

## Article

# Infrared Single-Frame Small Target Detection Based on Block-Matching

Yi Man <sup>1,2</sup>, Qingyun Yang <sup>3,4,\*</sup> and Tao Chen <sup>1</sup>

<sup>1</sup> Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup> Academy for Advanced Interdisciplinary Studies, Northeast Normal University, Changchun 130024, China

<sup>4</sup> Laboratory of Pinghu (Beijing Institute of Infinite Electric Measurement), Pinghu 314200, China

\* Correspondence: yangqy333@nenu.edu.cn; Tel.: +86-135-0447-5484

**Abstract:** The robust detection of small targets is one of the crucial techniques in an infrared system. It is still a challenge to detect small targets under complex backgrounds. Aiming at the problem where infrared small target detection is easily disturbed by complex backgrounds, an infrared single frame detection method based on a block-matching approach is proposed in this paper. Firstly, the input infrared image is processed by extracting blocks from it. A new infrared model is constructed by finding blocks that are similar to each such block. Then, the small target detection based on the block-matching model is formulated as an optimization problem of recovering low-rank and sparse matrices, which are effectively solved using robust principal component analysis. Finally, the results of processing are reconstructed to obtain the target and background images. A simple segmentation method is used to segment the target image. The experimental results from the actual infrared sequences show that the proposed method has better background suppression ability under complex backgrounds and better detection performance than conventional baseline methods.

**Keywords:** infrared image; block-matching; small target detection; low-rank matrix recovery



**Citation:** Man, Y.; Yang, Q.; Chen, T. Infrared Single-Frame Small Target Detection Based on Block-Matching. *Sensors* **2022**, *22*, 8300. <https://doi.org/10.3390/s22218300>

Academic Editors: Yuhan Liu and Xiaoyang Wang

Received: 27 September 2022

Accepted: 24 October 2022

Published: 29 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Infrared detection technology has many characteristics, such as anti-electromagnetic interference, a wide temperature range, and real-time observation. It has since been widely used in various applications, especially in the military and remote sensing [1]. Generally, small targets have a tiny proportion of pixels in infrared images. Meanwhile, target message loss is profound because of the low signal-to-noise ratio and poor contrast [2]. Background clutter and cloud also will increase the difficulty of infrared small target detection [3]. Therefore, infrared small target detection technology has great significance and is still a challenging study. The goals of researchers have been better real-time performance, a higher probability of detection, and a lower false-alarm rate.

Many scholars have recently proposed many effective methods to detect small targets. Currently, the practices of infrared small target detection can be divided into detection before track (DBT) and track before detection (TBD). Typical TBD methods include the dynamic programming method [4], the particle filter method [5], 3D matched filtering [6], and the pipeline filter method [7]. However, TBD methods combine the time and space information of sequence images to predict the target trajectory, which is challenging to apply in practice. In contrast, DBT methods detect small targets on a single frame, which have the advantages of fast detection speed and a high robustness to noise. The DBT methods can be summarized into three categories: the methods based on target features, the methods based on background features, and the methods based on low-rank sparse matrix recovery.

For methods based on target features, local information differences can be recognized in the human visual system [8]. Therefore, many small target detection methods based on the visual system are proposed [9–16]. Local Contrast Measure (LCM) is a usual method [9]. It combines the contrast mechanism of the human visual system with the derived kernel model to achieve target enhancement and background clutter suppression. However, LCM makes an excessive enhancement for noises with high brightness, resulting in a high false-alarm rate. To solve this problem, Han et al. proposed an improved Local Contrast Measure (ILCM) method [10]. To improve the detection performance in complex backgrounds, Han et al. proposed the NLCM method [11]. Wei et al. proposed MPCM from the perspective of image patch difference [12]. Deng et al. proposed the weighted local difference method (WLDM) [13], which can better separate the real targets from the interfering object. Han et al. proposed a multi-scale relative local Contrast ratio (RLCM) [14], which can efficiently detect small targets of different sizes in complex backgrounds. Guan et al. proposed the enhanced Local Contrast Measure (ELCM) [15], and Han et al. proposed the weighted enhanced local Contrast (WSLCM) [16] for infrared small target detection. For methods based on background features, Max-Mean and Max-Median [17], based on nonlinear filtering; and classical Top-Hat [18], based on morphological filtering, are traditional. They all use the background difference to obtain small targets. Bai [19] et al. studied new Top-Hat and applied it to infrared small target detection. Zhu [20] et al. combined Top-Hat with low-rank tensor completion to provide a more robust small target detection model. In addition, the two-dimensional least mean square (TDLMS) filter [21] and bilateral filter [22] also have good results using background prediction. However, these methods are not very effective in dealing with complex backgrounds.

With the development of matrix optimization theory, the low-rank sparse matrix recovery method is attracting more and more attention. This method can separate the target image from the background image according to the target's sparsity and the background's low rank. The most representative process is the infrared patch-image model (IPI) [23]. Dai et al. proposed a weighted infrared patch-image model (WIPI) [24] to suppress the edges better. The reweighted infrared patch-tensor model (RIPT) [25] and the non-negative infrared patch-image model via partial sum minimization for singular values model (NIPPS) [26] could provide more accurate background estimations. In terms of the matrix recovery improvement, some methods include low-rank sparse representation (LRSR) [27], adaptive weighted parameter [28], total variation regularization, principal component pursuit (TV-PCP) [29], and non-convex rank approximation minimization (NRAM) [30]. Rawat et al. proposed a method via total variation and partial sum minimization (TV-PSMSV) to recognize small targets in a highly complex background [31].

When the background of the infrared image is very complex, such as buildings, mountains, and trees, the low-rank property of the background is very poor, so the conventional low-rank sparse matrix recovery methods are not very effective. In order to resolve the aforementioned problem, this paper proposes a novel infrared small targets detection strategy based on block-matching. The proposed method uses non-local similarity to match and group image blocks to construct a new infrared model. Firstly, the original infrared image is divided into some reference blocks. Similar blocks are searched throughout the image for each reference block, and several block-matching groups are formed. Secondly, each block-matching group is formed into a matrix. We suppose the infrared background image is a low-rank matrix, and that the infrared target image is a sparse matrix. The small target detection problem is transformed into an optimization problem of recovering the low-rank and sparse matrices. Finally, the target image and background image are obtained via reconstruction. A simple threshold segmentation method is used to obtain a small target. Experimental results show that the proposed method has better stability under different backgrounds and better detection performance than the traditional baseline method.

The remainder of this paper is organized as follows. In Section 2, we introduce the process of the new infrared block-matching model construction and reconstruction, and we also analyze the model. Section 3 describes the small target detection method based on

block-matching. In Section 4, we present the experimental results of block size effects and comparison experiments in this paper. The conclusion is given in Section 5.

## 2. Infrared Block-Matching Model

In natural images, self-similarity generally exists between image blocks [32]. Many classical image denoising algorithms have been proposed by utilizing the self-similarity of images [33–35]. The distance between blocks usually expresses the similarity between two blocks. The smaller the distance between the two blocks, the higher the similarity. Block-matching is a common method in image denoising and motion estimation. It calculates the distance between a reference block and a block to be matched, finds some matching blocks closest to the reference block, and divides them into a set. Figure 1 shows some reference blocks in the image and matched blocks with high similarity.



**Figure 1.** Illustration of matching blocks of infrared images. Each image shows a reference block marked with red borders and a few of the blocks marked with blue borders matched to it.

