Article

Infrared Dim Small Target Detection Based on Nonconvex Constraint with L1–L2 Norm and Total Variation

Yu Shao [1,2], Xu Kang [1,2], Mingyang Ma [1], Cheng Chen [1], Sun He [1] and Dejiang Wang [1,*]

1   Key Laboratory of Airborne Optical Imaging and Measurement, Changchun Institute of Optics,
    Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China
2   University of Chinese Academy of Sciences, Beijing 100049, China
*   Correspondence: wangdj04@ciomp.ac.cn; Tel.: +86-0431-8670-8039

Abstract: Infrared dim small target detection has received a lot of attention, because it is a crucial component of the IR search and track systems (IRST). The robust principal component analysis (RPCA) is a common detection framework, which works with poor performance with complex background edges and sparse clutters due to the inappropriate approximation of sparse items. A nonconvex constraint detection method based on the difference between the L1 and L2 (L1–L2) norm and total variation (TV) is presented. The L1–L2 norm is a more accurate sparse item approximation of L0 norm, which can achieve a better description of the sparse item to separate the target from the complex backgrounds. Then, the total variation norm is conducted on the target image to suppress the sparse clutters. The new model is solved using the alternating direction method of multipliers (ADMM) method. Then, the subproblems in the model are tackled by the difference of convex algorithm (DCA) and the Newton conjugate gradient (Newton-CG) solving L1–L2 norm and TV norm, respectively. In the experiment, we conducted experiments on multiple and single target datasets, and the proposed model outperforms the state-of-the-art (SOTA) methods in terms of background suppression and robustness to accurately detect the target. It can achieve a higher true position rate (TPR) with a low false position rate (FPR).

Keywords: L1–L2 norm; nonconvex optimization; alternating direction method of multipliers; infrared small target detection

1. Introduction

The Infrared search and track systems (IRST), which outperform conventional radar in early-warning, precision guidance, and surveillance applications [1], rely heavily on the infrared small target detection task. On the one hand, the target, is far from the sensor, so it takes up few pixels in the IR image. Additionally, due to the same reason, the radiation attenuation of the target is serious, resulting in a low signal to noise ratio (SNR) of the target. On the other hand, the target is obscured by the intricate background and noise. As a result, it is challenging to detect dim and small targets without shape or texture information.

Decades ago, the RPCA framework was proposed for small target detection, and data optimization-based methods have received a lot of attention because of their excellent detection performance. The low rank and sparse characteristics are respectively represented by the nuclear norm and the L0 norm. The L1 norm always replaces the L0 norm as the sparse measure; however, the loose approximation may cause under punishment of the sparse item [2]. In addition, the target image will remain at the edge clutters and sparse clutters, inducing high false detection rates in the complex background.

Therefore, to obtain more accuracy and robust detection, the improvement of the sparse item was conducted. In [3], the Lp norm was considered a better approximation to the L0 norm, which can recover the sparse signals better when $0 < p < 1$. For the RPCA framework, the Lp norm is supposed to become a sparser component, and, if the value of p

is smaller, the solution is sparser. Thus, the nonconvex optimization of the Lp norm can achieve superior performance for the small target detection tasks. However, the setting of p makes a decisive role in the robustness of the Lp norm model. The value of p is close to 1, and the optimization results are similar to L1. When p is close to 0, the solution is much sparser. The false alarms will be high when the initial value of p is high in the complicated background, and the low value of p may cause missed detection in the uniform background. Therefore, the model based on the Lp norm is inadequate to apply in multifarious circumstances with IR small target detection tasks.

Furthermore, the SOTA methods based on matrix recovery can well achieve detection with homogenous scenes, while they are still affected by the sparse clutters, such as cloud edges, sun flash, and sea clutters. The sparse constraint is poor robustness to distinguish the target from the clutters and a high false alarm due to the heavy residual in the target image. TV norm [4] has already been applied to the yield of image denoise. Ideally, the target image is clean, except for the target after the separation, but the sparse constraint cannot work well with the sparse clutters, and the target image will remain as little sparse clutters to interfere with the detection task. Accordingly, we can employ the TV norm on the target image to suppress the sparse clutters in the target image.

Inspired by this, a parameter free sparse constraint model is supposed to be employed to detect the infrared small target and we propose a difference of L1 and L2 norm add total variation regularization on the target image for the detection. The method is good at complex background resistance, high detection precision, and sparse norm parameter free.

The novelties of this paper are as follows:

(1)   The difference of L1 and L2 norms is applied in the field of the IR dim and small target detection, which is parameter free, and the nonconvex optimization of the L1-L2 norm can achieve sparser target image restoration;

(2)   A total variation (TV) regularization is conducted on the sparse target image, which is to constrain the sparse clutters and decrease the residuals in the target image;

(3)   The difference between convex algorithm (DCA) [5] and Newton conjugate gradient (CG) [6] methods based on the alternating direction method of multipliers (ADMM) are presented to solve the nonconvex model. DCA is used to solve the difference between L1 and L2 norms. In addition, the CG is supposed to solve the total variation regularization, which converges quickly.

The following is the structure of the rest of this paper: Section 2 summarizes the relevant research of single-frame infrared small target detection and briefly introduces the current problems of SOTA; Section 3 introduces the proposed model and its solution. Section 4 introduces the comparison algorithm, test data and experimental results on multiple real infrared sequences and single frames. Section 5 discusses the improvement effect of the proposed algorithm compared with SOTA algorithm. The conclusion is provided in Section 6.

2. Related Work

DBT systems are better in real-time and can reject false alarms caused by single frame detection results thanks to a tracking algorithm. Therefore, the robustness of the single-frame detection algorithm is crucial for the performance of the DBT systems [7]. There are numerous researchers dedicated to single-frame detection approaches, which are roughly divided into four categories.

2.1. The Background Suppression-Based Method

This kind of method based on the assumption of the infrared image is that the target is isolated from a relatively continuous background and detection of the target via the suppression of the background. Then max-mean\max-median method [8] and the morphology opening method [9] are proposed to remove the background. Although these methods have the advantage of a low complexity calculation, the estimation accuracy will be greatly affected under the complex background. To solve the above problems, advanced technol-

ogy based on background suppression is proposed to improve the target and eliminate the background at the same time. Laplacian of Gaussian (LoG) scale-space [10], the average absolute gray difference (AAGD) [11], and the facet model [12] were proposed to highlight of small and weak targets disturbed by background clutter. However, such methods not only enhance the target but also highlight the edge and texture clutter in the background.

2.2. The Human Visual System-Based Method

The human visual system-based method is used to detect targets by the characteristics of local saliency of the target. This is motivated by the human visual system, which shows good robustness when dealing with the target detection task. The local contrast method (LCM) was introduced by Chen et al. [13]. The feature of local contrast calculates with a nested structure, which can cause the block effect and the background suppression to be inferior. Han J et al. [14] improved the local contrast method for high detection speed based on the subblock division of the IR image, and the relative local contrast measure (RLCM) [15] was introduced with a difference-ratio form to calculate the local contrast. Wei Y et al. [16] presented a multiscale patch-based contrast method (MPCM) that suppresses background clutter by utilizing the difference between the target patches and the background patches in nine directions and in more than one scale. The weighted strengthened local contrast measure (WSLCM) [17] can better suppress the background by designing the weight function, which fully utilizes the characteristics of background, target, and the dissimilarity between the target and background. Ye S et al. [18] advanced the high-boost-based local contrast algorithm in multiscale (HB-MLCM) to strengthen the dim target embedded in the heavy background clutters with the improved high-boost filter. Guan X et al. [19] proposed enhanced local contrast in Gaussian scale space (GSS-ELCM) to solve the problem, which was the changing of the target size. LCM and its extended methods are widely used, as is their low complexity computation. Obviously, these methods mainly take advantage of the local brightness being higher than the surroundings. The prior knowledge is only in the spatial space, and the character is too simple and singular to make the method robust in the complex background.

2.3. The Data Optimization Based-Method

Different from the conventional detection, the optimization-based methods consider the IR image, which is composed of the background component, target component, and noise component. Meanwhile, these methods always assume that the background component is low rank when the target component is sparse, and the noise component is the Gaussian distribution with zero mean. Afterward, the dim small target detection task is converted into a robust principal component analysis (RPCA) [20] optimization problem. The low rank characteristic of the background image has been improved with the proposal of the infrared patch image (IPI) model by Gao C et al. [21], and the nuclear norm characterize the low rank of the background when the L1 norm is used to characterize the sparsity of the target image. Wang X et al. [22] presented an RPCA of joint total variational regularization on the background image of reducing the false alarm caused by edges in the background. Dai Y et al. [23] also suppressed the background edges with the local steering kernel weighted IPI (WIPI) and increased penalties for background edges. At the same time, a reweighted infrared patch tensor (RIPT) [24] model was introduced, which showed a new patch model adopting the tensor structure. Guan X et al. [2] proposed a tensor model to improve the IPT with a nonconvex item, which was the replacement of tensor nuclear norm (TNN) [25], and the local contrast energy was applied in the model. Furthermore, the Laplace function was used to approximate the L0 norm. Zhang L et al. [26] introduced a rank approximation minimization, which is a nonconvex regularization term. Additionally, the L2,1 norm was introduced to constrained structure noise. Zhang T et al. [27] presented the Lp norm [28], which is a nonconvex item that can strengthen the sparse item constraint. This category can achieve a better detection due to the reasonable assumption and model structures.

2.4. The Deep Learning-Based Method

This type of algorithm is data driven, and many of the samples are trained to extract deep features of the target and use a classifier to detect the target. The ability of automatic learning target features of deep learning makes it develop rapidly in the infrared dim small target detection field. The convolutional neural network (CNN) [29] was used to train the filters to classify the background and target [30]. Then, the infrared small target detection module was designed to add to the CNN to acquire a better detection performance [31]. The image filtering module was also the same. Dai Y et al. [32] presented an attentional local contrast network (ALCNet), which incorporated a local contrast module and a bottom-up attention module. Hou Q et al. [33] proposed infrared small target detection, U-Net (ISTDU-Net), based on a U-shaped structure. By suppressing the background with a fully connected layer in the jump connection and improving the characterization capability of small targets, ISTDU-Net achieves excellent detection performance. Even though a deep learning-based method can perform well in detection with less assumptions, these methods need many samples containing various scenes to train, and there are few open access and appropriate datasets of the infrared small target.

Table 1 summaries the various types of infrared small target detection methods and their typical methods, which are mentioned in the related work.

Table 1. Summary of typical relevant algorithms.

