

Infrared Small Target Detection Based on a Group Image-Patch Tensor Model

Lanlan Yang^{ID}, Peng Yan^{ID}, Meihui Li^{ID}, Jianlin Zhang^{ID}, and Zhiyong Xu

Abstract—Despite many years of research, the accuracy of small target detection is still restricted by the complex background. In addition, the contradiction between the computation complexity reduction and detection performance improvement makes it hard to apply some high-performance methods toward practical application. In this letter, a novel group image-patch tensor (GIPT) model is proposed to solve the above-mentioned problems. First, based on the structure tensor theory, a newly designed measurement criterion is employed to divide the image pixels into three types: point area, flat area, and line area. Second, image pixels with line priors and point priors are assigned to different weights to reduce the effects of strong background edges. Third, a GIPT model is proposed to better explore the low-rank of the background, in which patches with the same feature type are rearranged into the same group for tensor decomposition. Moreover, the proposed GIPT model can make the optimization process run in parallel, which can greatly reduce the complexity of the algorithms. The experiment results show the superiority of the proposed method compared with other state-of-the-art algorithms.

Index Terms—Alternating direction method of multipliers (ADMM), group image-patch tensor (GIPT) model, infrared small target detection, structure tensor.

I. INTRODUCTION

INFRARED small target detection plays an important role in many practical applications such as early-warning, remote sensing, and infrared search and track (IRST). With the development of compressed sensing and sparse representation theories, low-rank sparse decomposition is widely used. But there are some difficulties.

- 1) The characteristics of the strong edges and the small targets are very similar, which can easily lead to target misjudgment.
- 2) The small targets get buried due to the low signal-clutter ratio (SCR). Therefore, many researchers have many proposals in recent years.

According to the form of data to be processed, the existing low-rank sparse decomposition algorithms can be divided into matrix-based and tensor-based methods. The matrix-based methods include infrared patch-image (IPI) model [1], nonnegative IPI model based on partial sum minimization of

Manuscript received November 23, 2021; revised December 21, 2021; accepted December 29, 2021. Date of publication January 4, 2022; date of current version February 4, 2022. This work was supported by the National Natural Science Foundation of China under Grant 62101529. (*Corresponding authors:* Peng Yan; Meihui Li.)

Lanlan Yang is with the Key Laboratory of Optical Engineering, Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610200, China, and also with the School of Optoelectronics, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: 1663462583@qq.com).

Peng Yan, Meihui Li, Jianlin Zhang, and Zhiyong Xu are with the Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610200, China (e-mail: 65493754@qq.com; limeihui@ioe.ac.cn; jlin_zh@163.com; xzy158@163.com).

Digital Object Identifier 10.1109/LGRS.2021.3140067

singular values (NIPPS) [2], nonconvex rank approximation minimization (NRAM) [3], and so on. These methods employ a fix-size window to slide on an infrared image from the left-top corner to the bottom-right corner. Then the obtained local image patches are vectorized to form a data matrix that is further used for matrix decomposition. By comparison, the tensor-based methods with more spatial information usually stack the local image-patches into a 3-D tensor and apply the Tucker algorithm for tensor decomposition. The tensor-based methods include the reweighted infrared patch tensor (RIPT) model [4] and the partial sum of the tensor nuclear norm (PSTNN) based on the infrared patch-tensor (IPT) model [5] and so on.

Generally, both the rank and the l_0 norm are nonconvex, so the current methods are adopted to try to solve the convex surrogate. For the low rank of background, the sum of the nuclear norm (SNN), tensor nuclear norm (TNN), tensor multimode expansion sum nuclear norm (TMESNN), and other tensor ranks are used. For the sparsity of target, some research studies have been conducted by adding local prior using the local structural tensor that allows simultaneous direction estimation and image structure analysis. Recently Lu *et al.* [6] proposed the tensor average rank and provided the theoretical guarantee for the exact recovery. Cao *et al.* [7] used the mode- k_1k_2 extension tensor tubal rank to explore the hidden information among different modes of the tensor. Zhao *et al.* [8] applied the morphological profiles to create three-order tensors to acquire more spatial and structural information. Kong *et al.* [9] used the tensor fibered nuclear norm based on the Log operator (LogTFNN) and added the prior information that is extracted by the local structure tensor to their proposed model. Zhou *et al.* [10] defined a novel spatial structure weight function to construct a spatial structural prior based on the nonlocal structure tensor.