### 2.1. Construction of Infrared Block-Matching Model

We construct the infrared block-matching model from an original image. The steps are as follows. Firstly, an input infrared image is clipped into several reference blocks. Then, for each reference block, a sliding window is used to search for similar blocks with the reference block from left and top to right and down. Block-matching is a method for finding blocks similar to a given reference one [34]. If the distance between the blocks and the reference one is smaller than a given threshold, they are mutually similar. Therefore, only blocks whose distance concerning the reference one is smaller than a threshold are deemed similar. In particular, we use the  $\ell^2$ -distance as a measure. The block distance can be calculated as:

$$d(Z_{xr}, Z_x) = \|Z_{xr} - Z_x\|_2^2, \quad (1)$$

where  $\|\cdot\|_2$  denotes the  $\ell^2$ -norm, and  $Z_{xr}$  and  $Z_x$  respectively represent the reference block and the block to be matched.

We can obtain the set composed of similar blocks from  $d(Z_{xr}, Z_x)$ :

$$S = \{d(Z_{xr}, Z_r) < \tau\}, \quad (2)$$

where  $\tau$  is a constant representing the maximum matching threshold of similar blocks. Finally, for each set, we transform each block into a column vector to form a new matrix.

### 2.2. Analysis of Infrared Block-Matching Model

A single infrared image is usually considered to be composed of three parts:

$$F_d = F_b + F_t + F_n, \quad (3)$$

where  $F_d$ ,  $F_b$ ,  $F_t$ , and  $F_n$  are the original image, background image, target image, and noise image, respectively. By constructing the infrared block-matching model for the infrared input images, we can transform the infrared image model into the corresponding infrared block-matching image model:

$$D = B + T + N, \quad (4)$$

where  $D$ ,  $B$ ,  $T$ , and  $N$  respectively represent the original image, background image, target image, and noise image in the infrared block-matching model.

Target block-matching T: Small targets are tiny in infrared images. Due to the low proportion of small targets, target block-matching T can be considered a sparse matrix. That is:

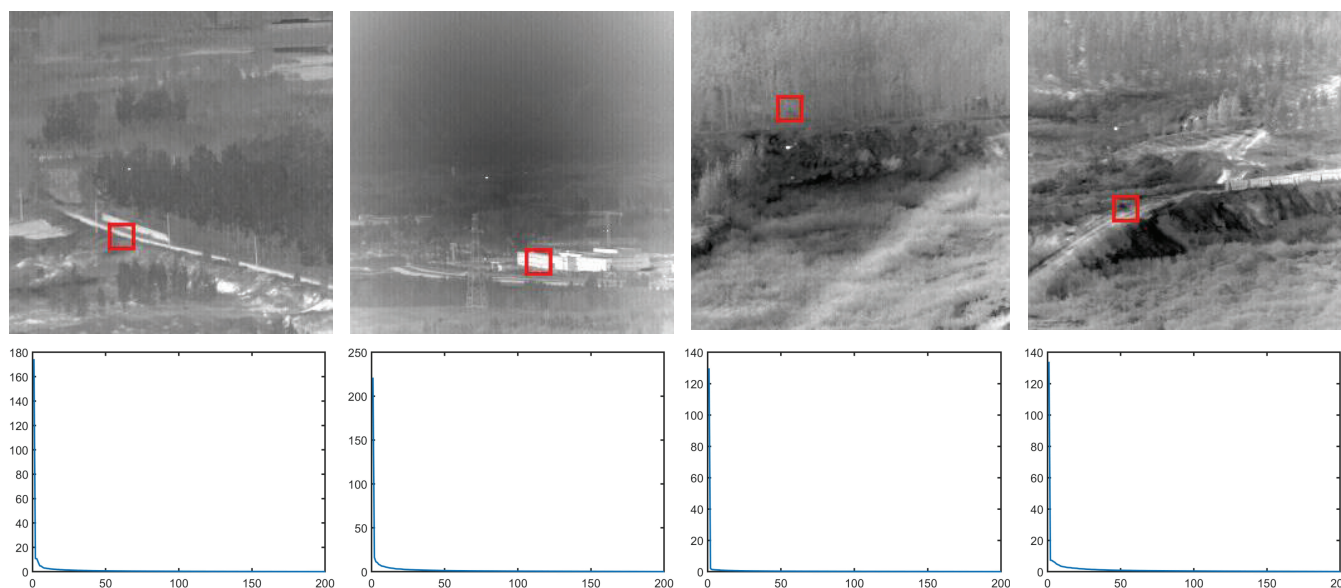
$$\|T\|_0 \leq k, \quad (5)$$

where  $\|\cdot\|_0$  denotes the  $\ell^0$ -norm and  $k$  is a constant determined by the size and number of the small target.

Background block-matching B: As previously discussed, the infrared background image usually has strong non-local self-similarity. We select three representative infrared background images. Additionally, the IBM model matrix is constructed for a reference block of each image. The reference block size is  $20 \times 20$ . Meanwhile, we calculate the singular value of the matrix. Figure 2 shows that their singular values all rapidly decrease to zero. Therefore, we can consider background block-matching B as a low-rank matrix. That is:

$$\text{Rank}(B) \ll r, \quad (6)$$

where  $r$  is a constant representing the complexity of the background image. The more complex the background, the greater the value of  $r$ .



**Figure 2.** Illustration of the low-rank property of background block-matching B. The first row are four images, each with a background reference block with a red border, and the second one are the singular values of background block-matching B constructed using four reference blocks.

Noise Block-Matching N: In this paper, we assume that the noise is additive white Gaussian noise and  $\|N\|_F \leq \delta$  for some  $\delta > 0$ . Thus, we have:

$$\|D - B - T\|_F \leq \delta, \quad (7)$$

where  $\|\cdot\|_F$  denotes the Frobenius norm.

Although the parameters  $k$ ,  $r$ , and  $\delta$  vary for different infrared images, we would not use them directly in the following sections. We will discuss this further in Section 3.1.

### 2.3. Reconstruction of Infrared Images from the Infrared Block-Matching Model

We reconstruct an image from an infrared block-matching model. The steps are as follows. Firstly, the similar blocks with each reference block are returned to their original positions in the infrared image. Because the selected similar blocks may overlap, a pixel

location would correspond to several values from different blocks. In this case, we use a 1D mean filter function to determine the final value of each pixel:

$$V = \text{mean}(x), \quad (8)$$

where  $V \in R$  and  $x$  is a vector containing the values of all similar blocks at one pixel.

### 3. Small Target Detection Based on Block-Matching

#### 3.1. Solution of Infrared Small Target Detection

According to Section 2, we can consider the target block-matching  $T$  as a sparse matrix and the background block-matching  $B$  as a low-rank matrix. Ideally, we transform the small target detection problem based on block-matching into the following low-rank and sparse decomposition problem:

$$\min_{B,T} \text{rank}(B) + \lambda \|T\|_0, \text{ s.t. } B + T = D, \quad (9)$$

where  $\lambda$  denotes the equilibrium factor; we choose the robust principal component analysis (RPCA) to solve (9) because it has better anti-noise interference ability and robustness. Because the parameters  $k$  and  $r$  in (5) and (6) vary for different infrared images, it is difficult to estimate them in advance. Robust principal component analysis can accurately and efficiently recover  $B$  and  $T$  in (9) under broad conditions, such as no relevant information about the rank of  $B$  or the support of  $T$  being given [36]. Therefore, our method does not need to obtain the parameters  $k$  and  $r$  of different images. Unfortunately, (9) is a nondeterministic polynomial hard (NP-hard) problem [37]. This problem cannot be solved using exact algorithms. Therefore, we need to find an efficient approximation algorithm. Here,  $\|B\|_*$  and  $\|T\|_1$  replace  $\text{rank}(B)$  and  $\|T\|_0$ , respectively. Then, we consider the situation of noise in an image. The above objective function can be transformed into the following convex optimization problem:

$$\min_{B,T} \|B\|_* + \lambda \|T\|_1 \text{ s.t. } \|D - B - T\|_F \leq \delta. \quad (10)$$

The traditional methods to solve the standard RPCA model are iterative thresholding (IT) [36] and accelerated proximal gradient (APG) [38]. In addition, exact augmented Lagrange multipliers (EALM) [39] and inexact augmented Lagrange multipliers (IALM) [40] are also commonly used. We choose IALM to solve (10) because of its faster operation speed and higher operation accuracy.