Kind of Methods	Typical Algorithms
The background suppression-based method	max-mean\max-median, morphology opening method, LoG, AAGD, facet model
The human visual system-based method	LCM, RLCM, MPCM, WSLCM, HB-MLCM, GSS-ELCM
The data optimization-based method	IPI, WIPI, RIPT,
The deep learning-based method	ALCNet, ISTDU-Net

3. Methodology

In this section, the difference between L1 and L2 norms is introduced and explain the advantages of the L1–L2 metric in the application of IR small target detection to approximate the L0 norm. Then, a novel L1–L2 norm add total variation regularization method for IR small target detection is presented. Finally, the solution of the nonconvex model based on ADMM is shown.

As mentioned above, Gao et al. presented the IPI model formulated as Equation (1). In the equation, the target image T and background image B have the property sparse and low rank, respectively. The N is the corresponding patch-images of the random noise. In addition, the original detection model is written as Equation (2), and the sparse characteristic is formulated with $\|T\|_0$, where $\| \bullet \|_0$ stands for the L0 norm which is the number of nonzero elements. The low rank characteristic is formulated as $rank(B)$.

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

$$\min_{B,T} rank(B) + \lambda\|T\|_0 \quad \text{s.t. } D = B + T, \tag{2}$$

In Formulation (3), there is the nuclear norm of the background patch-image, and the target patch-image is the L1 norm. The nuclear norm is the total of the matrix's singular value, and the L1 norm is the total of each matrix element's absolute values. The L0-norm is to acquire the sparest solution; however, minimizing the L0 norm optimization problem is an NP hard problem. The L1 norm is always regarded as the convex approximation of the L0 norm. $\| \bullet \|_1$ stands for the L1 norm which is the sum of absolute values of all elements.

$$\min_{B,T} \|B\|_* + \lambda \|T\|_1 \quad \text{s.t. } D = B + T, \tag{3}$$

3.1. Enhanced Sparsity of L1–L2 Metric

In [34], a convex relaxation of L1 to L0 has attracted extensive attention in IR small target detection. The L1 norm could not guarantee that the optimal solution is sparse since the intersection of affine subspace of the L1 norm and a level set is possibly not the unique point, which means that all the points on the segment are the optimal solution [35]. Therefore, L1 norm is too loose to constrain the sparse component and leads to the residual remained in the target image.

The nonconvex measure Lp norm is studied in [35,36] for replacing the L1 norm. Due to its curved level set, the defect of the L1 norm can be well avoided. The same is true with L1–L2 norm. However, the Lp norm is a non-Lipschitz continuous metric, and an additional smoothing operation is supposed to conduct in minimization, avoiding division by zero and enhancing sparsity [36]. Although the nonconvex function overcomes the defect of the L1 norm, the solving process of the nonconvex function is more challenging, and a prior unknown parameter p is extremely important for solving the model.

Figure 1 illustrated the level set of L1, Lp, and L1–L2. In the geometrical aspect, optimization of sparse functions based on equality constraint is supposed to obtain an interception of an affine subspace, and the intersection of the level set of the subspace and the plane is closest to zero [27].
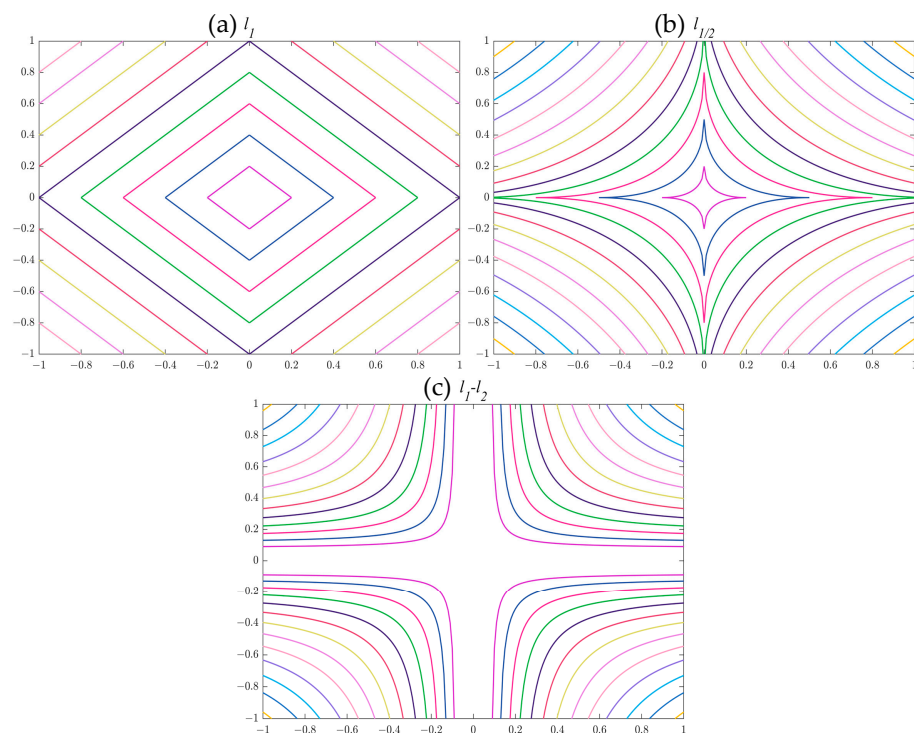

Figure 1. level set of sparsity metric (**a**) L1, (**b**) L1/2, (**c**) L1–L2.

In this paper, the L1–L2 metric is introduced for sparse target image recovery, which is nonconvex yet Lipschitz continuous. The L1–L2 metric has been applied in compress sensing [35] and hyperspectral unmixing [37] to recover the sparse signal. For a matrix X, the L1–L2 metric is given by $\|X\|_{1-2} \triangleq \|X\|_1 - \|X\|_2$ where $\| X \|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$. Compared with the Lp norm, the L1–L2 metric is parameter free and has similar convergence rates.

3.2. Total Variation Regularization

It is obvious that the IR image has poor image quality due to the abundant noise with sparse characteristics. At the same time, the building and cloud edges fluctuate greatly and will make part of the edges as the sparse components. Unfortunately, the detection of the target is also affected by these. The sparse clutters in the IR image will remain in the target image causing the false alarm.

The TV norm is successfully applied in the image denoise yield. The denoise model is formulated as follows:

$$X = \underset{X}{\mathrm{argmin}}\|Y - X\|_2^2 + \lambda\|X\|_{TV} \tag{4}$$

where Y and X represent the observe image and clean image, respectively. λ is an adjustment coefficient. $\| \bullet \|_{TV}$ represents the TV norm, which can be signified as Formulation (5).

$$X \in \mathbb{R}^{m \times n}, \ \|X\|_{TV} = \sum_{i,j}^{m,n} \sqrt{|x_{i+1,m} - x_{i,m}|^2 + |x_{n,j+1} - x_{n,j}|^2} = \sum_{i,j}^{m,n} \left\|grad(x)_{i,j}\right\| \tag{5}$$

In Equation (5), m and n are the row and column numbers of the clean image and $x_{i,j}$ is the pixel value of an image, which is the i-th row and j-th column pixel. $grad(x)_{i,j}$ is the gradient of pixel at (i, j) position in the clean image.

Therefore, the TV norm is supposed to constrain the target image, since the enhanced sparsity constraint will leave little noise in the target image occupied by small pixels.

3.3. Proposed Method

As Equation (1) describes, the infrared image is regarded as a linear model, which is the sum of the target component, background component, and noise component, when the noise is approximately regarded as additive. This approach is extensively used in IR small target detection [38–40]. The target image can be recovered based on the sparse property of the target image and low rank property of the background image. In [21], the original image is reconstructed with the patch images, which are obtained by the sliding window on the original image. The operation transforms the original image into a patch image and can enhance the low rank of the background.

On this basis, the L1–L2 metric constraint was employed on the target image for more sparsity solution. The objective function is changed as Equation (6).

$$\underset{B,T}{\min}\|B\|_* + \lambda\|T\|_{1-2} \quad \text{s.t. } D = B + T \tag{6}$$

For enhancing the robustness of the algorithm to resist the structure noise, a total variation regularized constraint of target image was applied. Then, we proposed an infrared small target detection method based on nonconvex constraint based on L1–L2 norm and total variation. The final objective function is rewritten as Equation (7).

$$\begin{aligned}\underset{B,T}{\min}\|B\|_* + \lambda_1\|T\|_{1-2} + \lambda_2\|F\|_{TV} \\ \text{s.t. } D = B + T, \quad T = F\end{aligned} \tag{7}$$

where $\| \bullet \|_{1-2}$ is denoted as $\| \bullet \|_1 - \| \bullet \|_2$.

3.4. Solution of the Proposed Model

In this section, the objective function is solved with the ADMM method, and the solution is shown. The equation constraint optimization problem is converted into an unconstrained optimization problem. Thus, Equation (7) can be rewritten as the augmented Lagrangian function as Equation (8).

$$L(D, B, T, F, \gamma_1, \gamma_2, \rho) = \|B\|_* + \lambda_1 \|T\|_{1-2} + \lambda_2 \|F\|_{TV} + \langle \gamma_1, F - T \rangle + \langle \gamma_2, D - B - T \rangle + \frac{\rho}{2} \left(\|F - T\|_F^2 + \|D - B - T\|_F^2 \right) \quad (8)$$

where $\langle \bullet \rangle$ denotes the inner product of two matrices, $\|\bullet\|_F$ is the Frobenius norm equaling to the square root of the square sum of matrix elements, γ_1 and γ_2 are the Lagrangian multiplier matrixes, and ρ is a penalty factor.

Based on the ADMM method, an alternative iteration is utilized to minimize the Lagrange function Then, the optimization problem is divided into three subproblems, and we can solve each subproblem independently. Moreover, the solution to subproblems is shown separately.

(a)　subproblem of B

The iteration function of B in the $(k + 1)$ step is as follows:

$$\begin{aligned} B^{k+1} &= \underset{B}{\arg\min} L(B, T^k, \gamma_2^k) \\ &= \underset{B}{\arg\min} \|B\|_* + \left\langle \gamma_2^k, D - B - T^k \right\rangle + \frac{\rho}{2} \left\| D - B - T^k \right\|_F^2 \\ &= \underset{B}{\arg\min} \|B\|_* + \frac{\rho}{2} \left\| D - B - T^k + \frac{\gamma_2^k}{\rho} \right\|_F^2 \end{aligned} \quad (9)$$

The above formulation has a close-form solution obtained by a singular value thresholding shrinkage operator [41].

$$B^{k+1} = U^k S_{1/\rho}(\Sigma) V^{k^T} \quad (10)$$

where U, Σ, V are acquired with the singular value decomposition (SVD) of Σ and $S_{1/\rho}[\bullet]$ is the soft thresholding operator, which is written in Equation (11).

$$S_\mu[x] = \begin{cases} x - \mu & if \ x > \mu \\ x + \mu & if \ x > \mu \\ 0 & others \end{cases} \quad (11)$$

(b)　subproblem of T

The iteration function of T in the $(k + 1)$ step is as follows:

$$\begin{aligned} T^{k+1} &= \underset{T}{\arg\min} L(B^{k+1}, T, F^k, \gamma_1^k, \gamma_2^k) \\ &= \underset{T}{\arg\min} \lambda_1 \|T\|_{1-2} + \left\langle \gamma_1^k, F^k - T \right\rangle + \left\langle \gamma_2^k, D - B^{k+1} - T \right\rangle + \frac{\rho}{2} \left(\left\| F^k - T \right\|_F^2 + \left\| D - B^{k+1} - T \right\|_F^2 \right) \end{aligned} \quad (12)$$

However, the above formulation is a nonconvex function, due to the $\|T\|_{1-2}$ item being nonconvex. As mentioned in the above section, $\|X\|_{1-2} \triangleq \|X\|_1 - \|X\|_2$, which can be minimized by the DCA [42]. DCA can directly linearize the objective function instead of adding constraints.