However, there are two defects in current algorithms. On the one hand, matrix-based methods are time-consuming while tensor-based methods reduce the complexity of the algorithm, but it is not fast enough. On the other hand, the priors introduced by the structure tensor are used in the improved method to suppress the strong edges, but it cannot effectively suppress the discontinuous strong edges, which leads to the final missing or false detection. It motivated us to solve the detection model parallel with tensor grouping to reduce the complexity and improve the priors to further suppress strong edges. In addition, the group image-patch tensor (GIPT) model used low-rank within groups based on grouping tensor to better explore the background low-rank. Then, we proposed a novel GIPT model that includes the following three innovations:

- 1) Employing a new criterion to measure the pixel-wise features based on the structure tensor theory, in which the line feature and point feature are extracted and performed as a local prior for the target component to achieve background interference suppression.

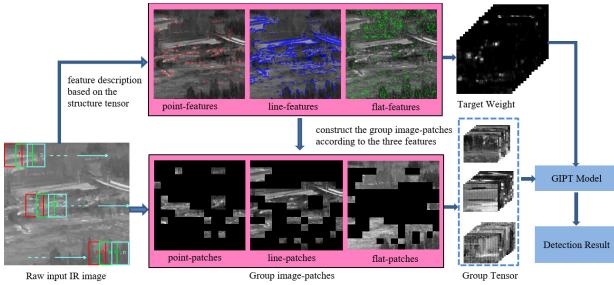


Fig. 1. Flowchart of the proposed algorithm.

Algorithm 1 ADMM Solver to the Proposed Method

Input: \mathbf{D} : Original image, \mathcal{D} : Original image-patch tensor
Output: $\mathbf{B}^k, \mathbf{T}^k$

1: Initialization: $\mathcal{T}^0 = \mathcal{B}^0 = 0, \mathcal{W}_{sw} = 1, c_1 = 0.5, c_3 = 100, c_2 = c_4 = 10, k = 0, \mu^0 = \frac{125}{\|\mathcal{D}\|_2}, \beta = 25 * \mu^0, \rho = 1.05, \eta = 2, \mu^{\max} = 10^{10}, \varepsilon = 10^{-7}, \lambda = \lambda_l / \sqrt{\max(n_1, n_2) * n_3}$

2: The original image is divided into three features by Eq.(7)

3: Complete tensor grouping based on the patch features calculated by Eq.(8)

4: The weight of target \mathcal{W} is constructed by Eq.(11)

5: While not converge do

- Fix the others and update \mathcal{B}^{k+1} by Eq.(16)
- Fix the others and update \mathcal{T}^{k+1} by Eq.(17)
- Fix the others and update \mathcal{W}^{k+1} through update \mathcal{W}_{sw}^{k+1}
- Fix the others and update $\mathcal{Y}^{k+1}, \mu^{k+1}$ by Eq.(18), Eq. (19)
- Check the convergence conditions:

$$\frac{\|\mathcal{B}^{k+1} + \mathcal{T}^{k+1} - \mathcal{D}\|_F^2}{\|\mathcal{D}\|_F^2} < \varepsilon, \text{ or } \|\mathcal{T}^{k+1}\|_0 = \|\mathcal{T}^k\|_0$$

update k: $k = k + 1$

6: End while

7: \mathcal{B}^k and \mathcal{T}^k are reconstructed to images

- 2) Introducing a patch-wise feature identification method. The background is assumed to be group low-rank according to the patch-wise features.
- 3) The weighted sparse target term and the group low-rank background term are applied to form the GIPT model. A parallel optimization method based on alternating direction method of multipliers (ADMM) is proposed to solve this model.

II. TRADITIONAL IMAGE PATCH-TENSOR MODEL

In this letter, we denote tensors as \mathcal{X} , matrices as \mathbf{X} , vectors as \mathbf{x} , and scalars as x . For a three-way tensor, its (i, j, k) th entry is denoted as \mathcal{X}_{ijk} . Similar to the matrix norm, the l_0 -norm of the tensor can be calculated as: $\|\mathcal{X}\|_0 = \#\{(i|\mathcal{X}_{ijk} \neq 0)\}$, the l_1 -norm and Frobenius norm of the tensor can be calculated as: $\|\mathcal{X}\|_1 = \sum_{ijk} |\mathcal{X}_{ijk}|$, $\|\mathcal{X}\|_F = (\sum_{ijk} |\mathcal{X}_{ijk}|^2)^{1/2}$, and $\|\mathcal{X}\|_*$ denotes the TNN.

