma - \mu = \begin{cases} 
\sigma - \mu, & \sigma > \mu, \\
0, & \text{otherwise}.
\end{cases}

(13)

Here, X = U\Sigma V^T is the singular value decomposition of matrix X, and \Sigma = \text{diag}(|\sigma|_{1 \leq r}) are the r largest singular values of matrix X.

(2) For Z_2 subproblem, it can be translated to
\[
\min_{Z_2} \lambda_2 \|Z_2\|_{W_2, \beta}^2 + \frac{\|T - \frac{1}{\beta}Y_2\|^2}{2}. \tag{14}
\]

The optimization problem (14) is convex and can be solved by using the well-known soft thresholding operator:
\[
Z_2 = \text{softTh}_{\frac{1}{\beta}Y_2}(T - \frac{1}{\beta}Y_2). \tag{15}
\]

where
\[
\text{softTh}_{\epsilon}(X) = \text{sign}(X)\max\{0, |X| - \epsilon\}, \tag{16}
\]
where |·| denotes the absolute value, and all operations are implemented by componentwise. For \( t \in \mathbb{R} \), the signum function \text{sign}(t) is defined as
\[
\text{sign}(t) = \begin{cases} 
+1, & t > 0, \\
0, & t = 0, \\
-1, & t < 0.
\end{cases} \tag{17}
\]

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

(3) For Z_3 subproblem, it can be translated to
\[
\min_{Z_3} \lambda_3 \|Z_3\|_1^2 + \frac{\|\nabla T - \frac{1}{\beta}Y_3\|^2}{2}. \tag{18}
\]

Optimization problem (18) can also be solved by using the soft thresholding operator. Because \( \nabla T = [\nabla_x X, \nabla_y X]^T \), equation (18) can be solved by two equations:
\[
Z_{3x} = \text{softTh}_{\frac{1}{\beta}Y_3}(\nabla_x T - \frac{1}{\beta}Y_3), \tag{19}
\]

where \( \nabla_x \) and \( \nabla_y \) are the horizontal and vertical first-order forward finite difference operators, respectively.

(4) For B subproblem, it can be translated to
\[
\min_B \left\| Z_1 - B + \frac{1}{\beta}Y_1 \right\|_F^2 + \left\| D - T - B + \frac{1}{\beta}Y_4 \right\|_F^2. \tag{21}
\]

It is a quadratic programming problem and has an analytical solution:
\[
B = \frac{1}{2} \left( Z_1 + D - T + \frac{1}{\beta}(Y_1 + Y_4) \right). \tag{22}
\]

(5) For T subproblem, it can be translated to
\[
\min_T \left\| Z_2 - T + \frac{1}{\beta}Y_2 \right\|_F^2 + \left\| \nabla T - \left( Z_3 + \frac{1}{\beta}Y_3 \right) \right\|_F^2 \tag{23}
\]

\[
\| T - (D - B + \frac{1}{\beta}Y_4) \|_F^2. \]

Optimization problem (23) is also a quadratic programming problem. The equivalent liner system of the model (23) can be expressed as
\[
(2I + \nabla^T \nabla)T = Z_2 + \nabla^T Z_3 + \frac{1}{\beta}(Y_2 + \nabla^T Y_3 + Y_4) + D - B, \tag{24}
\]

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):
\[
T = \mathcal{F}^{-1}\left\{ \mathcal{F}\left( \frac{\beta \zeta_1 + \nabla^T \zeta_2 + \zeta_3}{2\beta + \beta \mathcal{F}(\nabla^T \nabla)} \right) \right\}, \tag{25}
\]

where \( \mathcal{F}^{-1} \) and \( \mathcal{F} \) denote the inverse FFT and FFT operators, respectively, and \( \zeta_1 = Z_3 + D + B, \zeta_2 = \beta Z_3 + Y_3, \) and \( \zeta_3 = Y_2 + Y_4. \)
6 Mathematical Problems in Engineering

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 \( D \in \mathbb{R}^{a \times n} \) (we set \( m = a^2, n = t \)) is constructed from the original image \( f_D \in \mathbb{R}^{M \times N} \) 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 \( T \) and the background patch-image \( B \) are obtained from the patch-image \( D \) by applying Algorithm 1.