Formulation (12) is decomposed into the difference between two functions $\underset{T}{\arg\min} F(T) = G(T) - H(T)$, where

$$\begin{cases} G(T) = \lambda_1 \|T\|_1 + \left\langle \gamma_1^k, F^k - T \right\rangle + \left\langle \gamma_2^k, D - B^{k+1} - T \right\rangle + \frac{\rho}{2} \left(\left\| F^k - T \right\|_F^2 + \left\| D - B^{k+1} - T \right\|_F^2 \right) \\ H(T) = \lambda_1 \|T\|_2 \end{cases} \quad (13)$$

For facilitating the linearization of $H(T)$, we take an approximation of the L1–L2 metric, which is formulated as $\|X\|_{1-2} \triangleq \|X\|_1 - \|X\|_F^2$ so that $H(T) = \lambda_1 \|T\|_F^2$. The linearization of $H(T)$ is shown as follows:

$$H(T) = H(T^k) + \left\langle T - T^k, \partial H(T^k) \right\rangle \quad (14)$$

Then, the iteration solution of T can be resolved after the linearization of $H(T)$. The solution is given as follows:

$$
\begin{aligned}
T^{k+1} &= \operatorname*{argmin}_{T} G(T) - H(T) \\
&= \operatorname*{argmin}_{T} G(T) - H(T^k) - \left\langle T - T^k, \partial H(T^k) \right\rangle \\
&= \operatorname*{argmin}_{T} \lambda_1 \|T\|_1 - 2\lambda_1 tr(T^T T^k) + \left\langle \gamma_1^k, F^k - T \right\rangle + \left\langle \gamma_2^k, D - B^{k+1} - T \right\rangle + \frac{\rho}{2}\left(\left\| F^k - T \right\|_F^2 + \left\| D - B^{k+1} - T \right\|_F^2 \right) \\
&= \operatorname*{argmin}_{T} \lambda_1 \|T\|_1 + \frac{\rho}{2}\left\| \frac{1}{2}\left(F^k + D - B^{k+1} + \frac{\gamma_1^k + \gamma_2^k + 2\lambda_1 T^k}{\rho} \right) - T \right\|_F^2
\end{aligned}
\tag{15}
$$

The above equitation has a closed-form solution [43], which is shown as

$$
S_{\lambda_1/\rho}\left[\frac{1}{2}\left(F^k + D - B^{k+1} + \frac{\gamma_1^k + \gamma_2^k + 2\lambda_1 T^k}{\rho} \right) \right]
\tag{16}
$$

where $S_{\lambda_1/\rho}[\bullet]$ is a soft-thresholding operator defined as Equation (11).

(c)   subproblem of F

The iteration function of F in the $(k+1)$ step is as follows:

$$
\begin{aligned}
F^{k+1} &= \operatorname*{argmin}_{F} L(F, T^{k+1}, \gamma_1^k) \\
&= \operatorname*{argmin}_{F} \lambda_2 \|F\|_{TV} + \left\langle \gamma_1^k, F - T^{k+1} \right\rangle + \frac{\rho}{2}\left\| F - T^{k+1} \right\|_F^2 \\
&= \operatorname*{argmin}_{F} \lambda_2 \|F\|_{TV} + \frac{\rho}{2}\left\| F - T^{k+1} + \frac{\gamma_1^k}{\rho} \right\|_F^2
\end{aligned}
\tag{17}
$$

The Formulation (17) is a convex function and an unconstrained optimization problem so that the minimization for (15) can be easily achieved. Then, the optimization problem is corresponding to acquiring the F, where $\nabla L(F) = 0$. The derivatives of the objective function of F are shown as Equations (18) and (19):

$$
\nabla L(F) = \lambda_2 \nabla(\|F\|_{TV}) + \rho\left(F - T^{k+1} + \frac{\gamma_1^k}{\rho} \right)
\tag{18}
$$

$$
\nabla(\|F\|_{TV}) = div\left(grad(x)_{i,j} / \left\| grad(x)_{i,j} \right\| \right)
\tag{19}
$$

where $div[\bullet]$ represents the divergence operator. To avoid the gradient of the input image being 0, the TV norm is supposed to modulate with a parameter δ. Thus, the TV norm is shown as

$$
X \in \mathbb{R}^{m \times n}, \ \|X\|_{TV} = \sum_{i,j}^{m,n} \sqrt{\left| x_{i+1,m} - x_{i,m} \right|^2 + \left| x_{n,j+1} - x_{n,j} \right|^2 + \delta^2}
\tag{20}
$$

The Newton method is widely used to solve $\nabla L(F) = 0$, The iteration solution is given as follows:

$$
x^{k+1} = x^k - d
\tag{21}
$$

where k is the number of iterations, x^k denotes the k-th iteration result, and d means the descending direction, which is obtained by solving $H_k d = \nabla L(F^k)$ with the Newton method. In addition, H_k is the Hessian matrix at k-th iteration. However, for large scale optimization problems, the computational complexity of the inverse Hessian matrix is high. Then, the CG method is employed to solve the descending direction of the iteration equation. Thus, a solving operator $F_\rho[\bullet]$ is defined as follows to solve (17):

$$
F_{\lambda_2/\rho}\left[T^{k+1} - \frac{\gamma_1^k}{\rho} \right] = \operatorname*{argmin}_{F} \lambda_2 \|F\|_{TV} + \left\| F - T^{k+1} + \frac{\gamma_1^k}{\rho} \right\|
\tag{22}
$$

Thus, the function (22) will be solved by the Newton-CG, which is summarized in Algorithm 1.

The solving process of the nonconvex function (7) combination with the ADMM method and the solution details are summarized in Algorithm 2.

Algorithm 1: Newton-CG algorithm for solving TV norm.

Input: F^k, $temp = T^{k+1} - \gamma_1^k/\rho$,
Output: F^{k+1}
Initialize: $F^k = temp$
While not converged do
Compute $\nabla L\left(F^k\right)$
$\nabla L(F) = \lambda_2 \nabla(\|F\|_{TV}) + \rho\left(F - T^{k+1} + \gamma_1^k/\rho\right)$;
Compute $\nabla^2 L\left(F^k\right)$ (approximate Hessian matrix);
Compute d;
Solving $\nabla^2 L\left(F^k\right)d = \nabla L(F^k)$ with CG method;
Compute F^{k+1}
$F^{k+1} = F^k - d$;
Check the convergence conditions
$\left\|F^{k+1} - F^k\right\| < 10^{-3}$;
Update k
$k = k + 1$;
end

Algorithm 2: ADMM solver to the proposed model.

Input: Patch image D, λ_1
Output: Target image T, Background image B
Initialize: $B^0 = D$, $T^0 = F^0 = \gamma_1^0 = \gamma_2^0 = 0$, $\rho^0 = 1/(5*\text{std}(A(:)))$, $\lambda_2 = 0.02$;
While not converged do
%Update B^{k+1}
$B^{k+1} = U^k S_{1/\rho}(\Sigma)V^{k^T}$;
%Update T^{k+1}
$T^{k+1} = S_{\lambda_1/\rho}\left[\frac{1}{2}\left(F^k + D - B^{k+1} + \frac{\gamma_1^k+\gamma_2^k+2\lambda_1 T^k}{\rho}\right)\right]$;
%Update F^{k+1}
$F^{k+1} = F_{\lambda_2/\rho}\left[T^{k+1} - \frac{\gamma_1^k}{\rho}\right]$;
%Update γ_1^{k+1}, γ_2^{k+1} and ρ^{k+1}
$\gamma_1^{k+1} = \gamma_1^k + \rho^k\left(D - B^{k+1} - T^{k+1}\right)$;
$\gamma_2^{k+1} = \gamma_2^k + \rho^k\left(F^{k+1} - T^{k+1}\right)$;
$\rho^{k+1} = 1.5 * \rho^k$;
% Judge the convergence conditions
$\left\|D - B^{k+1} - T^{k+1}\right\|_F / \|D\|_F < 10^{-7}$;
% Update k
$k = k + 1$;
end

3.5. The Procedure of the Proposed Method

The flow chart of the complete procedure of the proposed method is shown in Figure 2. Additionally, the particular algorithmic procedure can be summed up as follows:

1. Convert the original infrared image into a patch image through a sliding window with a length of *len* and a step of *step* The *len* and *step* value will be discussed in the next section;
2. Parameters initialize of *lambda1*. The influence of the parameters on the experiments is discussed in Section 3;
3. The patch image input Algorithm 1 and target patch image T solves until the iterative convergence. During iteration, the T and F iteration expressions are solved with DCA and Newton-CG methods, respectively;
4. The target patch image is restored with the inverse process of step 1;
5. The target detection utilizes threshold segmentation, and the segmentation is shown as Equation (23), and μ and δ denotes the mean and variance value of the separated target competent.
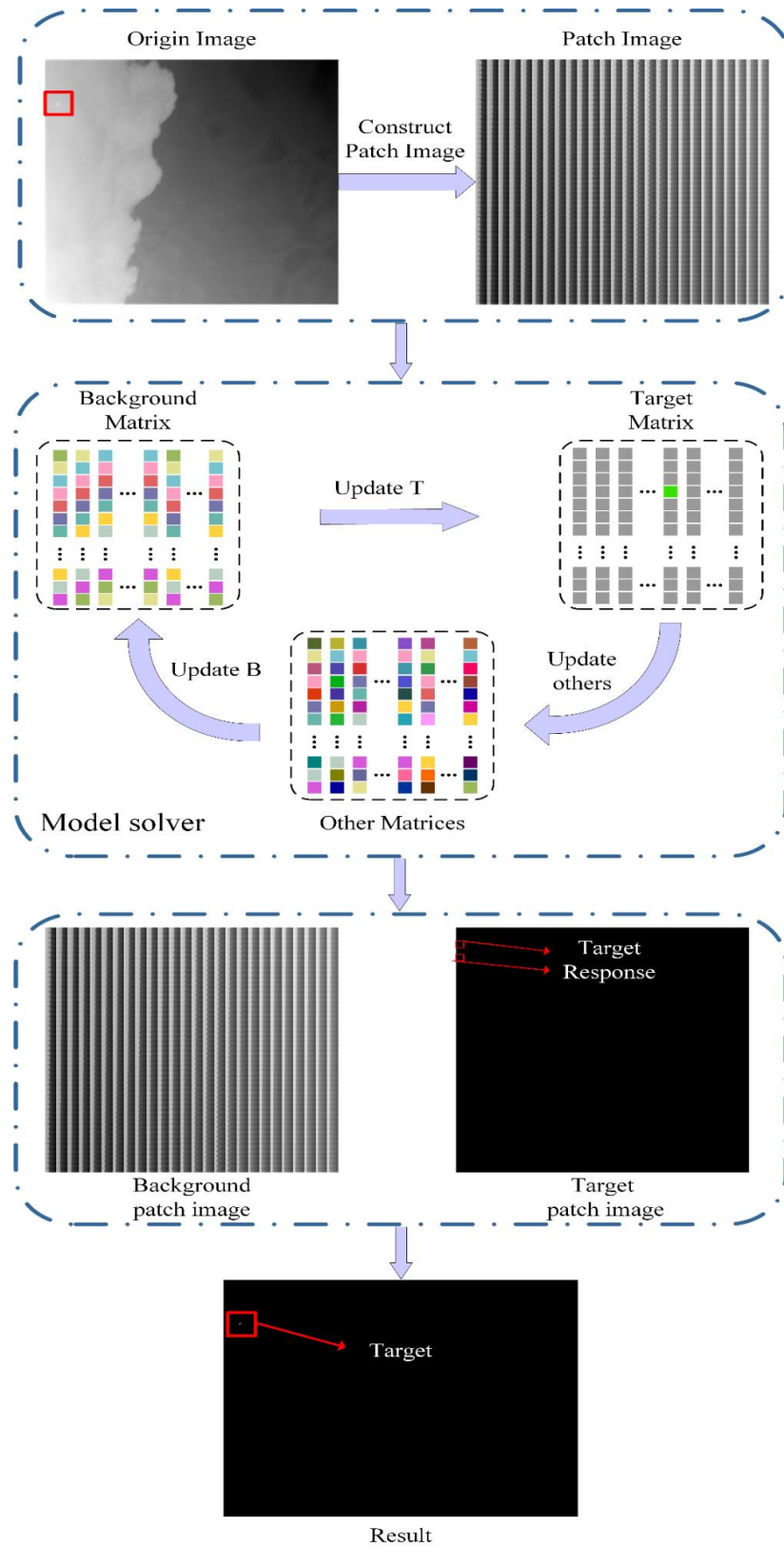
$$T = \mu + k * \sigma \tag{23}$$

Figure 2. The detection pipeline of the proposed model.

4. Experiments and Results

The specific experimental content will be discussed in this part. At first, the compared algorithms with their parameter settings and experimental datasets will be presented. Then, several quantitative indicators are introduced. Subsequently, the best parameters of the proposed model are selected through experiments. Finally, the datasets are tested by the proposed method and the baselines methods.

4.1. Experimental Setting

The proposed model is supposed to compare with eight SOTA methods, including local contrast measure (LCM), multiscale patch-based contrast measure (MPCM), absolute directional mean difference (ADMD) [44], infrared patch-image model (IPI), nonconvex optimization with Lp-norm constraint (NOLC), nonconvex rank approximation minimization joint L2,1 norm (NARM), and partial sum of the tensor nuclear norm (PSTNN) [45] model. The compared methods' parameters choice was set as the authors suggested, which is given in Table 2.

Table 2. Details of seven compared algorithms.

Algorithm	Parameter Setting
LCM	Slide windows size: 3×3
MPCM	Slide windows size: 1, 2, 3, 4
ADMD	Slide windows size: 3, 5, 7, 9
IPI	Patch size: 30×30 Slide *step*: 10 *lambda*: 1/sqrt (min (m, n))
NOLC	Patch size: 30×30 Slide *step*: 10 *lambda*: 1/sqrt (max (m, n)) $p = 0.5$
NARM	Patch size: 30×30 Slide *step*: 10 *lambda*: 1/sqrt (min (m, n))
PSTNN	Patch size: 40×40 Slide *step*: 40 *lambda*: 0.7/sqrt (min (m, n))

The proposed model is conducted on eight different scenes, four real sequences, and four single frame scenes are contained. Each frame image of the sequence datasets and the last two single frames contain only one target, while the first two single frame images contain more than one target. The scenes of the test datasets consist with various complex interference, which mainly involve the background edges with large brightness fluctuations and flickering background clutters, such as sequence 1,3 and single frame 1,4. The details of the test images are listed in Table 3.

Table 3. Detailed descriptions of the test datasets.

Test Image	Frame Number	Size	Image Description
sequence1	50	589×418	The background contains a lot of broken clouds and banded clouds, and the target occupies few pixels
sequence2	50	480×359	The image contains a highlighted building background, and the target is dim
sequence3	50	256×239	The background contains cumulus and edge clutter with high brightness, and there are few target pixels
sequence4	50	256×200	The background mainly contains banded clouds
single frame1		198×134	The background of the image is relatively uniform, but there is some convex interference. The image contains two dim targets, which are close to each other
single frame2		128×128	The image is based on the ground and contains three targets, which are vulnerable to background interference during detection
single frame3		233×161	There are buildings in the image background, and the details of buildings are easy to interfere with the detection of targets occupying a few pixels
single frame4		128×128	The target is buried in the ground object background and seriously disturbed by the weed texture

4.2. Evaluation Metrics

In order to objectively illustrate the effectiveness of the proposed method, quantitative evaluation indicators are introduced. The signal-to-clutter ratio gain (SCRG) and the background suppression factor (BSF) can quantitatively describe the target enhancement and background suppression ability of the algorithms, respectively. In addition, the receiver operating characteristic (ROC) curve shows the relationship between the true positive rate (*TPR*) and false positive rate (*FPR*).

SCRG is defined as:

$$SCRG = \frac{SCR_{out}}{SCR_{in}} \tag{24}$$

where *SCR* is an indicator to measure the significance of objectives. The numerator represents the *SCR* processed by the algorithm, while the denominator represents the *SCR* of the original image

SCR is formulated Equation (25):

$$SCR = \frac{|\mu_t - \mu_b|}{\sigma_b} \tag{25}$$

where μ_t and μ_b indicate the target area's and the neighbor region's average pixel values. σ_b indicates the standard deviation of the neighbor area around the target position. The region size of the target and neighbor is set to 10×10 and 40×40 in the experiment.

BSF is formulated Equation (26):

$$BSF = \frac{\sigma_{in}}{\sigma_{out}} \tag{26}$$

where σ_{in} and σ_{out} indicates the standard deviation of the surrounding background in the input image and the standard deviation of the surrounding background in the output image. According to the formula of *SCRG* and *BSF*, the target enhancement ability and background noise suppression ability of the algorithm can be measured quantitatively. The larger *SCRG* and *BSF*, the better performance of the algorithm.

4.3. ROC Curve

ROC curve is another extensively employed evaluation in the single frame detection field, which can utilize the relationship between the *FPR* and *TPR* to illustrate the detection ability of the algorithm. *TPR* is supposed to demonstrate the detection proportion of the correct detection. *FPR* demonstrates the proportion of false alarm response which is detected as a target. As mentioned above, the algorithm has better performance when its ROC curve is closer to up and left. The *TPR* is formulated Equation (27):

$$TPR = \frac{\text{counts of true detections}}{\text{counts of real targets}} \tag{27}$$

FPR is formulated Equation (28):

$$FPR = \frac{\text{counts of false detections}}{\text{counts of image pixels}} \tag{28}$$

4.4. Parameter Analysis

The sliding window size *len*, window sliding *step*, and *lambda1* are important variables in the proposed model that have a major impact on the low rank, sparse, and iteration. For the purpose of obtaining the best results of the real dataset, it is essential to conduct sufficient experiments on the selection of key variables. Figure 3 is the ROC curve of the comparison with the parameters on the four real IR image sequences.

As depicted in Figure 3, the ROC curves obtained from the experiments from the sequences 1–4 from up to down. The ROC curves from left to right are the comparison among the *len*, *step,* and *lambda1.* We take the value of *len* as 20, 30, 40, 50, and 60, and discuss the impact on the optimization results under the condition of only changing the window size. It is seen that with the increasing of the value of *len*, the detection rate is increased, but when the value is 60, the detection performance begins to degrade. As the window size increases, the content in the patch is more and more abundant, and the sparse characteristic becomes more obvious. However, when the size is more than 60, the patch information is redundant, causing the sparse characteristic to degrade. This means that there is a reasonable value of *len*, which can make an ideal experimental result. Through the ROC curve in the first column in Figure 3, we can see that, for the fixed FPR, the TPR is the highest of the *len* value at 50.
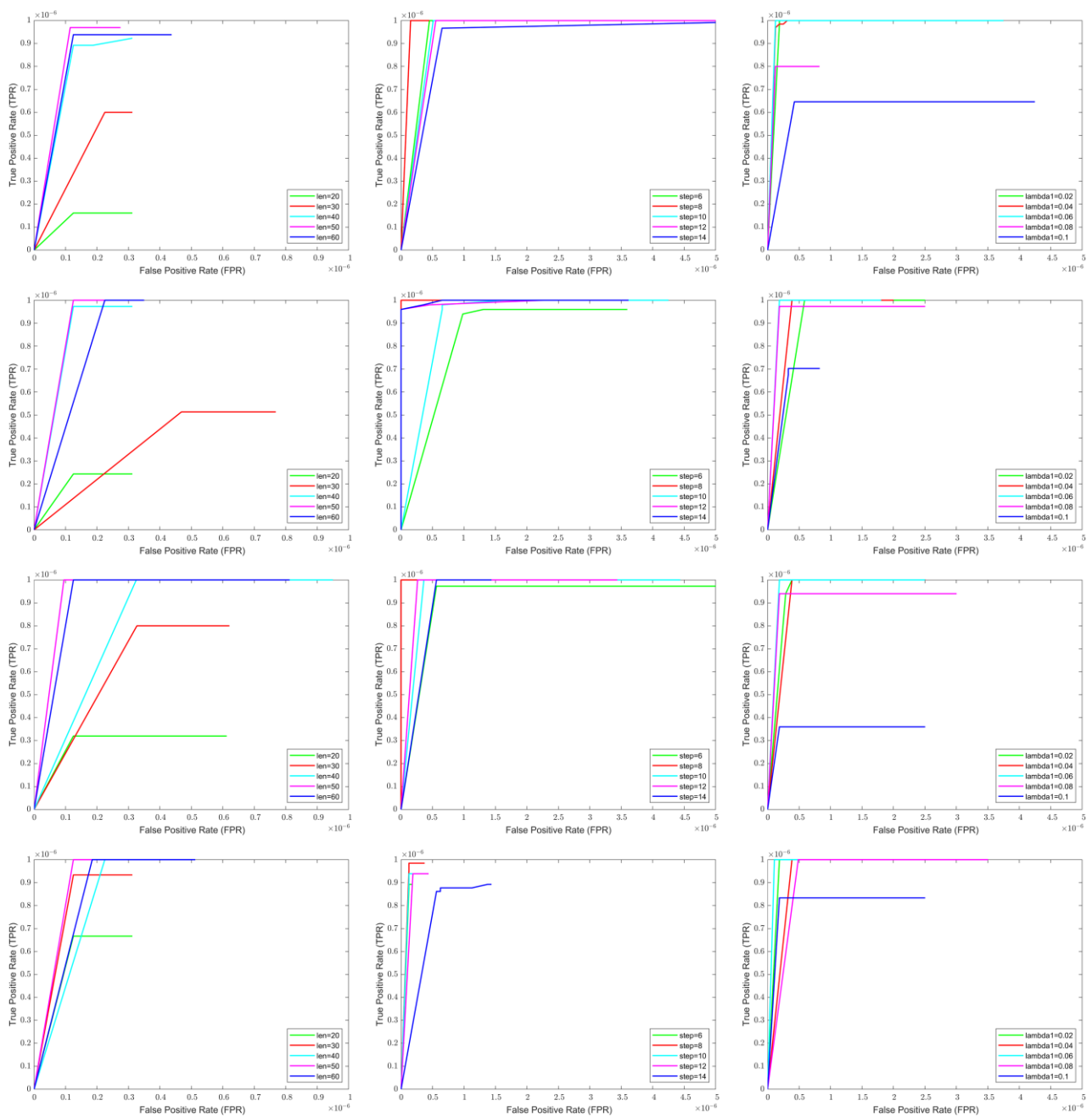
Figure 3. Parameter setting comparison.