Construct the augmented Lagrange function for (10):

$$\Gamma(B, T, Y, \mu) = \|B\|_* + \lambda \|T\|_1 + \langle Y, D - B - T \rangle + \frac{\mu}{2} \|D - B - T\|_F^2, \quad (11)$$

where  $Y \in R^{m \times n}$  denotes the Lagrange multiplier,  $\mu > 0$  is the penalty parameter, and  $\langle \bullet \rangle$  denotes the standard inner product.  $\mu$  is varied while solving the above optimization, and (10) is equivalent to (11) for some value  $\mu(\delta)$  [41]. When  $Y = Y_K$  and  $\mu = \mu_k$ , the optimization problem  $\min_{B,T} L(B, T, Y, \mu_k)$  is solved using alternate direction methods.

#### 3.2. Target Segmentation

After obtaining the target image, we need to extract a small target from the image. We adopt the same target segmentation threshold as [23]. In this paper, we only detect bright targets. The adaptive threshold  $t_{up}$  is defined by:

$$t_{up} = \max(v_{\min}, \mu + k\sigma), \quad (12)$$

where  $\mu$  and  $\sigma$  respectively denote the average and standard deviation of the target image, and  $K$  and  $V_{\min}$  are constants determined experientially.

### 3.3. The Entire Procedure of Small Target Detection

Figure 3 shows the whole process of the small target detection method proposed in this paper. The steps are: (1) According to Section 2.1, we construct the block-matching D from the input infrared image. (2) We use IALM to decompose the block-matching D into the low-rank background block-matching B and sparse block-matching target T. (3) According to Section 2.3, we reconstruct the target image  $F_t$  and background image  $F_b$ . (4) We obtain the final image by using a simple threshold segmentation on the target image.

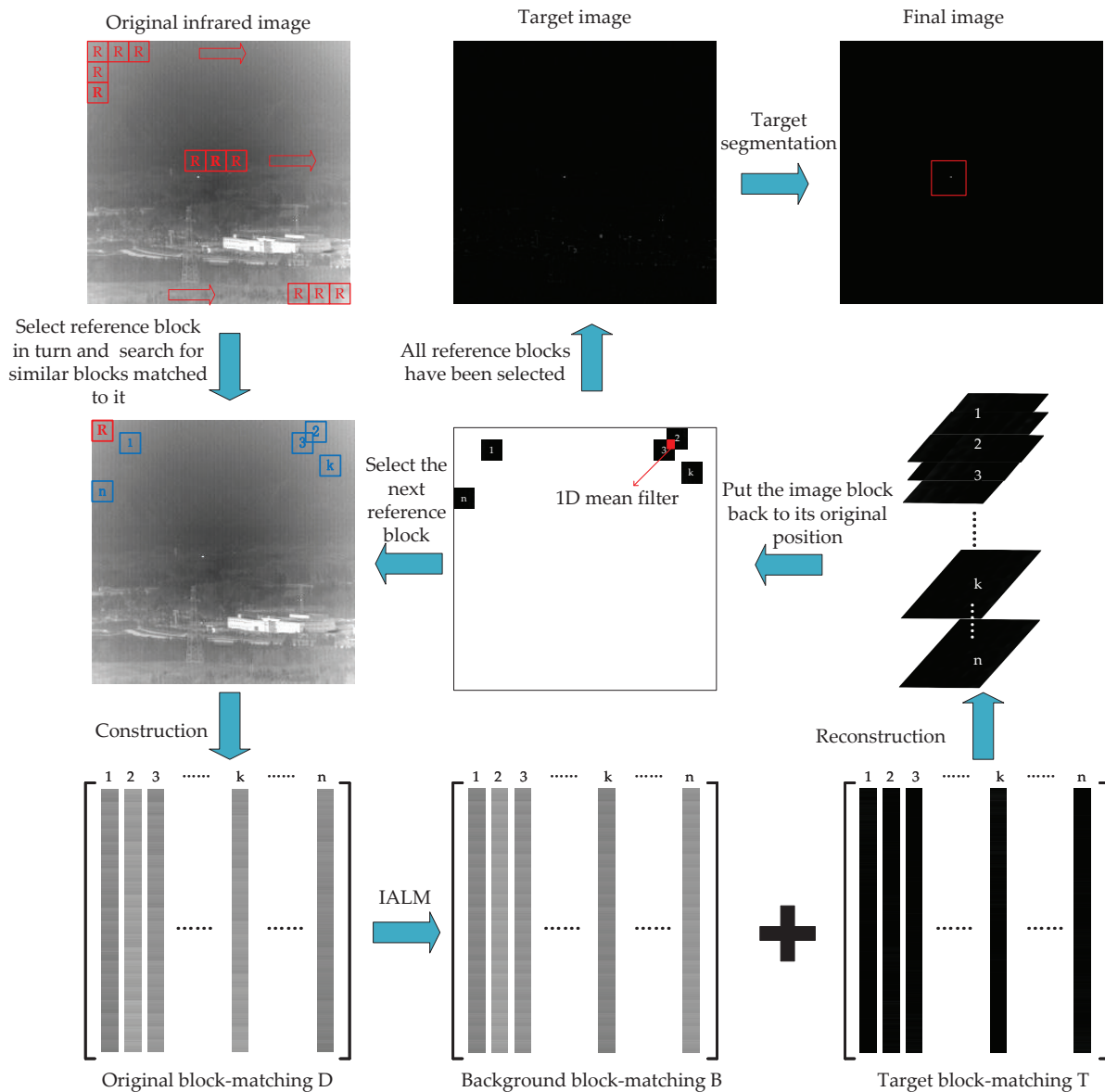


Figure 3. The overview of the proposed method in this paper.