3. The target image \( f_T \) and background image \( f_B \) are reconstructed from the patch-image \( T \) and \( B \), 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

\[
f_D(x, y) = \begin{cases} 
\max(r f_T(x - x_0, y - y_0), f_B(x, y)), & x \in (1 + x_0, n + x_0), y \in (1 + y_0, m + y_0), \\
\min(f_B(x, y)), & \text{otherwise.}
\end{cases}
\]

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

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

\[
t_{up} = \mu + k \sigma,
\]

where \( \mu \) and \( \sigma \) 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 \( f_T(x, y) > t_{up} \); 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.
(1) **Input:** infrared patch-image matrix \( D \in \mathbb{R}^{m \times n} \), \( \lambda_1, \lambda_2, C, \beta \).

(2) **Output:** infrared patch-image matrix \( B, T \).

(3) **Initialization:** \( B_0 = \text{zeros}(m,n) \), \( T_0 = \text{zeros}(m,n) \), \( Y_1^0 = Y_2^0 = Y_3^0 = Y_4^0 = \text{zeros}(m,n) \), \( \beta_{\text{max}} = 10^7, k = 0 \).

(4) while not converged do do

(5) **Updating procedure for** \( Z_{k+1}^1 \)

\[
Z_{k+1}^1 = \text{SVT}_{w_1}^{(1/\beta)} (B^k - (1/\beta)Y_2^k);
\]

(6) **Updating procedure for** \( Z_{k+1}^2 \)

\[
Z_{k+1}^2 = \text{softTh}_{1/\beta} (T^k - (1/\beta)Y_4^k);
\]

(7) **Updating procedure for** \( Z_{k+1}^3 \)

\[
Z_{k+1}^3 = \text{softTh}_{1/\beta} (V_x T^k - (1/\beta)Y_3^k);
\]

(8) **Updating procedure for** \( Z_{k+1}^4 \)

\[
Z_{k+1}^4 = \text{softTh}_{1/\beta} (V_y T^k - (1/\beta)Y_1^k);
\]

(9) **Updating procedure for** \( T^{k+1} \)

\[
T^{k+1} = \mathcal{F}^{-1} \{(\mathcal{F} (\beta Z_{k+1}^3 + \nabla_T T^k + C_3^k) + \nabla_T) / (2\beta + \beta \mathcal{F} (\nabla_T \mathcal{F} \nabla_T))\};
\]

(10) **Updating procedure for** \( \beta^{k+1} \)

\[
\beta^{k+1} = \min \{\rho \beta^k, \beta_{\text{max}}\};
\]

(11) **Check the convergence conditions**

\[
||D - T^{k+1}||_F / D_F < \epsilon;
\]

(12) **Update** \( k \leftarrow k + 1 \).

(13) **Update Lagrangian multipliers** \( Y_1^k, Y_2^k, Y_3^k, Y_4^k \) using (8) and (9), respectively;

(14) **Update** \( w_{B_k}^k \) and \( w_{T_k}^k \) using (8) and (9), respectively;

(15) **Updating procedure for** \( B^{k+1} \)

\[
B^{k+1} = (1/2)(Z_{k+1}^3 + D - T^k + (1/\beta^k)(Y_1^k + Y_2^k));
\]

(16) **Updating procedure for** \( B^{k+1} \)

\[
B^{k+1} = (1/2)(Z_{k+1}^3 + D - T^k + (1/\beta^k)(Y_1^k + Y_2^k));
\]

