TNLRS: Target-Aware Non-local Low-Rank Modeling with Saliency Filtering Regularization for Infrared Small Target Detection

Hu Zhu, *Member, IEEE,* Haopeng Ni, Shiming Liu, Guoxia Xu, *Member, IEEE,* and Lizhen Deng, *Member, IEEE,*

Abstract—Recently, infrared small target detection problem has attracted substantial attention. Many works based on local low-rank model have been proven to be very successful for enhancing the discriminability during detection. However, these methods construct patches by traversing local images and ignore the correlations among different patches. Although the calculation is simplified, some texture information of the target is ignored, and targets of arbitrary forms cannot be accurately identified. In this paper, a novel target-aware method based on a non-local low-rank model and saliency filter regularization is proposed, with which the newly proposed detection framework can be tailored as a non-convex optimization problem, therein enabling joint target saliency learning in a lower dimensional discriminative manifold. More specifically, non-local patch construction is applied for the proposed target-aware low-rank model. By combining similar patches, we reconstruct them together to achieve a better generalization of non-local spatial sparsity constraints. Furthermore, to encourage target saliency learning, our proposed saliency filtering regularization term based on entropy is restricted to lie between the background and foreground. The regularization of the saliency filtering locally preserves the contexts from the target and surrounding areas and avoids the deviated approximation of the low-rank matrix. Finally, a unified optimization framework is proposed and solved with the alternative direction multiplier method (ADMM). Experimental evaluations of real infrared images demonstrate that the proposed method is more robust under different complex scenes compared with some state-of-the-art methods.

Index Terms—Infrared Small Target Detection, Non-local Low-rank Modeling, Saliency Filter Regularization, ADMM.

I. INTRODUCTION

INFRARED small target detection is a key technology for many applications, including maritime surveillance systems [1], missile tracking and interception systems [2], and forest warning systems [3]. Generally, infrared small target detection is still a challenging problem mainly due to 1): the target is always small and does not have any other textures or shape features, and 2): the complex background has a low signal-to-clutter ratio caused by cluttered noise. Recently, many
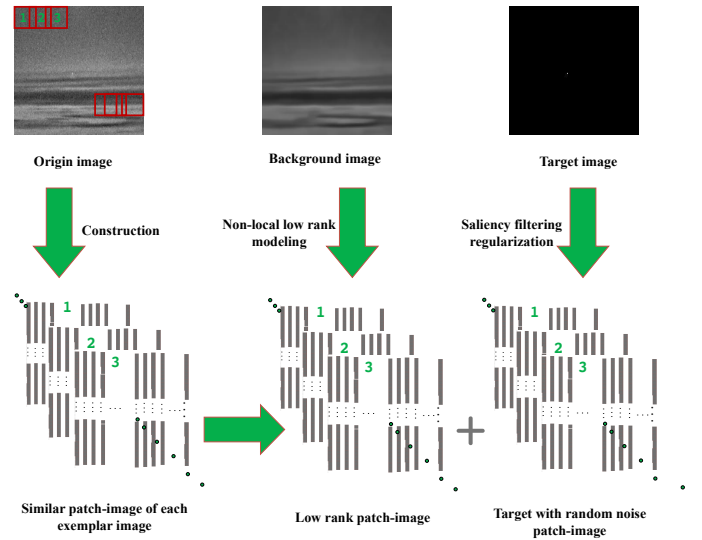
Fig. 1. The overall framework of TNLRS is presented. The first line shows the original image, the separated background image and the target image, and the second line presents a flow chart of the proposed method.

traditional detection methods have been proposed such as three-dimensional matched filters [4] and multi-scale adaptive matched filters [5]. These methods exploit the spatial-temporal information of the target, which is tractable in the specific case of static backgrounds. However, in the practical application of infrared search and track (IRST) systems [6]–[8], the imaging background changes rapidly due to the high speed of many targets such as aircrafts and missiles. The performance of the local low-rank model is rapidly degraded. Therefore, research on infrared small target detection is of great significance and has attracted widespread attention in recent years.