## 4. Experiments

### 4.1. Experimental Settings

Evaluation Metrics: In this paper, we choose background suppression factor (BSF), local signal-to-noise ratio gain (LSNRG), and signal-to-clutter ratio gain (SCRG) to evaluate the performance. Three evaluation metrics are calculated in a local region concerning the target, as illustrated in Figure 4. The target size is  $a \times b$  and the neighborhood width is  $d$ , which is a constant. We set  $d = 20$  in this paper. BSF represents the inhibitory ability to the background. The higher the value of BSF, the better the background suppressing performance. BSF is defined as follows:

$$BSF = \frac{\sigma_{in}}{\sigma_{out}}, \quad (13)$$

where  $\sigma_{in}$  and  $\sigma_{out}$  represent the local background neighborhood standard variances of the input and output images, respectively. LSNRG measures the local signal-to-noise ratio (LSNR) gain, where a higher value of LSNRG denotes a better performance. LSNR is defined as follows:

$$LSNR = \frac{P_T}{P_B}, \quad (14)$$

where  $P_T$  and  $P_B$  represent the maximum target pixel value and the maximum local background neighborhood pixel value, respectively.

$$LSNRG = \frac{LSNR_{out}}{LSNR_{in}}, \quad (15)$$

where  $LSNR_{out}$  and  $LSNR_{in}$  stand for the LSNR values of the output and input images. SCRG represents the validity of target enhancement. The stronger the target enhancement ability, the larger the value. SCR is defined as follows:

$$SCR = \frac{|\mu_t - \mu_b|}{\sigma_b}, \quad (16)$$

where  $\mu_t$  is the average pixel value of the target, and  $\mu_b$  and  $\sigma_b$  are the average pixel value and the standard deviation of the local background neighborhood, respectively.

$$SCRG = \frac{SCR_{out}}{SCR_{in}}, \quad (17)$$

where  $SCR_{out}$  and  $SCR_{in}$  stand for the SCR values of the output and input images. In addition to the three evaluation indicators, the receiver operating characteristic (ROC) curve is also an important metric. The ROC curve can directly reflect the accuracy of detection and help to find the appropriate threshold. For small target detection, the larger the area enclosed by the curve and abscissa, the better the method's performance. Additionally, the higher detection probability for the same false-alarm ratio means a better performance. The ROC curve is plotted using the detection probability  $P_d$  and the false-alarm rate  $F_a$ .  $P_d$  and  $F_a$  are defined as follows:

$$P_d = \frac{\text{number of true detections}}{\text{number of actual targets}} \quad (18)$$

$$F_a = \frac{\text{number of false pixels detected}}{\text{number of pixels in all images}}. \quad (19)$$

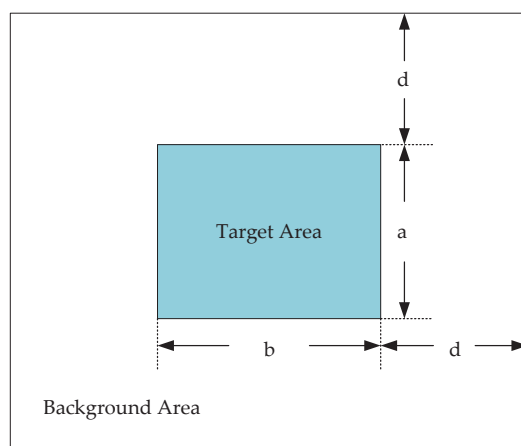


Figure 4. Target area and background neighborhood area of the small target.

**Baseline Methods:** In this paper, we choose Max-Mean [17], Max-Median [17], Top-Hat [18], IPI [23], and TV-PCP [29] as the baseline methods. The detailed parameters of these methods are shown in Table 1.

**Experimental Datasets:** In this paper, we select four infrared sequences for the experiments. All sequences are from public datasets [42]. Detailed information on the infrared sequence is shown in Table 2.

**Table 1.** Detailed parameters of baseline methods.

Methods	Parameters
Max-Mean [17]	Filter size: $15 \times 15$
Max-Median [17]	Filter size: $15 \times 15$
Top-Hat [18]	Shape: square, filter size: $15 \times 15$
IPI [23]	Patch size: $50 \times 50$ , sliding step: 10, $\lambda = \frac{1}{\sqrt{\min(m,n)}}$ , $\varepsilon = 10^{-7}$
TV-PCP [29]	Patch size: $50 \times 50$ , sliding step: 14, $\lambda = 0.005$ , $\lambda_2 = \frac{1}{\sqrt{\min(m,n)}}$ , $\beta = 0.025$ , $\gamma = 1.5$ , maxIter = 250, tol = $5 \times 10^{-6}$

**Table 2.** Detailed description of experimental sequences.

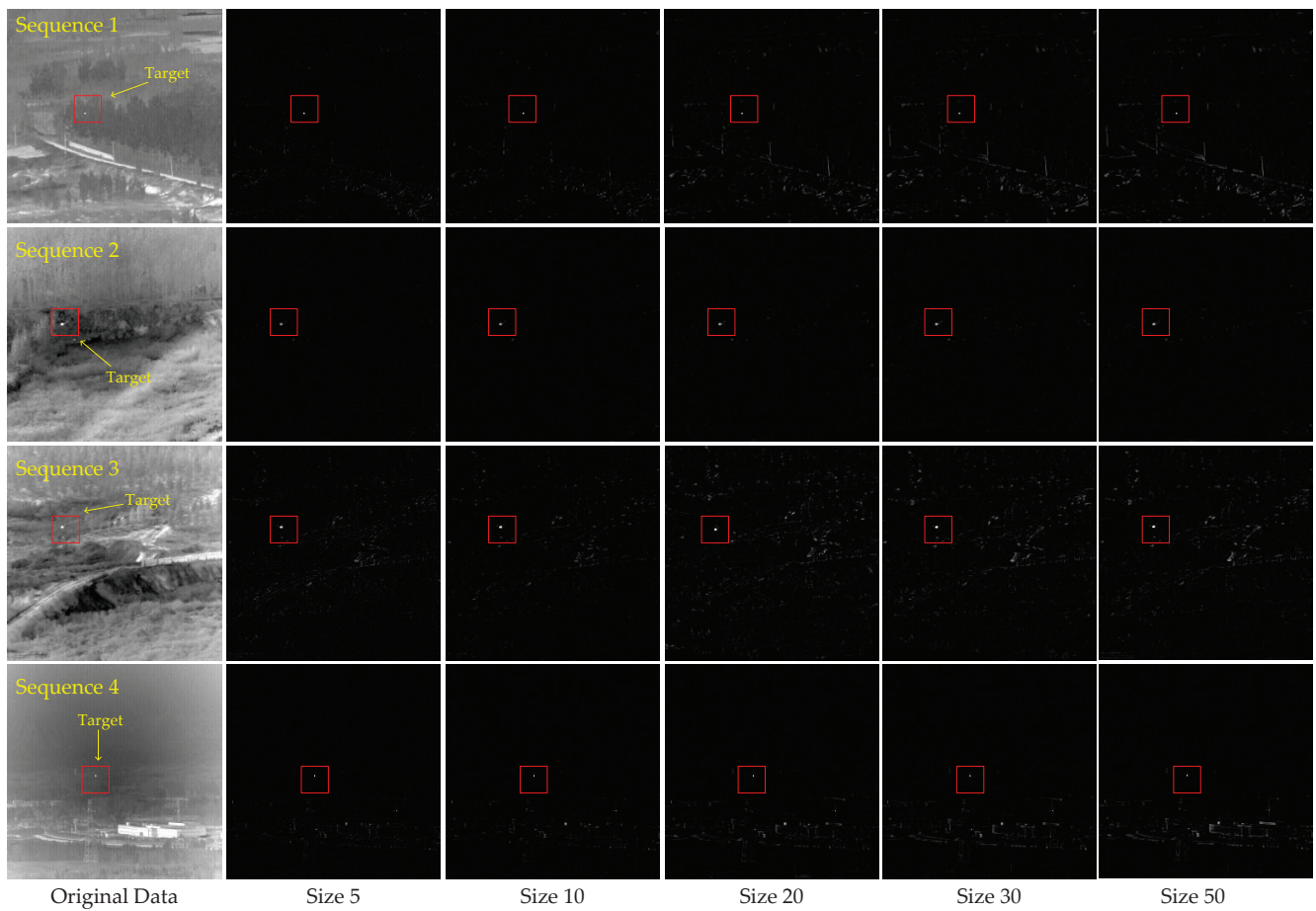
N0.	Frame	Image Size	Characteristics
1	200	$256 \times 256$	Complex background with trees and highlighted interference
2	200	$256 \times 256$	Complex background with thickets
3	200	$256 \times 256$	Complex ground background with trees and road
4	200	$256 \times 256$	Complex background with constructions