Low-rank sparse decomposition theory of tensors formulates the target-background separation task as an optimization problem of recovering low-rank and sparse tensors. However, minimizing the rank and l_0 norm is NP-hard, and its convex surrogate can be written as

$$\min_{\mathbf{B}, \mathbf{T}} \|\mathbf{B}\|_* + \lambda \|\mathbf{T}\|_1 \quad \text{s.t. } \mathcal{D} = \mathbf{B} + \mathbf{T} \quad (1)$$

where \mathcal{D} , \mathbf{B} , and \mathbf{T} denote the original image patch-tensor, background patch-tensor, and target patch-tensor, respectively. λ is a constant. Equation (1) is called the IPT model. Generally, minimizing the TNN may cause some unavoidable biases and there are background residuals with strong edges in the target image.

III. PROPOSED METHODS

The flowchart of the proposed algorithm is shown in Fig. 1. First, a feature descriptor is proposed to distinguish the three features of the original image. Then the construction of the tensor grouping and the target weight is completed based on these three features. Finally, target detection is achieved using the proposed GIPT model.

A. Feature Description Based on Structure Tensor

Compared with gradient, the structure tensor obtains more abundant local structural features by the smoothing process and measures the homogeneity of orientations within the neighborhood of a pixel. Supposing that the original image is \mathbf{D} , the classic linear structure tensor is calculated as follows:

$$\mathbf{J} = \mathbf{K}_\rho * (\nabla \mathbf{D}_\sigma \otimes \nabla \mathbf{D}_\sigma) = \begin{pmatrix} \mathbf{K}_\rho * \mathbf{I}_x^2 & \mathbf{K}_\rho * \mathbf{I}_x \mathbf{I}_y \\ \mathbf{K}_\rho * \mathbf{I}_x \mathbf{I}_y & \mathbf{K}_\rho * \mathbf{I}_y^2 \end{pmatrix} = \begin{pmatrix} \mathbf{J}_{11} & \mathbf{J}_{12} \\ \mathbf{J}_{21} & \mathbf{J}_{22} \end{pmatrix} \quad (2)$$

where \mathbf{K}_ρ is a Gaussian kernel function with variance ρ , \mathbf{D}_σ represents Gaussian smoothing filter with variance σ for the original image, $\mathbf{I}_x = (\partial \mathbf{D}_\sigma / \partial x)$, $\mathbf{I}_y = (\partial \mathbf{D}_\sigma / \partial y)$, represents the gradient of \mathbf{D}_σ along the x - and y -directions, respectively. $*$ is convolution operation, ∇ is the gradient, \otimes is Kronecker Product, and the structure tensor is set to 2×2 .

The features of the image can be reflected by the two highest eigenvalues of the structure tensor when λ_1 and λ_2 satisfy different situations [5]: 1) $\lambda_1 \geq \lambda_2 \gg 0$: Locating at corner region; 2) $\lambda_1 \gg \lambda_2 \approx 0$: Locating at edge region; and 3) $\lambda_1 \approx \lambda_2 \approx 0$: Locating at the flat region. To distinguish three situations more accurately, we proposed a function as (3), and it is divided into two parts as follows:

$$\Theta = \frac{1}{2} e^{\frac{\lambda_1}{\lambda_2 + 0.001} - 1} [e^{\lambda_1(\lambda_1 - \lambda_2)} + e^{\lambda_2(\lambda_1 - \lambda_2)}] \quad (3)$$

$$\Theta_1 = e^{\frac{\lambda_1}{\lambda_2 + 0.001} - 1}, \quad \Theta_2 = \frac{1}{2} [e^{\lambda_1(\lambda_1 - \lambda_2)} + e^{\lambda_2(\lambda_1 - \lambda_2)}] \quad (4)$$

where Θ_1 is used as an amplification factor measuring the gap between λ_1 and λ_2 . When Θ tends to infinity, it means that the pixel is at the edge region and it is classified to line-feature. When Θ is near 1, it can be considered as flat-feature, and the rest of the original area is considered as point-feature. However, due to the interference of noise and excessive amplification of Θ_1 , it is necessary to select appropriately threshold values as the judgment is as follows:

$$\text{seg}_p = c_1 \max(Z), \quad \text{seg}_f = c_2 \min(Z) \quad (5)$$

$$\text{seg}_{p1} = c_3 \min(S), \quad \text{seg}_{f1} = c_4 \min(S) \quad (6)$$

where c_1, c_2, c_3, c_4 are all constants. Z and S represent the remaining values after deleting infinity in Θ and Θ_1 , respectively. Through the search to meet the above-discussed conditions

$$\begin{cases} \Theta(i, j) > \text{seg}_p \text{ and } \Theta_1(i, j) \leq \text{seg}_{p1}, & (i, j) \in \text{point} \\ \Theta(i, j) = \infty \text{ and } \Theta_1(i, j) > \text{seg}_{p1}, & (i, j) \in \text{line} \\ \Theta(i, j) \leq \text{seg}_f \text{ and } \Theta_1(i, j) \leq \text{seg}_{f1}, & (i, j) \in \text{flat} \end{cases} \quad (7)$$

where (i, j) represents the pixel position. The image is divided into point-feature, line-feature, and flat-feature when located at the corner region, the edge region, and the flat region.

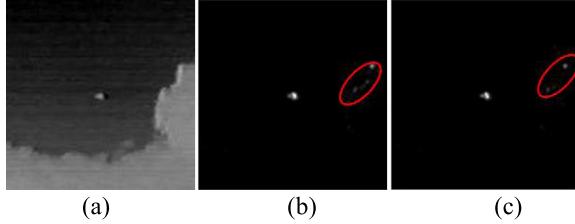


Fig. 2. Comparison of different prior maps. (a) Original image. (b) Prior map used in PSTNN. (c) Prior map used in the proposed model.

B. Background Model

In complex scenes, there are not enough patches with strong edges, which lead to poor background low-rank performance. Therefore, we use the grouping method to make the background satisfy low rank in the similarity space. Naturally, the image patches with large similarities are divided into the same group. The probability that the i th patch belonging to the three features can be calculated by

$$P_i(p, l, f) = \frac{k_i(p, l, f)}{\max(k_p, k_l, k_f)} \quad (8)$$

where $k_i(p, l, f)$ represents the number of pixels belonging to point-feature, line-feature, and flat-feature in the i th image patch. k_p , k_l , and k_f represent the number of pixels belonging to point-feature, line-feature, and flat-feature in all image patches.

To classify the strong edges as group-line definitely, we increased the probability that the image patch satisfying $\Theta_1 = \infty$ and $\Theta_2 = \infty$. Then, the group image-patches consist of point-patches, line-patches, and flat-patches according to the maximum probability.

C. Target Model

Due to the deficiency of RIPT, the target will over-shrink and disappear. Zhang and Peng [5] proposed an improved local prior map simultaneously encoded with target-related and background-related information. Its prior can be expressed as

$$\mathbf{W}_l = \max(\lambda_1, \lambda_2) \cdot \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}. \quad (9)$$

However, when we meet a complex environment, strong edges are magnified. As shown in Fig. 2(b), the area marked by the red ellipse should be suppressed. Subsequently, to enhance the small targets of point-feature, and suppress the strong edges of line-feature, we proposed a new prior

$$\begin{aligned} \mathbf{W}' &= \mathbf{W}_l \odot (\eta * \mathbf{W}_e + \mathbf{W}_p) \\ \mathbf{W}_e(i, j) &\left\{ \begin{array}{ll} 0, & \text{if } \mathbf{D}(i, j) \in \text{line} \\ \tau, & \text{else} \end{array} \right. \\ \mathbf{W}_p(i, j) &\left\{ \begin{array}{ll} \tau, & \text{if } \mathbf{D}(i, j) \in \text{point} \\ 0, & \text{else} \end{array} \right. \end{aligned} \quad (10)$$

where η is a constant, τ is a nonzero constant and is set to 2. \odot denotes the Hadamard product. To speed up the convergence rate, we adopted the reweighted scheme as well. The tensor form of the target weight is

$$\mathbf{W} = \mathbf{W}' \odot \mathbf{W}_{sw}. \quad (11)$$

The sparsity weight is defined: $\mathbf{W}_{sw}^{k+1} = (1/|\mathbf{T}^k| + \varepsilon)^{-1}$, where $\varepsilon > 0$ is a small number to avoid division by zero, and $k+1$ denotes the $(k+1)$ th iteration.