For the sliding *step*, the variable decreases and the larger the overlapping area of the patch. Therefore, the value of the sliding *step* mainly affects the rank attribute of the constructed patch image. Furthermore, the smaller *step* will increase the dimension of the patch image and the number of calculations increases. The second column ROC curves in the Figure 3 are the comparison of the parameter of *step*. In addition, the value of the *step* is set 6, 8, 10, 12, and 14. In the experiment, the other parameters are invariant. As seen in the ROC curve, we can conclude that the value of the sliding *step* is 8 and can achieve a better detection performance.

As for *lambda1*, which is a sensitive parameter for the detection ability, the ROC curves of the comparison are shown in the third column. The value of *lambda1* is set to 0.02, 0.04, 0.06, 0.08, and 0.1. The ROC curve indicates that there is an appropriate *lambda1* in a reasonable range to achieve the best performance on the test data. When the value of *lambda1* is 0.06, the model can detect the targets with the lowest FPR.

4.5. Comparison to SOTA

The above section discusses the vital parameters of the model, and, to confirm the robustness of the proposed model, in this section, we trial the proposed method with the SOTA methods on four real IR sequences and four single scenes. The comparison methods are listed in Table 1, including local contrast method (LCM), multiscale patch-based contrast measure (MPCM), absolute directional mean difference algorithm (ADMD), infrared patch image (IPI), nonconvex optimization with Lp-norm constraint (NOLC), nonconvex rank approximation minimization (NRAM), and partial sum of tensor nuclear norm (PSTNN). Figures 4 and 5 show the performance of all algorithms on the sequence images, and Figures 6 and 7 show the performance of all algorithms on the single frame images.

The outcomes of the proposed method and SOTA algorithms are depicted in Figure 4. The 3D displays of the results can intuitively learn about all algorithms' detection abilities. Figure 5 displays the outcomes of comparison methods and the original IR images in gray. We make the target position in the images stand out with red rectangles and expand the target area in the images' corners for better display.

It is obvious that the SOTA methods can eliminate the background and enhance the target to a certain extent. Nevertheless, LCM, MPCM, and ADMD have the detection performance and is relatively worse at the same FPR. Because these methods are proposed on the basis of simple assumptions, their detection ability is very poor, which is the Gaussian-like target, and the target is located at a uniform region. By comparison, the other methods are based on the low rank and sparse recovery assumption, which can obtain better results. Figures 4 and 5 show that the IPI and NARM has many remaining clutters after processing. The clutters mainly obtain the edges and high bright part of the architecture. As for IPI, the simple L1 norm is employed to constrain the sparse item, resulting in a worse result, while the corner and isolate clutters are also sparse in the recovery. For NRAM, providing the surrogate, the sparse item with a weighted L1 norm for more accuracy limit the sparse item. However, it still encounters strong edge inference. The PSTNN-based IPT [46] model introduces a surrogate of the tensor rank item and unfolding of the patch-tensor to capture the low rank property. Similar to NRAM, it focuses on the low rank constraint and neglects the limit on the sparse regulation. Thus, the performances of both NRAM and PSTNN are not well suited to resist the salient clutters. The NOLC method presents the Lp norm to surrogate the L1 norm as an effective sparse constraint item and can achieve relatively better detection results compared with the methods analyzed above. However, as shown in Figure 4, there are few residual clutters in the results.

For further analysis of the abilities of all algorithms, the experiments are conducted on four single frame IR images. Figures 6 and 7 show the 3D display and gray display of the results of the single frame images, respectively. In Figure 7, we made the target position in the images stand out with red rectangles and expanded the target area in the images' corners for better display in the single target scenarios. As described in Table 2, the first and second frames contain more than one target, and the targets in the third and fourth frames are tiny and are severely interfered by the background texture. Generally speaking, the performance of optimization-based algorithms is better than the HVS-based methods at the same FPR. Then the NOLC method can achieve better background suppression by benefiting its nonconvex sparse item constraint. However, the proposed method can obtain a superior result on the test datasets in background elimination and target strengthening than NOLC. The defect of NOLC of the Lp norm is poor robustness, for it may cause a missed detection or a little higher FPR.

The ROC comparison for all methods of four real IR image sequences is advanced in Figure 8 to demonstrates the advantages of the advanced method. The horizontal axis and vertical axis are the FPR, defined as Formulation (28) and the TPR defined as Formulation (27), respectively. As seen in Figure 8, the LCM always shows the lowest TPR with the highest FPR among the comparison algorithm due to the simple local contrast feature extraction with the nine cells nest structure. The MPCM performs a little better than LCM, which can suppress the uniform background but is terrible at eliminating the complicated background. This illustrates the improved background suppression ability of MPCM compared with LCM, while LCM mainly enhanced the target with local contrast features. ADMD can achieve effective detection results on all sequences images compared with the HVS-based and background suppression-based methods. Then, the PSTNN, IPI, and NRAM have similar performance. The FPR is controlled at the low level, and the TPR is at a relatively high level. Furthermore, the NRAM keeps the lowest FPR among the three methods. The performance of NOLC is that it is the more high-performance optimization-based algorithm, which is polluted with high bright clutters causing little FPR. In general, the proposed method can obtain a better ROC on all the test sequence images, which is the lowest FPR and highest TPR among the other SOTA methods.
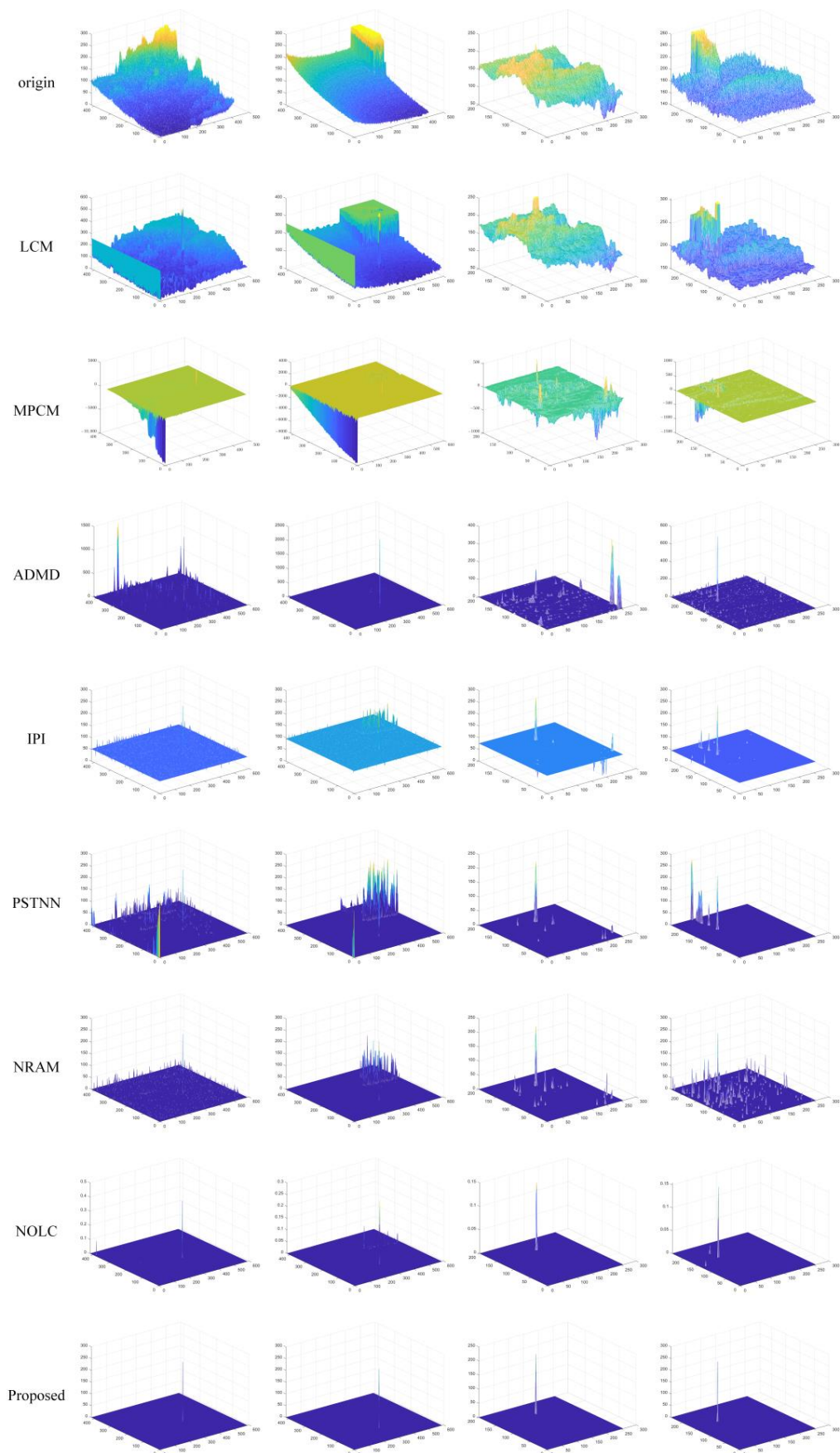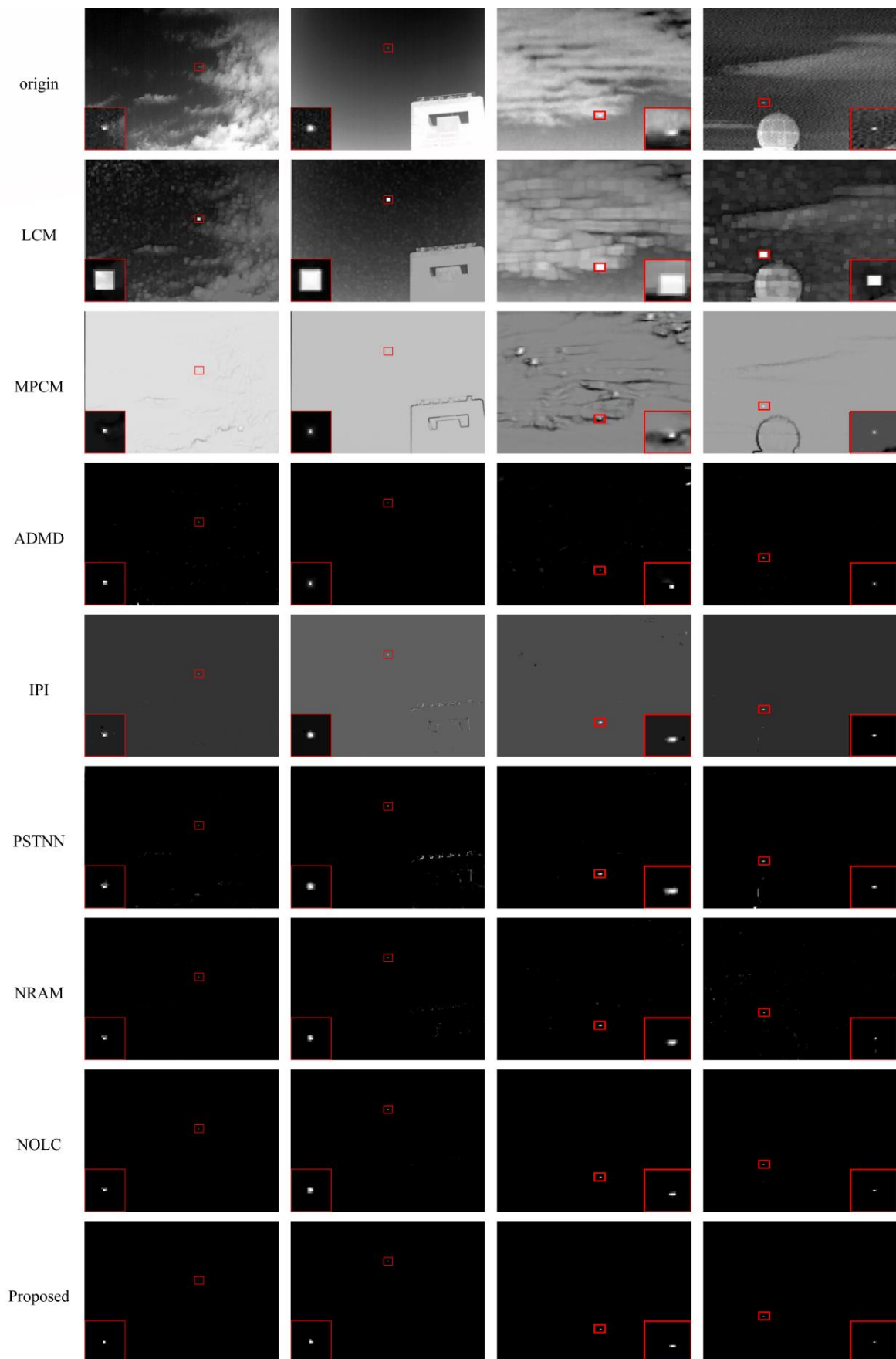
Figure 4. 3D display of sequences, original IR image, and SOTA method processing results.