#### 4.2. Effect of Block Size

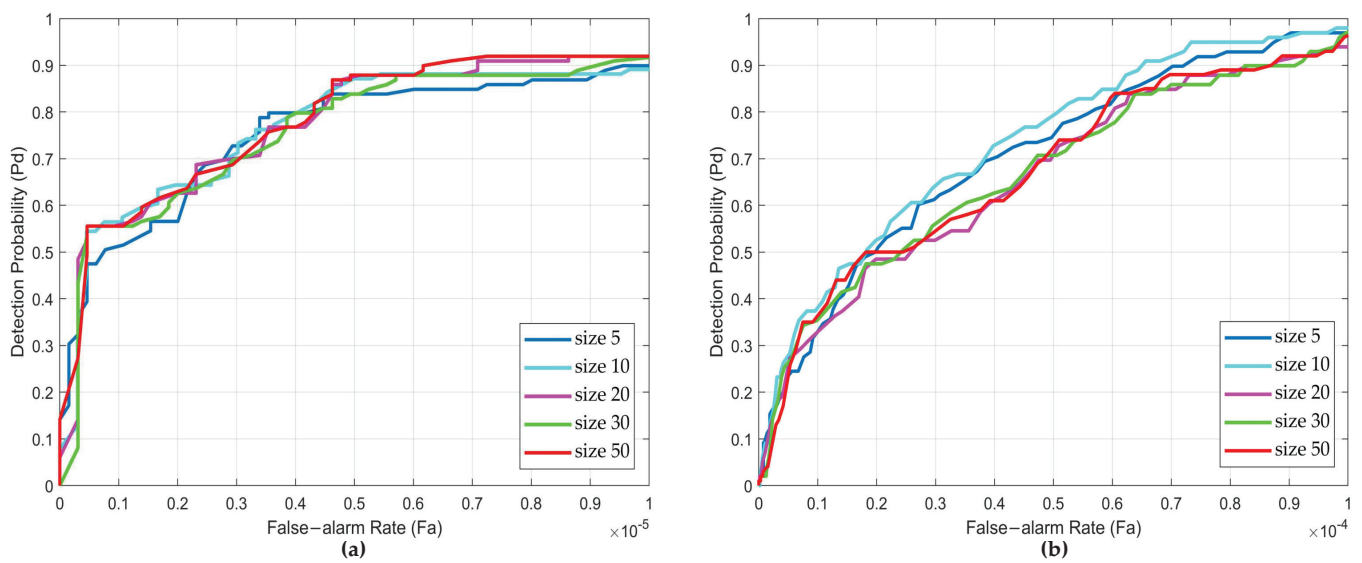
In our method, the size of the reference block is an important parameter. Different block sizes will affect the algorithm performances for different datasets. Thus, we need to adjust it to obtain better detection results. We set the square block size as  $5 \times 5$ ,  $10 \times 10$ ,  $20 \times 20$ , and  $50 \times 50$ . Figure 5 shows the separated target images according to the different sizes of the reference block. We adjust the contrast of the images for clearer viewing. As seen from Figure 5, the smaller the block size, the better the background suppression performance, and the less the background residue of the target images for the four sequences. This is because the smaller the block size, the higher the matching degree between the matching block and the reference block. This results in a sparser image of the block-matching model when using a smaller block size. However, as the image of the constructed block-matching model is sparser, the low-rank property of targets will be lost to a certain extent. Therefore, the contrast between the target and background also decreases.

To further discuss the effect of block size, we display the ROC curve for different block sizes. The ROC curves for the four sequences are shown in Figure 6. It can be seen from Figure 6 that the ROC curves of different sequences under different block sizes have little differences. In particular, for Seq.2 and Seq.3, the smaller the block size, the better the performance. However, the pattern is not evident for Seq.1 and Seq.4. This is because Seq.2 and Seq.3 have a larger target size than Seq.1 and Seq.4. According to the above analysis, the low-rank property of the target will be lost to a certain extent. As the size of the target becomes bigger, this effect becomes more pronounced. It should be noted that the smaller the block size, the longer it takes to obtain the separated target images. In combination with the above qualitative and quantitative analysis, we choose block size  $10 \times 10$  as the parameters of our method in the subsequent contrast experiments, based on the consideration of computational cost.

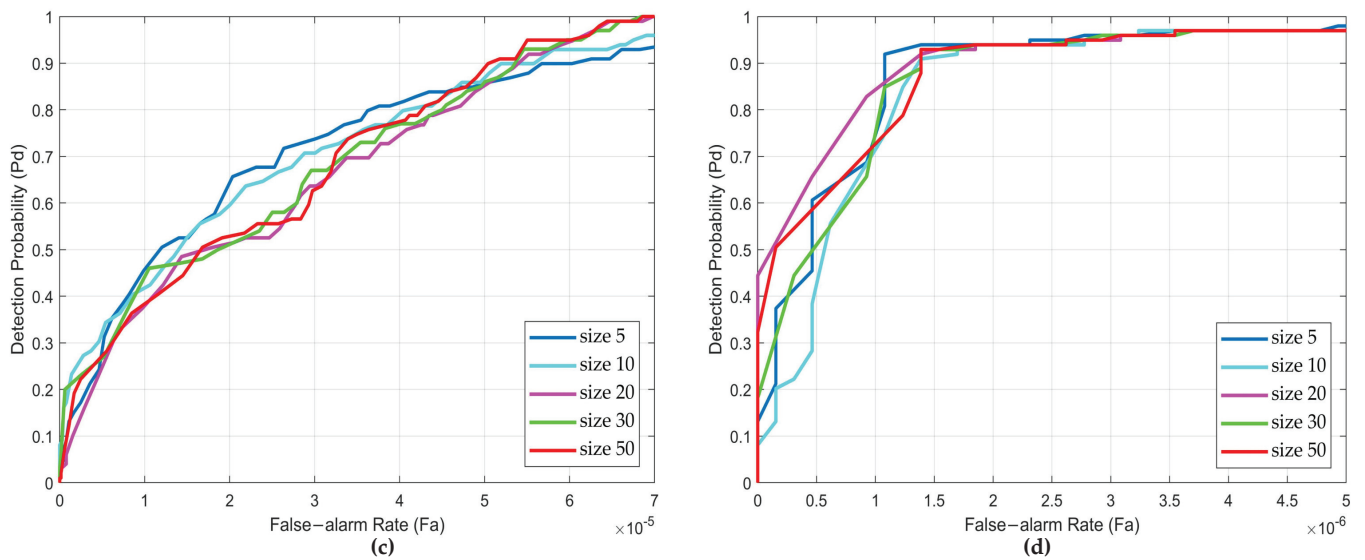




**Figure 5.** The representative images of the four real image sequences and the corresponding processed target images of different block sizes. From left to right are the original data and the processing results of size 5, size 10, size 20, size 30, and size 50.