From Fig. 2(c), we can see that although the discontinuous strong edges are not completely removed, they are suppressed to a certain extent in our prior.

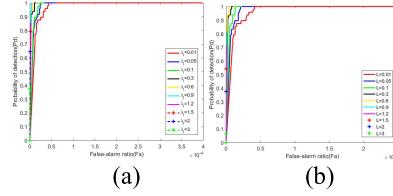


Fig. 3. ROC curves of different parameters. (a) ROC of different λ_l . (b) ROC of different L .

D. GIPT Model

Low rank within groups in which each group is independent with each other and global sparsity of targets while improving the target weight; our proposed GIPT model is

$$\min_{\mathbf{B}, \mathbf{T}} \sum_{i=1}^3 \beta \|\mathcal{B}_i\|_* + \lambda \|\mathcal{T} \odot \mathcal{W}\|_1 \quad \text{s.t. } \mathcal{D} = \mathcal{B} + \mathcal{T} \quad (12)$$

where \mathcal{D} denotes the original image-patch tensor, β and λ are constants, \mathcal{B}_i denotes the background tensor of group i . Our proposed model can better explore the low rank of the background and the sparsity of the target.

E. Optimization Method

The optimization in this letter is based on ADMM [11] algorithm because it decomposes the original problem into several small subproblems with fast convergence speed and high accuracy. The augmented Lagrangian function of (12) is defined as

$$\begin{aligned} L_\mu(\mathcal{B}, \mathcal{T}, \mathcal{W}, \mathcal{Y}) &= \sum_{i=1}^3 \beta \|\mathcal{B}_i\|_* + \lambda \|\mathcal{T} \odot \mathcal{W}\|_1 \\ &\quad + \langle \mathcal{Y}, \mathcal{B} + \mathcal{T} - \mathcal{D} \rangle + \frac{\mu}{2} \|\mathcal{B} + \mathcal{T} - \mathcal{D}\|_F^2 \end{aligned} \quad (13)$$

where \mathcal{Y} is the Lagrange multiplier, and $\langle \cdot \rangle$ denotes the inner product of the two tensors. Then the augmented Lagrangian function can be separated into several subproblems and they are updated in the $(k+1)$ th iteration

$$\mathcal{B}^{k+1} = \operatorname{argmin}_{\mathcal{B}} \sum_{i=1}^3 \beta \|\mathcal{B}_i\|_* + \frac{\mu^k}{2} \left\| \mathcal{B} + \mathcal{T}^k - \mathcal{D} + \frac{\mathcal{Y}^k}{\mu^k} \right\|_F^2 \quad (14)$$

$$\mathcal{T}^{k+1} = \operatorname{argmin}_{\mathcal{T}} \lambda \|\mathcal{T} \odot \mathcal{W}^k\|_1 + \frac{\mu^k}{2} \left\| \mathcal{B}^{k+1} + \mathcal{T} - \mathcal{D} + \frac{\mathcal{Y}^k}{\mu^k} \right\|_F^2. \quad (15)$$

Equation (14) can be solved by singular value shrinkage operator [12]

$$\mathcal{B}_i^{k+1} = \mathfrak{D}_{\frac{3\beta}{\mu^k}} \left(\mathcal{D}_i - \mathcal{T}_i^k - \frac{\mathcal{Y}_i^k}{\mu^k} \right), \quad i = 1, 2, 3. \quad (16)$$

Equation (15) can be solved by soft-threshold operator [13]

$$\mathcal{T}^{k+1} = S_{\frac{\lambda\mathcal{W}^k}{\mu^k}} \left(\mathcal{D} - \mathcal{B}^{k+1} - \frac{\mathcal{Y}^k}{\mu^k} \right). \quad (17)$$

\mathcal{Y} and μ update in the standard way

$$\mathcal{Y}^{k+1} = \mathcal{Y}^k + \mu^k (\mathcal{D} - \mathcal{B}^{k+1} - \mathcal{T}^{k+1}) \quad (18)$$

$$\mu^{k+1} = \min(\rho \mu^k, \mu^{\max}). \quad (19)$$

The entire optimization process and the parameters used are shown in Algorithm 1.

IV. EXPERIMENTAL RESULTS AND COMPARISONS

First, extensive experiments were conducted to evaluate the influence of key parameters on the detection performance.