(17) **Updating procedure for** \( B^{k+1} \)

\[
B^{k+1} = (1/2)(Z_{k+1}^3 + D - T^k + (1/\beta^k)(Y_1^k + Y_2^k));
\]

(18) **Updating procedure for** \( B^{k+1} \)

\[
B^{k+1} = (1/2)(Z_{k+1}^3 + D - T^k + (1/\beta^k)(Y_1^k + Y_2^k));
\]

(19) **Updating procedure for** \( B^{k+1} \)

\[
B^{k+1} = (1/2)(Z_{k+1}^3 + D - T^k + (1/\beta^k)(Y_1^k + Y_2^k));
\]

(20) **Updating procedure for** \( B^{k+1} \)

\[
B^{k+1} = (1/2)(Z_{k+1}^3 + D - T^k + (1/\beta^k)(Y_1^k + Y_2^k));
\]

(21) **Updating procedure for** \( B^{k+1} \)

\[
B^{k+1} = (1/2)(Z_{k+1}^3 + D - T^k + (1/\beta^k)(Y_1^k + Y_2^k));
\]

(22) **Updating procedure for** \( B^{k+1} \)

\[
B^{k+1} = (1/2)(Z_{k+1}^3 + D - T^k + (1/\beta^k)(Y_1^k + Y_2^k));
\]

(23) **end While**

---

**Algorithm 1:** Target and background separation solver via ADMM.

---

**Figure 1:** The various steps of the proposed algorithm.

---

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:

\[
\text{SCRG} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}}}
\]

(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:

\[
\text{BSF} = \frac{\sigma_{\text{in}}}{\sigma_{\text{out}}}
\]

(31)
where $\sigma_{\text{in}}$ and $\sigma_{\text{out}}$ 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:

$$\text{CG} = \frac{\text{CON}_{\text{out}}}{\text{CON}_{\text{in}}}$$  \hspace{1cm} (32)

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

$$\text{CON} = |\mu_t - \mu_b|,$$  \hspace{1cm} (33)

where $\mu_t$ and $\mu_b$ 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:

$$P_d = \frac{\#\text{true targets detected}}{\#\text{total background targets}},$$

$$F_a = \frac{\#\text{false pixels detected}}{\#\text{total pixels in images}}.$$  \hspace{1cm} (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 $\lambda_1$ and $\lambda_2$ 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 $\ell_1$ norm, and TV term. We set the parameter $\lambda_1$ to be $1/\sqrt{\max(m,n)}$ the same as in PCP problem [38]. For $\lambda_2$, a large value of $\lambda_2$ means that the influence of TV regularization becomes more significant and that the clutters and noise can be suppressed remarkably. A small value of $\lambda_2$ may cause more clutters and noise residual in the target image. In our implementation, we
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. $\beta$ 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): $\text{NSDE} = \|\mathbf{D} - \mathbf{T}^{k+1} - \mathbf{B}^{k+1}\|_F/\|\mathbf{D}\|_F < \varepsilon$ 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 proper 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

### Table 3: Detailed parameter setting for the test methods.

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

In order to choose the appropriate selection of these parameters, the following cases are considered. For each measure of SCR, BSF, and CG, a higher score means that the background clutters and noise 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.
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 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.
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 (f1)–(f8), 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
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 $\ell_1$ norm, some non-target points are also suppressed. In a word, the combination of the TV regularization and the weighted $\ell_1$ norm is
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
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 $\ell_1$ 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 evolitional 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
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 \( O(MN L^2 \log L^2) \), 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 \( O(L^2 \log L^2) \). As for the Max-mean method, the total cost is \( O(MN L^2) \). 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, \( \ell_1 \) norm, and total variation norm and is solved by ADMM. For the proposed method, the computational complexities of updating the variables \( Z_1, Z_2, Z_{3x}, Z_{3y}, B, \) and \( T \) 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 \( \ell_1 \) 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 \( \ell_1 \) 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.

### 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