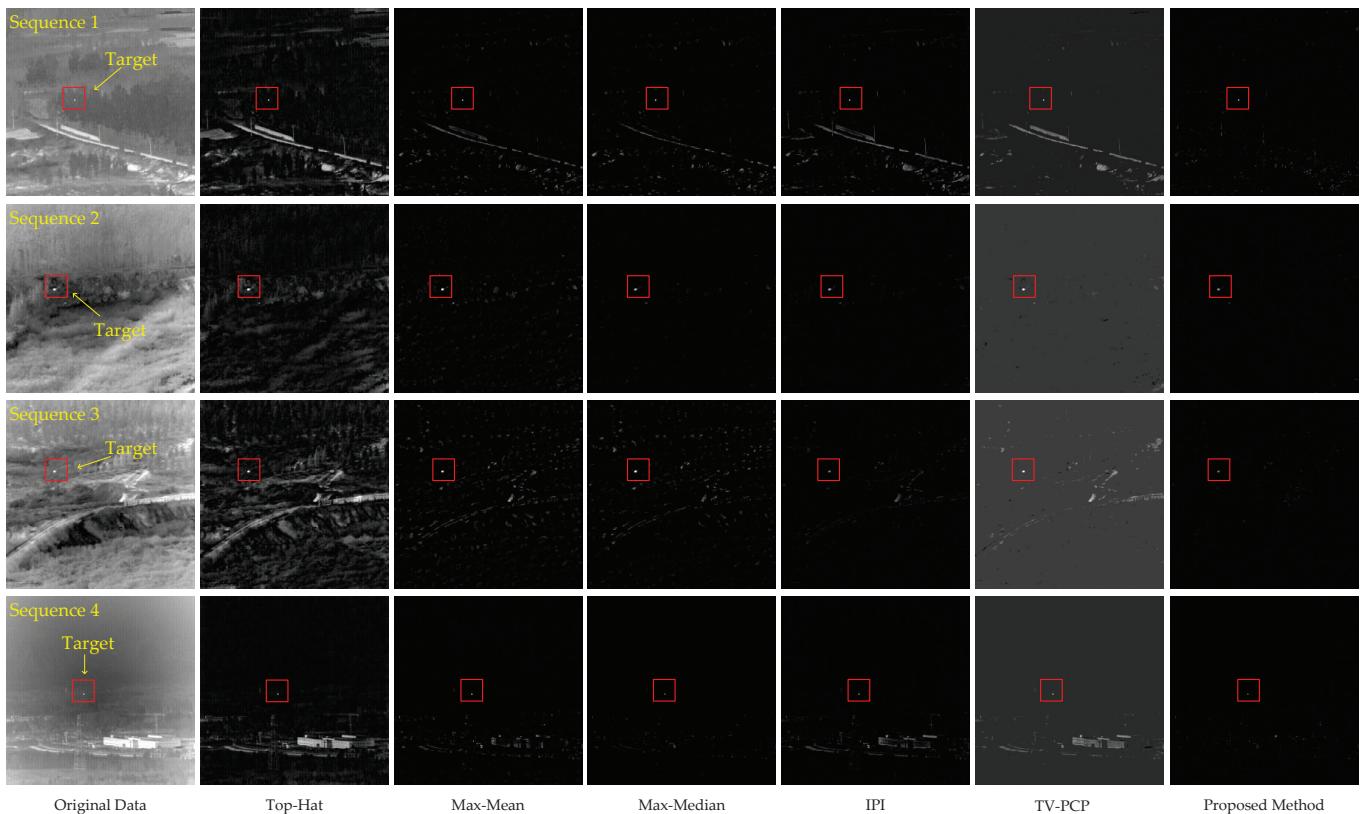
**Figure 6.** Cont.



**Figure 6.** The receiver operating characteristic (ROC) curves of different sizes for four real image sequences. (a) Sequence 1. (b) Sequence 2. (c) Sequence 3. (d) Sequence 4.

#### 4.3. Contrast Experiments

In order to verify the effectiveness of the proposed method, baseline methods including Top-Hat, Max-Mean, Max-Median, IPI, and TV-PCP are selected for comparison. We select a representative image from each of the four sequences and process them with the six methods. Firstly, we compare the background suppression performance. The detection results of the representative images of four sequences using different methods are shown in Figure 7.



**Figure 7.** The representative images of the four real image sequences and the corresponding processed target images of different methods. From left to right are the original data, the processing results of Top-Hat, Max-Mean, Max-Median, IPI, TV-PCP, and the proposed method.

It can be seen in Figure 7 that Top-Hat performs very poorly in the face of four sequences with complex backgrounds. It can enhance the target. However, because Top-Hat relies on the assumption of the size of target, it cannot distinguish the target from the background clutter well. At the same time, it can not screen noise that is close to the target size. Max-Mean and Max-Median can enhance the target, but the ability to suppress background clutter is weak. Max-Median is more robust than Max-Mean. According to the results of Seq.2 and Seq.4, it can be seen that Max-Median has a better inhibition effect on strong edges compared to Max-Mean. The common reason for why Top-Hat, Max-Mean, and Max-Median do not perform well is that they need to make preset assumptions about the target information. The last IPI, TV-PCP, and our method are all based on low-rank sparse matrix recovery. IPI enhances the background sparsity by constructing the patch-image. Therefore, the result is good in the face of the sequence (Seq.2), with a relatively smooth background transition. When the background is very complex with a lot of highlighted interference, the correlation between adjacent pixels in the image background is not strong. This results in a worse IPI performance in Seq.1, Seq.3, and Seq.4. TV-PCP brings TV regularization term to the traditional PCP model. It can better deal with non-uniform images. However, TV-PCP performs poorly when faced with strong edges and highlighted backgrounds such as IPI. This is because they are all based on the low-rank property of the constructed patch-image. In contrast, the proposed method in this paper is better at suppressing the nontarget components. Our method has the least residual background clutter in the experimental images, as it makes full use of the self-similarity of whole image. Although there is a very small amount of clutter because of the complex background, we can further use threshold segmentation to obtain small targets. Based on the above comparisons, the proposed method achieves the best background suppression effect among the six tested methods.