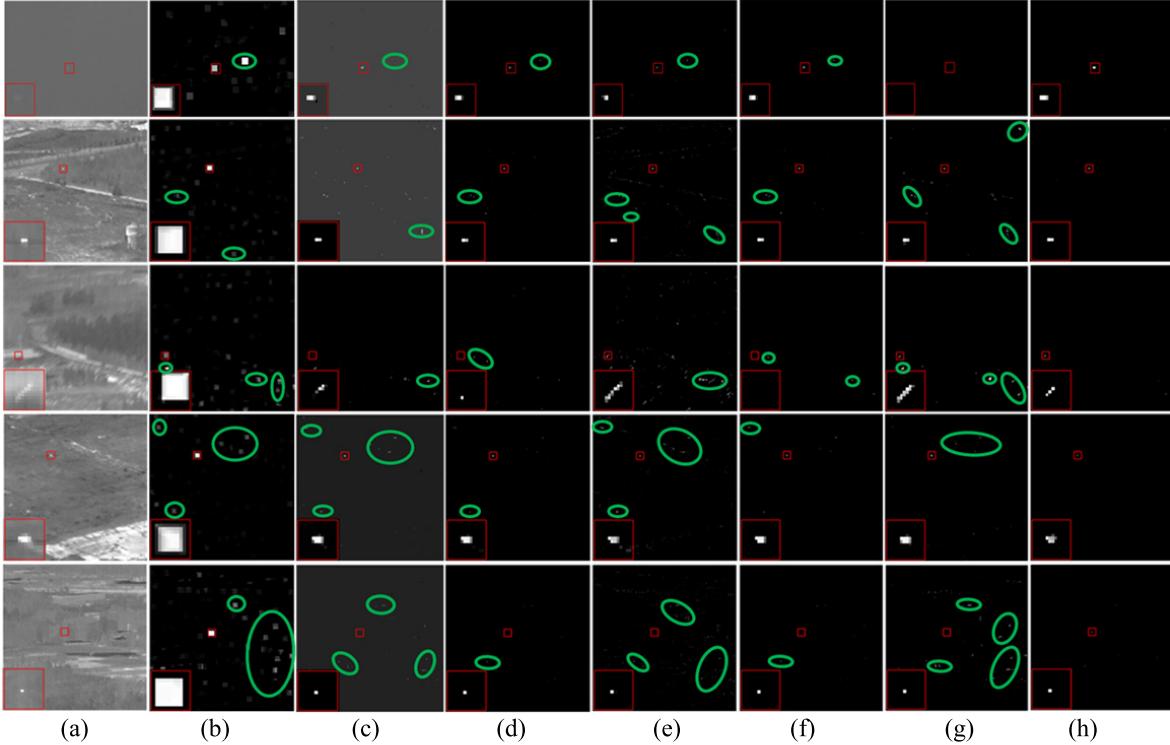


Fig. 4. Detection results of five sequences using different seven methods. (From left to right) Column (a)–(h) show the original image, detection images using relative local contrast measure (RLCM), IPI, NIPPS, RIPT, NRAM, PSTNN, and proposed GIPT methods, respectively. Targets are marked with red square boxes and false alarm areas are marked with green oval boxes.

TABLE I
DETAILS OF THE FIVE SEQUENCES

	Frames	Size	Target number	Average SCR
Seq.1	48	208x152	1	2.60
Seq.2	194	256x256	1	0.91
Seq.3	395	256x256	1	0.83
Seq.4	359	256x256	1	2.99
Seq.5	192	256x256	1	0.71

Then, we used five sequences for testing and the details are shown in Table I. All the experiments were implemented with MATLAB R2016a in Windows 10 based on AMD Ryzen 5 4600H 3.00 GHz CPU with 16G of RAM.

Evaluation indicators include the background suppression factor (BSF), signal-to-clutter gain (SCRG), and receiver operating characteristic (ROC) curves. The definitions of SCRG and BSF are as follows:

$$\text{SCRG} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}}}, \quad \text{BSF} = \frac{\sigma_{\text{in}}}{\sigma_{\text{out}}} \quad (20)$$

where SCR_{out} and SCR_{in} are the SCR [5] values of the reconstructed target image and the original image, respectively. σ_{in} and σ_{out} are the standard deviation of the original image and the reconstructed target image, respectively. The ROC curves can be drawn according to the False-alarm ratio (Fa) and the Probability of detection (Pd) as follows:

$$\text{Fa} = \frac{\text{number of detected false targets}}{\text{total number of pixels in the whole image}} \quad (21)$$

$$\text{Pd} = \frac{\text{number of detected true targets}}{\text{total number of real targets}} \quad (22)$$

where Fa is abscissa and Pd is ordinate.