Small target detection methods are divided into two categories based on prior knowledge of the target awareness domain: local background consistency prior [9], [10] and non-local background self-correlation prior [11]. In the former, it is noticeable that the local background is slowly transitioned, and the neighboring pixels are highly correlated. Thus, the targets would be highlighted in a smooth manner. Based on this assumption, methods have been proposed by comparing a pixel or a region with its neighbors such as the Top-Hat filter [12], [13], the Max-mean/Max-median filter [14], and the local entropy model (LEM) [15]. However, these methods always enhance the edges of the sea surface and

make it difficult to determine whether it is the target or the edge of the cloud during the detection process. To solve this problem, Deng et al. [16] proposed to differentiate the real target by weighting the multi-scale local difference contrast (WLDM). By weighting the local information entropy of the image [17], local areas have different importances even if they have similar pixel distributions. In addition, some methods based on the human visual system (HVS) [18] have been proposed regarding the saliency of the target such as the local contrast measure (LCM) [11], multi-scale patch-based contrast measure (MPCM) [19], novel weighted image entropy (NWIE) [20], derivative entropy-based contrast measure (DECM) [21], and entropy-based window selection model (EWSM) [22]. These methods generally assume that the background has local stability and that the target moves slowly. However, image backgrounds are not always spatially invariant and these background patches would have the similar characteristic across the local regions.

For the latter method, it is assumed that all background patches commonly share a single subspace or a mixture of low-rank subspace clusters [23]–[25]. Therefore, the segmentation of the target from the background can be seen as a recovery task from a noisy image under the perspective of saliency guiding. This problem can be effectively solved via principal component analysis (PCA) [26] for, e.g., the infrared patch-image model (IPI) [27]. At the same time, the complex noise in an actual scene may also be considered a sparse component by the IPI model, which produces numerous false positives. To solve this issue, Wang et al. [28] proposed a patch image model with local and global analysis (PILGA) to constrain the sparsity of noise patch images. However, the performance of these methods degrades rapidly while considering more heterogeneous backgrounds. Later, He et al. [29] proposed a low-rank and sparse representation model under the multi-subspace hypothesis. On the other hand, these methods select the pixels of the entire image to form a new low-rank matrix, which massively increases the computational complexity [30]–[32].

Based on the above discussion, the non-local background self-correlation prior-based methods have gained increasing attention from researchers for practical applications. On the one hand, the local algorithm analysis of the image can reduce the amount of calculation in target detection, but it greatly increases the false alarm rate. Current non-local algorithms [33]–[35] apply convolutional neural networks for target detection. Although this method is highly efficient, it requires a large number of data sets for training. On the other hand, local analysis can lead to the loss of the integrity of the image, and the original link between patches cannot be restored. For example, a typical local detection algorithm, such as [36], relies on a fixed-size encoding method to encode each image local patch. While there are similar background patches in different areas, they cannot match and show the connections among themselves. Non-local algorithms [37] do not have such problems, and the only cost is an increase in the amount of calculation. However, in practical use, this portion of the calculation time can be neglected, and the detection accuracy rate can be greatly improved. At the same time, by combining the links between the patches (such as in low-rank modeling), the amount of computation of the algorithm is also greatly reduced.

Recently, the saliency prototype has become a new research domain in the field of small target detection. The property of saliency in an image is determined as the local contrast of the region with respect to its neighborhood at various scales [38], [39]. The global consideration of an image can assign an approximate significance to unique regions like the attention mechanism of the human visual system, which can prominently highlight the target. The relational entropy-based saliency detection in images, such as [40], takes the statistical property of the Rayleigh quotient via a Pseudo-Wigner-Ville distribution, and the saliency weight of the target is too low, resulting in a thick cloud and target that cannot be distinguished on the resulting graph. Although the background is not static and the prior information of the target is not available, the saliency of the image can be achieved by enhancing the target and predicting the background. For the above such problems, we can consider increasing the saliency threshold of the target and combining the advantages of regularization to prevent model overfitting and transfer the detection of a target into an optimization problem. Image saliency [41] is the best way to distinguish between a target and the background. A small target is considered to be the area that destroys the visual continuity of the original background image. In the target-aware algorithm, we use image saliency as an important indicator for detecting the small targets. Thus, we design an entropy-based saliency filtering for small target detection, and the targets should have certain global and local characteristics. In other words, the targets should be independent in color and ultimately aggregate in spatial distribution [42], whereas the background should be the opposite. However, the application of local contrast methods is not very effective because the segmentation results readily classify some foreground targets into the background region, resulting in unsatisfactory final results.