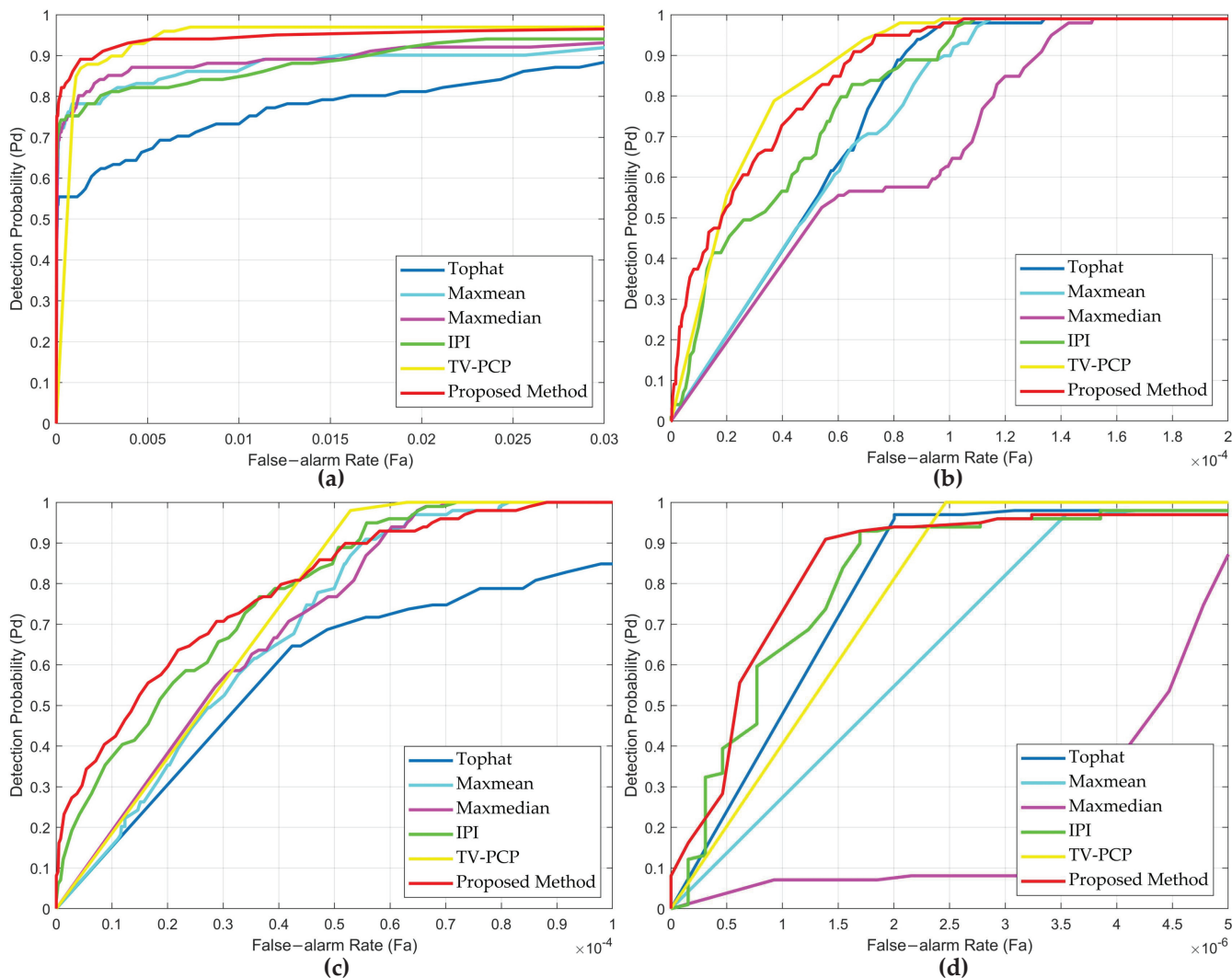
We calculate the metrics values without threshold segmentation to further evaluate these methods, and higher values denote better performance. The performance comparison of the six methods is shown in Table 3. Table 3 shows that the proposed method yields the best or second-best results for most of the image sequences, compared with the other methods. This suggests that the proposed method in this paper outperforms the baseline methods in terms of background suppression and target enhancement.

**Table 3.** Quantitative comparison of six methods for the images of Sequence 1–4.

	Top-Hat	Max-Mean	Max-Median	IPI	TV-PCP	Proposed Method
10th frame of Seq.1						
BSF	1.9150	2.2114	2.2241	2.7063	2.8176	<b>2.9542</b>
LSNRG	1.8555	2.4063	2.2001	2.8511	2.3334	<b>3.3821</b>
SCRG	1.8726	3.6916	4.1848	4.5444	4.6315	<b>4.8301</b>
10th frame of Seq.2						
BSF	1.4953	1.7571	2.1528	1.9926	1.8748	<b>2.2620</b>
LSNRG	1.0046	0.8935	<b>1.0860</b>	1.0236	1.0116	1.0704
SCRG	1.3216	2.3538	2.8490	2.6366	<b>3.1674</b>	2.8290
10th frame of Seq.3						
BSF	1.4840	1.5812	1.5812	1.8878	1.3768	<b>1.9611</b>
LSNRG	1.1988	1.3980	1.1080	1.4019	1.2662	<b>1.4470</b>
SCRG	1.0564	2.3859	2.4129	3.1280	2.8973	<b>3.1363</b>
10th frame of Seq.4						
BSF	1.7803	2.6111	2.8659	<b>3.0519</b>	2.3717	2.9747
LSNRG	2.0182	<b>3.0833</b>	2.4292	2.5882	1.4800	2.5807
SCRG	1.6963	3.9429	4.0535	<b>4.3271</b>	4.1347	4.2013

Finally, to further reveal the advantage of the proposed method, we display the ROC curve to reflect the methods' performance more intuitively. The ROC curves are shown in Figure 8. As seen in Figure 8, the detection rate of the proposed method is more

prominent. On the whole, it has a higher detection probability under the same false-alarm rate. In Seq.2, IPI, TV-PCP, and our method have achieved good results due to the relatively smooth background. TV-PCP performs well in Seq.1. However, the false-alarm rate was higher in Seq.3 and Seq.4, with more highlighted points. IPI also performs well in Seq.3 and Seq.4, but the detection rate is not as good as our method due to the poor ability of the algorithm to deal with the non-uniform background. Top-Hat, Max-Median, and Max-Mean all had unsatisfactory results due to their dependence on the assumption of target information. In particular, Max-Median does badly in Seq.4. Our method shows better detection performance when dealing with complex backgrounds.



**Figure 8.** The receiver operating characteristic (ROC) curves of different methods for four real image sequences. (a) Sequence 1. (b) Sequence 2. (c) Sequence 3. (d) Sequence 4.

## 5. Conclusions

This paper proposed an infrared small target detection method based on block-matching in images. This method takes full advantage of non-local similarity to construct an infrared block-matching model. Then, the small target detection task is transformed into an optimization problem of recovering low-rank and sparse matrices. This problem can be effectively solved via IALM. Our method reduces the dependence on the correlation between adjacent pixels in the image background. When the background of images has strong edges and highlighted clutter, the block-matching model has better low-rank property. This block-matching model can significantly suppress the background clutter and

noise to improve the performance of small target detection. The experimental results show that our proposed method not only obtains the clearest separated target images compared with the baseline methods, but also significantly improves quantitative parameters such as LSNRG, SCRG, BSF, and the ROC curve. In the future, we will study the feasibility of the block-matching model in 3D or multi-dimensional space. We will also learn the matrix recovery algorithm applicable to the proposed method further to improve the speed and applicability.