Figure 5. Gray display of sequences, original IR image, and SOTA method processing results.
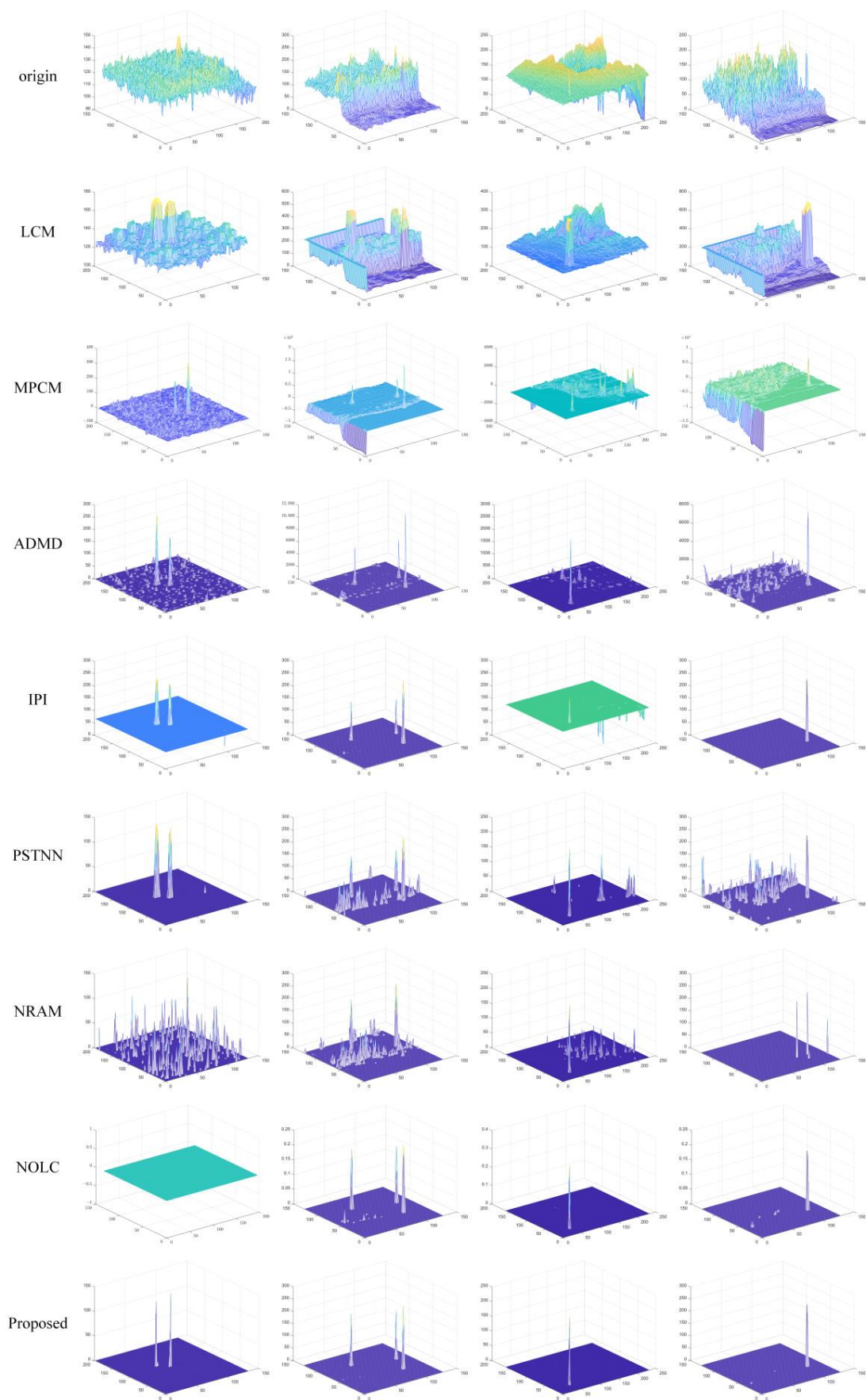
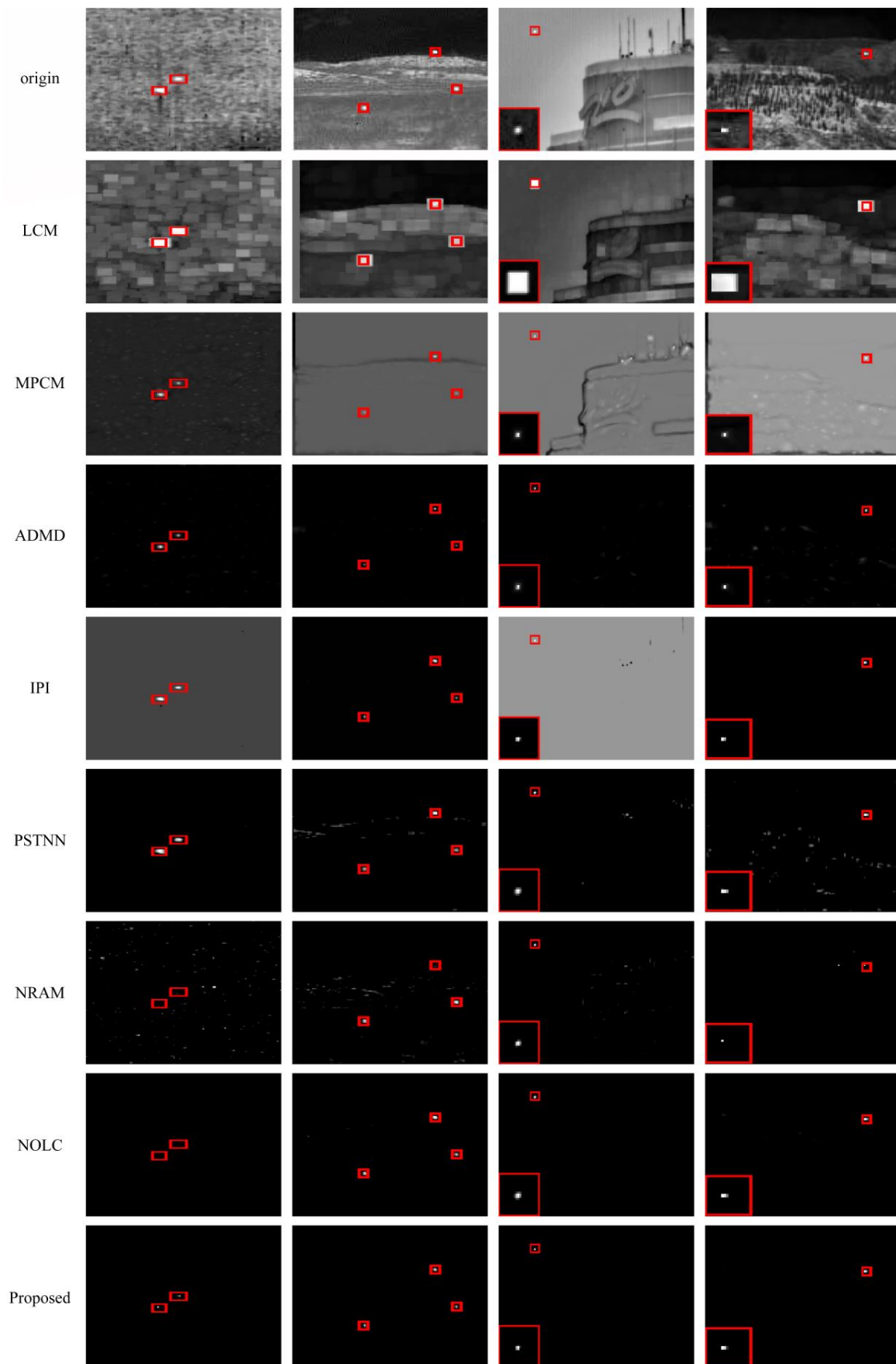**Figure 6.** 3D display of single frame, original IR image, and SOTA method processing results.