A. Analyses on Key Parameter

λ is a compromising parameter that usually affects the robustness of different scenes. On one hand, reducing the false

alarm rate of the target image requires a larger λ , on the other hand, increasing the detection rate of the target image requires a smaller λ .

To evaluate the effect of λ on the detection performance, the ROC curves with different values of λ_l are tested and shown in Fig. 3. From Fig. 3(a), we see the detection ability first increases and then decreases with the increase of λ_l . Therefore, we limited the maximum value of λ_l to 2. Since it controls the tradeoff between the target patch-tensor and the background patch-tensor, we associated it with the image-patch number and set: $\lambda_l = \min((L \times \text{length}(\text{line-patches})/\text{length}(\text{point-patches})), 2)$. From Fig. 3(b), we see the ROC curves reach the highest first when $L = 0.6$, so we choose this value.

B. Comparison of Other Algorithms

To verify the superiority of the proposed algorithm, the six state-of-the-art algorithms are included for comparison and the parameters used in the experiments are optimal values in corresponding literature, and we chose the best patch for each sequence in our method.

Fig. 4 shows the detection results that reflect the great suppressing to strong edges in our algorithm. From the target image, whether it is full of point noise or strong edges in the background, our algorithm handles it well and there is almost no background residual. Table II shows the average SCRG and BSF for sequences one–five. From the final results, our method has superior performance than other methods, especially in target enhancement. Through Fig. 5 we can see that the Fa using our method has decreased a lot compared with others.

It is hard to coordinate real-time and good detection. Therefore, we discuss the computational efficiency of the seven methods as shown in Table III. The size of the image, vectorized matrix, and image-patch tensor are $M \times N, m \times$

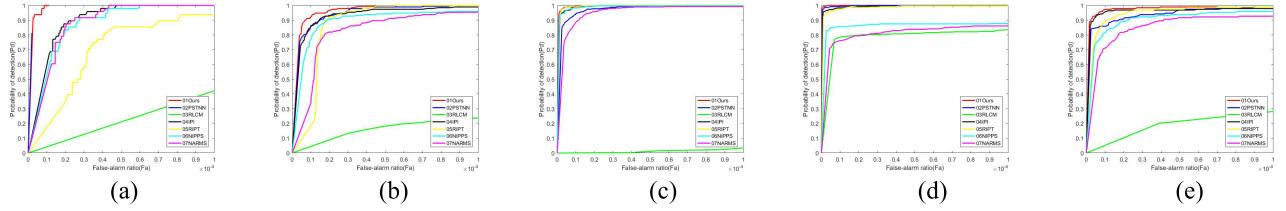


Fig. 5. ROC curves of seven algorithms for (a) Seq.1, (b) Seq.2, (c) Seq.3, (d) Seq.4, and (e) Seq.5.

TABLE II
SCRG AND BSF VALUES OF SEVEN METHODS

Method	Seq1		Seq2		Seq3		Seq4		Seq5	
	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
RLCM[14]	3.11	0.04	4.86	0.37	13.92	0.61	2.83	1.83	6.62	0.39
IPI[1]	195.49	44.02	750.44	344.28	5328.04	576.25	298.21	293.53	154.71	110.44
NIPPS[2]	236.86	60.93	1418.72	794.76	6334.73	1127.06	192.58	286.51	2027.92	1354.92
RIPT[4]	197.57	62.50	357.30	192.52	2405.83	182.30	20.39	20.68	68.51	47.68
NRAM[3]	<u>327.81</u>	59.70	1612.54	<u>925.30</u>	<u>7067.95</u>	<u>1340.50</u>	441.29	611.24	1867.24	1366.57
PSTNN[5]	202.27	52.68	<u>1970.22</u>	638.27	6891.97	518.65	897.38	<u>664.47</u>	454.93	209.72
Ours	470.05	<u>62.38</u>	2091.19	1240.74	9309.18	1918.24	<u>735.83</u>	894.89	2206.07	1059.45