**Author Contributions:** Conceptualization, Q.Y. and T.C.; methodology, Y.M.; software, Y.M.; validation, Y.M.; investigation, Q.Y.; resources, Q.Y. and T.C.; writing—original draft preparation, Y.M.; writing—review and editing, Q.Y. and T.C.; funding acquisition, T.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Open Fund of Laboratory of Pinghu.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hong, R.; Xiang, C.; Liu, H.; Glowacz, A.; Pan, W. Visualizing the knowledge structure and research evolution of infrared detection technology studies. *Information* **2019**, *10*, 227. [[CrossRef](#)]
2. Rawat, S.S.; Verma, S.K.; Kumar, Y. Review on recent development in infrared small target detection algorithms. *Procedia Comput. Sci.* **2020**, *167*, 2496–2505. [[CrossRef](#)]
3. Guan, X.; Zhang, L.; Huang, S.; Peng, Z. Infrared small target detection via non-convex tensor rank surrogate joint local contrast energy. *Remote Sens.* **2020**, *12*, 1520. [[CrossRef](#)]
4. Johnston, L.A.; Krishnamurthy, V. Performance analysis of a dynamic programming track before detect algorithm. *IEEE Trans. Aerosp. Electron. Syst.* **2002**, *38*, 228–242. [[CrossRef](#)]
5. Salmond, D.J.; Birch, H. A particle filter for track-before-detect. In Proceedings of the 2001 American Control Conference (Cat. No. 01CH37148), Arlington, VA, USA, 25–27 June 2001; Volume 5, pp. 3755–3760.
6. Reed, I.S.; Gagliardi, R.M.; Stotts, L.B. Optical moving target detection with 3-D matched filtering. *IEEE Trans. Aerosp. Electron. Syst.* **1988**, *24*, 327–336. [[CrossRef](#)]
7. Dong, L.; Wang, B.; Zhao, M.; Xu, W. Robust infrared maritime target detection based on visual attention and spatiotemporal filtering. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3037–3050. [[CrossRef](#)]
8. Huang, K.; Tao, D.; Yuan, Y.; Li, X.; Tan, T. Biologically inspired features for scene classification in video surveillance. *IEEE Trans. Syst. Man Cybern. Part B* **2010**, *41*, 307–313. [[CrossRef](#)]
9. Chen, C.P.; Li, H.; Wei, Y.; Xia, T.; Tang, Y.Y. A local contrast method for small infrared target detection. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 574–581. [[CrossRef](#)]
10. Han, J.; Ma, Y.; Zhou, B.; Fan, F.; Liang, K.; Fang, Y. A robust infrared small target detection algorithm based on human visual system. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 2168–2172.
11. Qin, Y.; Li, B. Effective infrared small target detection utilizing a novel local contrast method. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1890–1894. [[CrossRef](#)]
12. Wei, Y.; You, X.; Li, H. Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recognit.* **2016**, *58*, 216–226. [[CrossRef](#)]
13. Deng, H.; Sun, X.; Liu, M.; Ye, C.; Zhou, X. Small infrared target detection based on weighted local difference measure. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 4204–4214. [[CrossRef](#)]
14. Han, J.; Liang, K.; Zhou, B.; Zhu, X.; Zhao, J.; Zhao, L. Infrared small target detection utilizing the multiscale relative local contrast measure. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 612–616. [[CrossRef](#)]
15. Guan, X.; Peng, Z.; Huang, S.; Chen, Y. Gaussian scale-space enhanced local contrast measure for small infrared target detection. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 327–331. [[CrossRef](#)]
16. Han, J.; Moradi, S.; Faramarzi, I.; Zhang, H.; Zhao, Q.; Zhang, X.; Li, N. Infrared small target detection based on the weighted strengthened local contrast measure. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1670–1674. [[CrossRef](#)]
17. Deshpande, S.D.; Meng, H.E.; Venkateswarlu, R.; Chan, P. Max-mean and max-median filters for detection of small targets. *Proc. SPIE* **1999**, *3809*, 74–83.
18. Tom, V.T.; Peli, T.; Leung, M.; Bondaryk, J.E. Morphology-based algorithm for point target detection in infrared backgrounds. *Proc. SPIE* **1993**, *1954*, 2–11.

19. Bai, X.; Zhou, F. Analysis of new top-hat transformation and the application for infrared dim small target detection. *Pattern Recognit.* **2010**, *43*, 2145–2156. [[CrossRef](#)]
20. Zhu, H.; Liu, S.; Deng, L.; Li, Y.; Xiao, F. Infrared small target detection via low-rank tensor completion with top-hat regularization. *IEEE Trans. Geosci. Remote Sens.* **2019**, *58*, 1004–1016. [[CrossRef](#)]
21. Hadhoud, M.M.; Thomas, D.W. The two-dimensional adaptive LMS (TDLMS) algorithm. *IEEE Trans. Circuits Syst.* **1988**, *35*, 485–494. [[CrossRef](#)]
22. Bae, T.W.; Sohng, K.I. Small target detection using bilateral filter based on edge component. *J. Infrared Millim. Terahertz Waves Vol.* **2010**, *31*, 735–743. [[CrossRef](#)]
23. Gao, C.; Meng, D.; Yang, Y.; Wang, Y.; Zhou, X.; Hauptmann, A. G. Infrared patch-image model for small target detection in a single image. *IEEE Trans. Image Process.* **2013**, *22*, 4996–5009. [[CrossRef](#)] [[PubMed](#)]
24. Dai, Y.; Wu, Y.; Song, Y. Infrared small target and background separation via column-wise weighted robust principal component analysis. *Infrared Phys. Technol.* **2016**, *77*, 421–430. [[CrossRef](#)]
25. Dai, Y.; Wu, Y. Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3752–3767. [[CrossRef](#)]
26. Dai, Y.; Wu, Y.; Song, Y.; Guo, J. Non-negative infrared patch-image model Robust target-background separation via partial sum minimization of singular values. The title of the cited article. *Infrared Phys. Technol.* **2017**, *81*, 182–194. [[CrossRef](#)]
27. He, Y.; Li, M.; Zhang, J.; An, Q. Small infrared target detection based on low-rank and sparse representation. *Infrared Phys. Technol.* **2015**, *68*, 98–109. [[CrossRef](#)]
28. Wang, C.; Qin, S. Adaptive detection method of infrared small target based on target-background separation via robust principal component analysis. *Infrared Phys. Technol.* **2015**, *69*, 123–135. [[CrossRef](#)]
29. Wang, X.; Peng, Z.; Kong, D.; Zhang, P.; He, Y. Infrared dim target detection based on total variation regularization and principal component pursuit. *Image Vis. Comput.* **2017**, *63*, 1–9. [[CrossRef](#)]
30. Zhang, L.; Peng, L.; Zhang, T.; Cao, S.; Peng, Z. Infrared small target detection via non-convex rank approximation minimization joint  $l_{2,1}$  norm. *Remote Sens.* **2018**, *10*, 1821. [[CrossRef](#)]
31. Rawat, S.S.; Alghamdi, S.; Kumar, G.; Alotaibi, Y.; Khalaf, O.I.; Verma, L.P. Infrared small target detection based on partial sum minimization and total variation. *Mathematics* **2022**, *10*, 671. [[CrossRef](#)]
32. Zontak, M.; Irani, M. Internal statistics of a single natural image. In Proceedings of the CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011; pp. 977–984.
33. Buades, A.; Coll, B.; Morel, J.M. A non-local algorithm for image denoising. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 20–25 June 2005; pp. 60–65.
34. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.* **2007**, *16*, 2080–2095. [[CrossRef](#)]
35. Mairal, J.; Bach, F.; Ponce, J.; Sapiro, G.; Zisserman, A. Non-local sparse models for image restoration. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 29 September–2 October 2009; pp. 2272–2279.
36. Candès, E.J.; Li, X.; Ma, Y.; Wright, J. Robust principal component analysis? *J. ACM (JACM)* **2011**, *58*, 1–37. [[CrossRef](#)]
37. Hstad, J. Tensor rank is np-complete. *J. Algorithms* **1990**, *31*, 644–654. [[CrossRef](#)]
38. Lin, Z.; Ganesh, A.; Wright, J.; Wu, L.; Chen, M.; Ma, Y. Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix. In Proceedings of the Computational Advances in Multi-Sensor Adaptive Processing, Aruba, Dutch Antilles, 13–16 December 2009.
39. Lin, Z.; Chen, M.; Ma, Y. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *arXiv* **2010**, arXiv:1009.5055v3.
40. Yuan, X.; Yang, J. Sparse and low-rank matrix decomposition via alternating direction methods. *Preprint* **2009**, *12*, 1–16.
41. Zhou, Z.; Li, X.; Wright, J.; Candès, E.; Ma, Y. Stable principal component pursuit. In Proceedings of the 2010 IEEE International Symposium on Information Theory, Austin, TX, USA, 13–18 June 2010; pp. 1518–1522.
42. Hui, B.; Song, Z.; Fan, H.; Zhong, P.; Hu, W.; Zhang, X.; Lin, J.; Su, H.; Jin, W.; Zhang, Y.; et al. Infrared dim and small aircraft target detection and tracking dataset from infrared image in ground/air background. *Chin. Sci. Data* **2020**, *5*, 286–297.