Figure 7. Gray display of single frame, original IR image, and SOTA method processing results.
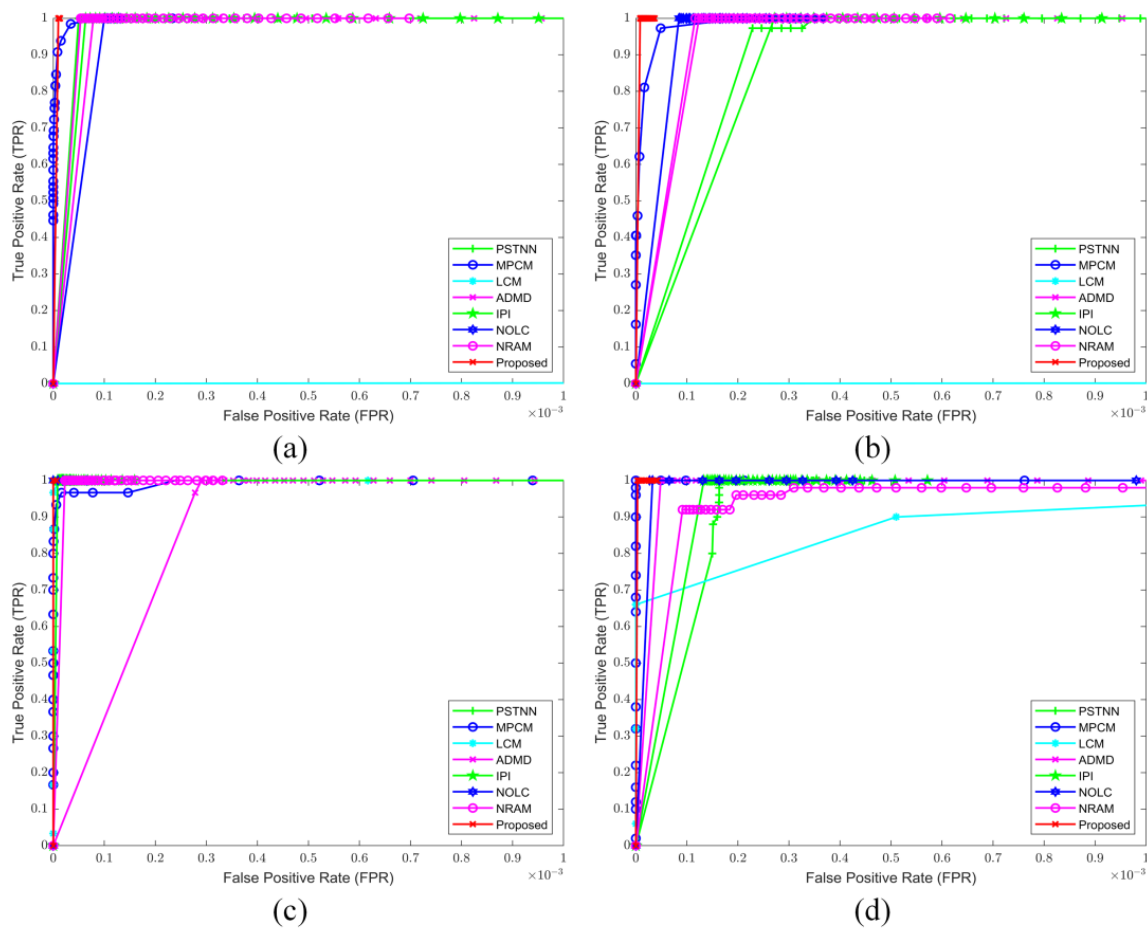
Figure 8. (**a**–**d**) ROC curves of SOTA methods on four real sequences.

For the further quantitative evaluation of the efficiency of the advanced method, we utilized the SCRG and BSF indexes to measure the abilities of target enhancement and background suppression, respectively. The definitions of the indexes are introduced in Formulations (24) and (26). As shown in Table 4, the metrics of the proposed methods outperformed the other methods. Intuitively, the bold display shows the best two indicators. In addition, it is obvious that the optimization-based approaches show better abilities than the other methods.

Table 4. SCRG and BSF values of the SOTA methods.

		LCM	MPCM	ADMD	IPI	PSTNN	NRAM	NOLC	Proposed
Seq 1	SCRG	0.6766	3.3325	2.8766	3.0071	**3.7805**	1.1525	2.6787	**Inf**
	BSF	1.4943	4.3979	**30.5796**	4.6997	11.8668	13.5845	24.4370	**37.8915**
Seq 2	SCRG	2.8994	2.9210	42.4387	46.9946	28.8361	5.2754	**Inf**	**Inf**
	BSF	2.0465	4.3853	4.8713	4.7876	12.1526	16.0935	**17.9242**	**22.7435**
Seq 3	SCRG	3.8201	0.6461	73.6852	109.5999	140.9454	14.9330	**Inf**	**Inf**
	BSF	1.5474	1.4455	0.8282	1.7376	1.7940	**2.6039**	2.4205	**2.7145**
Seq 4	SCRG	3.5833	4.6747	39.1615	197.6377	14.6742	16.7123	**Inf**	**Inf**
	BSF	1.1544	2.8494	11.1216	3.7819	6.2014	10.3059	**11.2824**	**12.9538**

To demonstrate the real-time of the proposed method, we will analyze the computational complexity of the eight algorithms. Assuming that the size of the patch image is m×n and the original IR image size is M × N, the LCM adopts 3 × 3 template for computing so that the complexity is O(MN). Taking the multiscale into consideration and setting k sliding windows of different sizes, the final complexity is O(k^{3}MN) where k is the number of the scales. The MPCM is the same as LCM, which calculates pixel by pixel, so the computation is also O(k^{3}MN). For ADMD, the computation

cost is mainly caused by the nested sliding window. The size of the inner window is 3×3, and the out window is multi-scale and is set to k. Thus, the final computation is $O(8k^3MN)$. For the IPI model, the time-consuming part is from the matrix SVD with a complexity of $O(mn^2)$. In addition, the NARM and NOLC have the same computational complexity with the IPI. The PSTNN needs to construct a tensor with the size $n_1 \times n_2 \times n_3$, and the calculation complexity is depended on the tensor SVD and FFT operation. The FFT only conducts on the frontal slice, which the size is $n_1 \times n_2$. Finally, the complexity of PSTNN is $O(n_1n_2n_3 \log(n_1n_2) + n_1n_2^2[(n_3 + 1)/2])$. The computational complexity of the proposed method relies on the matrix SVD and Newton-CG, and the complexity is $O(nm^2)$. It is obvious that the computational complexity of the proposed method is closely related to the size of the patch image. Furthermore, the methods based on low-rank recovery have a computational cost of $O(nm^2)$ on average. Therefore, the proposed algorithm can achieve a relatively acceptable computational complexity. The computational complexity of all methods in the experiments is summarized in Table 5.

Table 5. Comparison of computational complexity of all methods.

Method	LCM	MPCM	ADMD	IPI	PSTNN	NRAM	NOLC	Proposed
Complexity	$O(k^3MN)$	$O(k^3MN)$	$O(8k^3MN)$	$O(nm^2)$	$O(n_1n_2n_3 \log(n_1n_2) + n_1n_2^2[(n_3 + 1)/2])$	$O(nm^2)$	$O(nm^2)$	$O(nm^2)$

5. Discussion

The background and target separation with optimization-based methods are extensively employed in the aspect of infrared small target detection. More and more researchers improve the robustness of the model with the surrogate of the nuclear norm for accurately estimating the rank of the background image or adding the additional regularization for constraining the remaining noise. The NOLC analysis is the influence of the difference between rank and sparse constraint for the IR target detection. However, the complex background edges and sparse clutters seriously interfere with the dim small target detection. In order to improve the IR small target detection model robustness under the RPCA framework, we utilized the L1–L2 norm as the sparse regularization and the total variation item to work on the target image. In Section 4, the experiments on the IR image sequences and single frames revealed that the sparse measure with L1–L2 norm could achieve an approving result in the detection tasks, though the targets were disturbed with the complex background edges and sparse clutters.

Comparing with other optimized based-methods, IPI is the beginning of the optimization-based model, which only employs the nuclear norm to estimate the rank of the background and the L1 norm as the sparse item. Therefore, the simple regularizations are inadequate to handle the complicated scenes. Making full use of the nonlocal prior information, the PSTNN extends the matrix structure to tensor structure. It mainly works at the background recovery accurately with the improved tensor nuclear norm. However, the component of the background is not strictly low rank, causing unsatisfactory recovery results facing the complex background. As shown in the experiment outcomes in Section 3, PSTNN is sensitive to the strong edges, and the tensor model ability will be limited in these scenes. As an extension of IPI, NRAM introduced the structural noise suppression regularization, which employed the L21 norm to achieve the row sparsity. However, the patch size plays an important role in structural noise elimination. If the sliding window size does not match the noise region, the effect of the regularization may not work well. Then, the NOLC improves the robustness from the perspective of the target which is different from the above mentioned one. The sparse item with the Lp norm can obtain sparser results than the L1 norm. However, the Lp norm causes missed detection and false alarms due to the selection of the value of *p*. The proposed method improves the robustness of the sparse item with the L1–L2 norm and adds the total variation norm to decrease the interference from strong clutters on the target perspective.