TABLE III
COMPARISON OF COMPLEXITY AND AVERAGE RUNNING TIME (IN SECONDS) OF THE SEVEN METHODS

Method	Complexity	Seq1	Seq2	Seq3	Seq4	Seq5
RLCM	$O(S^3 MN)$	3.33	10.06	10.04	10.00	10.10
IPI	$O(mn^2)$	1.28	7.45	6.80	7.28	7.55
NIPPS	$O(mn^2)$	1.62	7.65	7.20	7.58	7.03
RIPT	$O(n_1 n_2 n_3 (n_1 n_2 + n_2 n_3 + n_1 n_3))$	0.73	5.38	5.36	4.88	6.04
NRAM	$O(mn^2)$	1.03	4.61	4.42	4.14	4.24
PSTNN	$O(n_1 n_2 n_3 \log(n_1 n_2 n_3) + n_1 n_2^2 (n_3 + 1)/2)$	0.39	0.80	0.81	0.77	0.82
Ours	$O(n_1 n_2 l' \log(n_1 n_2 l') + n_1 n_2^2 (l' + 1)/2)$	0.06	0.13	0.13	0.15	0.15

n , and $n_1 \times n_2 \times n_3$, respectively. S is the processing window scale. After grouping, the size of the three patch-tensors are $n_1 \times n_2 \times l_1$, $n_1 \times n_2 \times l_2$, and $n_1 \times n_2 \times l_3$ where $l_1 + l_2 + l_3 = n_3, l' = \max(l_1, l_2, l_3)$. Naturally, the time complexity of our algorithm can be reduced and verified by our faster running time.

V. CONCLUSION

To cope with the strong edges in the background, we proposed a GIPT model and incorporated it with a target weight. Experiments show that our method has an overwhelming advantage in running time, enhancing the target, and suppressing the background. We can conclude that our method is suitable for images with a complex environment and small targets.

However, the proposed method is sensitive to the patch size of the image suffering from the destroyed local structure by unsuitable size. Therefore, the construction of adaptive patch size is our future work.

REFERENCES

- [1] C. Q. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4996–5009, Dec. 2013.
- [2] Y. Dai, Y. Wu, Y. Song, and J. Guo, "Non-negative infrared patch-image model: Robust target-background separation via partial sum minimization of singular values," *Infr. Phys. Technol.*, vol. 81, pp. 182–194, Mar. 2017.
- [3] L. Zhang, L. Peng, T. Zhang, S. Cao, and Z. Peng, "Infrared small target detection via non-convex rank approximation minimization joint $l_{2,1}$ norm," *Remote Sens.*, vol. 10, no. 11, 2018.
- [4] Y. Dai and Y. Wu, "Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3752–3767, Aug. 2017.
- [5] L. Zhang and Z. Peng, "Infrared small target detection based on partial sum of the tensor nuclear norm," *Remote Sens.*, vol. 11, no. 4, p. 382, Feb. 2019.
- [6] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis with a new tensor nuclear norm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 925–938, Jan. 2020.
- [7] Z. Cao, X. Kong, T. Zang, Q. Zhu, S. Cao, and Z. Peng, "Infrared dim target detection via mode-k1k2 extension tensor tubal rank under complex ocean environment," *ISPRS J. Photogramm. Remote Sens.*, vol. 181, pp. 167–190, Dec. 2021.
- [8] M. Zhao, W. Li, L. Li, P. Ma, Z. Cai, and R. Tao, "Three-order tensor creation and tucker decomposition for infrared small-target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2021, Art. no. 5000216.
- [9] X. Kong et al., "Infrared small target detection via nonconvex tensor fibered rank approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2021, Art. no. 5000321.
- [10] F. F. Zhou, Y. Wu, and Y. Y. Dai, "Infrared small target detection via incorporating spatial structural prior into intrinsic tensor sparsity regularization," *Digit Signal Proc.*, vol. 111, no. 5, Apr. 2021, Art. no. 102966.
- [11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Nov. 2010.
- [12] J. Cai, Cand, E. J. Xe, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, 2010.
- [13] E. T. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for l_1 -minimization: Methodology and convergence," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1107–1130, 2008.
- [14] J. Han, K. Liang, B. Zhou, X. Zhu, J. Zhao, and L. Zhao, "Infrared small target detection utilizing the multiscale relative local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 612–616, Apr. 2018.