The proposed detection method utilizes the parameter free sparse item to surrogate the L0 norm, and we constrain the clutters in the target image with the total variation. Thus, the model can succeed against the challenging clutters and achieve a great detection effect.

The performance of the proposed model is illustrated by experiments. At first, we analyzed the vital parameters in the model and confirmed the optimal parameters. Then experiments were conducted on four sequences and four single frame images. Compared with the SOTA methods, the outcomes of the advanced model are the best among all the algorithms and consistent with the analysis results. The advanced model works well in both target detection accuracy and background

elimination. As the results were analyzed in the experiment, the computational complexity of our algorithm did not improve. If the image size of the experiment is larger, the algorithm will take longer to run compared to the filter-based method.

6. Conclusions

In summary, the proposed approach of incorporating the L1–L2 norm and total variation regularization on the target image has significantly improved the detection accuracy of the IR small target detection model. The L1–L2 norm enhances the model's sparsity constraint, while the TV norm strengthens its ability to reject sparse clutter interference. The resulting nonconvex model demonstrates strong performance in detecting targets and eliminating clutter, as evidenced by the ROC curve with high TPR and low FPR compared to other SOTA methods. Additionally, quantitative metrics like SCRG and BSF further validate the effectiveness of the proposed technique. Overall, this study offers a promising approach for enhancing small target detection in infrared imaging applications.

Although it is true that the proposed method has an average level of computational complexity compared to other low-rank recovery methods, it may still be more time-consuming compared to traditional filtering methods. Therefore, it is important to take into consideration the trade-off between computational efficiency and accuracy when choosing a suitable approach for the task. Additionally, further research should be conducted to improve and optimize the proposed method in order to overcome its potential limitations.

Author Contributions: Methodology, Y.S.; Validation, M.M.; Writing—original draft, Y.S.; Writing— review & editing, X.K.; Supervision, S.H.; Project administration, D.W.; Funding acquisition, C.C. and S.H. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

References

1. Yavari, M.; Moallem, P.; Kazemi, M.; Moradi, S. Small Infrared Target Detection Using Minimum Variation Direction Interpolation. *Digit. Signal Process.* **2021**, *117*, 103174. [CrossRef]
2. 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]
3. Cai, Y. Weighted Lp—L1 Minimization Methods for Block Sparse Recovery and Rank Minimization. *Anal. Appl.* **2020**, *19*, 343–361. [CrossRef]
4. Peng, C.; Liu, Y.; Kang, K.; Chen, Y.; Wu, X.; Cheng, A.; Kang, Z.; Chen, C.; Cheng, Q. Hyperspectral Image Denoising Using Nonconvex Local Low-Rank and Sparse Separation with Spatial–Spectral Total Variation Regularization. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–17. [CrossRef]
5. Aragón Artacho, F.J.; Vuong, P.T. The Boosted Difference of Convex Functions Algorithm for Nonsmooth Functions. *SIAM J. Optim.* **2020**, *30*, 980–1006. [CrossRef]
6. Huang, J.; Almurib, H.A.F.; Kumar, T.N.; Lombardi, F. An Inexact Newton Method For Unconstrained Total Variation-Based Image Denoising by Approximate Addition. *IEEE Trans. Emerg. Top. Comput.* **2022**, *10*, 1192–1207. [CrossRef]
7. Balasingam, B.; Bar-Shalom, Y.; Willett, P.; Pattipati, K. Maximum Likelihood Detection on Images. In Proceedings of the 2017 20th International Conference on Information Fusion (Fusion), Xi'an, China, 10–13 July 2017; pp. 1–8.
8. Deshpande, S.D.; Er, M.H.; Venkateswarlu, R.; Chan, P. Max-Mean and Max-Median Filters for Detection of Small Targets. In *Signal and Data Processing of Small Targets*; SPIE: Santa Clara, CA, USA, 1999; Volume 3809, pp. 74–83.
9. Liu, R.; Wang, D.; Zhou, D.; Jia, P. Point Target Detection Based on Multiscale Morphological Filtering and an Energy Concentration Criterion. *Appl. Opt.* **2017**, *56*, 6796. [CrossRef]
10. Shao, X.; Fan, H.; Lu, G.; Xu, J. An Improved Infrared Dim and Small Target Detection Algorithm Based on the Contrast Mechanism of Human Visual System. *Infrared Phys. Technol.* **2012**, *55*, 403–408. [CrossRef]
11. Aghaziyarati, S.; Moradi, S.; Talebi, H. Small Infrared Target Detection Using Absolute Average Difference Weighted by Cumulative Directional Derivatives. *Infrared Phys. Technol.* **2019**, *101*, 78–87. [CrossRef]
12. Bai, X.; Bi, Y. Derivative Entropy-Based Contrast Measure for Infrared Small-Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 2452–2466. [CrossRef]
13. Chen, C.L.P.; Li, H.; Wei, Y.; Xia, T.; Tang, Y.Y. A Local Contrast Method for Small Infrared Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 574–581. [CrossRef]

14. 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. [CrossRef]

15. 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]

16. Wei, Y.; You, X.; Li, H. Multiscale Patch-Based Contrast Measure for Small Infrared Target Detection. *Pattern Recognit.* **2016**, *58*, 216–226. [CrossRef]

17. 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.* **2021**, *18*, 1670–1674. [CrossRef]

18. Shi, Y.; Wei, Y.; Yao, H.; Pan, D.; Xiao, G. High-Boost-Based Multiscale Local Contrast Measure for Infrared Small Target Detection. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 33–37. [CrossRef]

19. 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.* **2020**, *17*, 327–331. [CrossRef]

20. Yao, S.; Chang, Y.; Qin, X. A Coarse-to-Fine Method for Infrared Small Target Detection. *IEEE Geosci. Remote Sens. Lett.* **2019**, *16*, 256–260. [CrossRef]

21. 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]

22. 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]

23. 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]

24. 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]

25. Zhang, Z.; Ely, G.; Aeron, S.; Hao, N.; Kilmer, M. Novel Methods for Multilinear Data Completion and De-Noising Based on Tensor-SVD. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 3842–3849.

26. Zhang, L.; Peng, L.; Zhang, T.; Cao, S.; Peng, Z. Infrared Small Target Detection via Non-Convex Rank Approximation Minimization Joint L2,1 Norm. *Remote Sens.* **2018**, *10*, 1821. [CrossRef]

27. Zhang, T.; Wu, H.; Liu, Y.; Peng, L.; Yang, C.; Peng, Z. Infrared Small Target Detection Based on Non-Convex Optimization with Lp-Norm Constraint. *Remote Sens.* **2019**, *11*, 559. [CrossRef]

28. Chartrand, R.; Staneva, V. Restricted Isometry Properties and Nonconvex Compressive Sensing. *Inverse Probl.* **2008**, *24*, 035020. [CrossRef]

29. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef]

30. Fan, Z.; Bi, D.; Xiong, L.; Ma, S.; He, L.; Ding, W. Dim Infrared Image Enhancement Based on Convolutional Neural Network. *Neurocomputing* **2018**, *272*, 396–404. [CrossRef]

31. Ju, M.; Luo, J.; Liu, G.; Luo, H. ISTDet: An Efficient End-to-End Neural Network for Infrared Small Target Detection. *Infrared Phys. Technol.* **2021**, *114*, 103659. [CrossRef]

32. Dai, Y.; Wu, Y.; Zhou, F.; Barnard, K. Attentional Local Contrast Networks for Infrared Small Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 9813–9824. [CrossRef]

33. Hou, Q.; Zhang, L.; Tan, F.; Xi, Y.; Zheng, H.; Li, N. ISTDU-Net: Infrared Small-Target Detection U-Net. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]

34. Zhou, F.; Wu, Y.; Dai, Y.; Ni, K. Robust Infrared Small Target Detection via Jointly Sparse Constraint of L1/2-Metric and Dual-Graph Regularization. *Remote Sens.* **2020**, *12*, 1963. [CrossRef]

35. Yin, P.; Lou, Y.; He, Q.; Xin, J. Minimization of ℓ_{1-2} for Compressed Sensing. *SIAM J. Sci. Comput.* **2015**, *37*, A536–A563. [CrossRef]

36. Lou, Y.; Yin, P.; He, Q.; Xin, J. Computing Sparse Representation in a Highly Coherent Dictionary Based on Difference of L_1 L 1 and L_2 L 2. *J. Sci. Comput.* **2015**, *64*, 178–196. [CrossRef]

37. Sun, L.; Ge, W.; Chen, Y.; Zhang, J.; Jeon, B. Hyperspectral Unmixing Employing l_1-l_2 Sparsity and Total Variation Regularization. *Int. J. Remote Sens.* **2018**, *39*, 6037–6060. [CrossRef]

38. Kim, S.; Yang, Y.; Lee, J.; Park, Y. Small Target Detection Utilizing Robust Methods of the Human Visual System for IRST. *J. Infrared Millim. Terahertz Waves* **2009**, *30*, 994–1011. [CrossRef]

39. Zhao, M.; Li, W.; Li, L.; Hu, J.; Ma, P.; Tao, R. Single-Frame Infrared Small-Target Detection: A Survey. *IEEE Geosci. Remote Sens. Mag.* **2022**, *10*, 87–119. [CrossRef]

40. Zhang, J.; Zhang, B.; Liu, P. Infrared Small Target Detection Based on Salient Region Extraction and Gradient Vector Processing. In Proceedings of the ACM International Conference Proceeding Series; Association for Computing Machinery: New York, NY, USA, 2019; pp. 422–426.

41. Cai, J.-F.; Candès, E.J.; Shen, Z. A Singular Value Thresholding Algorithm for Matrix Completion. *SIAM J. Optim.* **2010**, *20*, 1956–1982. [CrossRef]

42. Zhang, F.; Yang, Z.; Wan, M.; Yang, G. Robust Principal Component Analysis Based on L1-2 Metric. In Proceedings of the 2017 4th IAPR Asian Conference on Pattern Recognition (ACPR), Nanjing, China, 26–29 November 2017; pp. 394–398.

43. Yuan, M.; Lin, Y. Model Selection and Estimation in Regression with Grouped Variables. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2006**, *68*, 49–67. [CrossRef]

44. Moradi, S.; Moallem, P.; Sabahi, M.F. Fast and Robust Small Infrared Target Detection Using Absolute Directional Mean Difference Algorithm. *Signal Process.* **2020**, *177*, 107727. [CrossRef]

45. Zhang, L.; Peng, Z. Infrared Small Target Detection Based on Partial Sum of the Tensor Nuclear Norm. *Remote Sens.* **2019**, *11*, 382. [CrossRef]

46. Zhang, X.; Ding, Q.; Luo, H.; Hui, B.; Chang, Z.; Zhang, J. Infrared Small Target Detection Based on an Image-Patch Tensor Model. *Infrared Phys. Technol.* **2019**, *99*, 55–63. [CrossRef]