In this paper, a novel method for infrared small target detection is proposed, and the overall framework is shown in Fig. 1. We propose to generalize the traditional infrared small target detection model [43] to the new infrared patch-based small target detection model by non-local low-rank modeling with saliency filtering regularization. In practice, for each exemplary image patch, we look for a set of image patches with similar conditions in the input image to form a data matrix. Each similar patch has a similar structure. Thus, the rank of the data matrix is low. Specifically, the size of the infrared target is small compared to the whole image. In light of the definition of local entropy in this paper, we can find that the local entropy of the target satisfies the sparsity. Overall, the detection task of an infrared small target is considered an optimization problem for recovering the sparse and low-rank matrices. We have listed the main contributions of this paper as follows:

- We propose to use the target-aware non-local low-rank model instead of the traditional target detection model, which reduces the influence of random noise on the real target and improves the performance of the target

detection.

- Saliency is determined as the local contrast of an image region with respect to its neighborhood at various scales. Inspired by [44] to more efficiently exploit the non-local sparsity of infrared images, we propose to take the saliency filtering regularization in the non-local low-rank model to enhance the anti-interference and the generalization property.
- We conduct the traditional low-rank and sparse recovery model for infrared target detection and introduce the detailed solution process based on the alternating direction multiplier method (ADMM) for optimization.

The organization of the remainder of this paper is presented as follows: In the second chapter, we introduce in detail the theoretical basis of our model. In the third chapter, we introduce in detail the proposed model and the related optimization solution process. In the fourth chapter, we perform many experiments and verify the superiority of our model from various aspects. In the fifth chapter, our conclusions and opinions are given.

II. PRIOR WORK ON INFRARED SMALL TARGET DETECTION

In this section, we introduced the theoretical basis of non-local low-rank and saliency filtering regularization-based TNLRS for infrared small target detection.

A. Classic low-rank matrix method

In the target detection algorithm, the detection time will greatly affect the actual application effect. To solve this problem, the low-rank matrix is applied to infrared small target detection. If the rank of the matrix is much smaller than its number of rows or columns, then such a matrix is called a low-rank matrix(as shown in Fig. 2). If the input data matrix is composed of two features(such as the background and the target), where one of which has a sparse property and the other has a low-rank characteristic, then one can recover the low-rank component of the matrix by the convex optimization method.

The global low-rank saliency detection algorithm first reconstructs the image foreground significant target according to the natural image foreground target and the background brightness and color difference; then, the low-rank decomposition is used to suppress the non-significant region in the image. The algorithm is divided into three steps: contrast extraction, initial saliency map generation and global low-rank decomposition. This method performs global low-rank decomposition on the initial saliency map of the image to obtain the low-rank part and the sparse part of the initial saliency map. The calculation method is as follows:

$$\min rank(f_B) + ||f_T||_0 \qquad (1)$$

where f_B and f_T correspond to the low-rank and sparse parts respectively. The global saliency map is obtained by subtracting the low-rank portion after decomposition from the initial saliency map (as shown in Fig. 3).
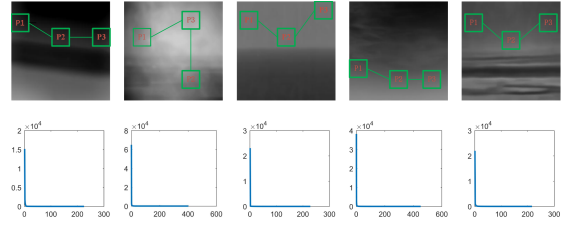

Fig. 2. The illustration of the non-local low-rank property. The first row shows five representative background images. The second row includes the singular values of the corresponding background patch images.
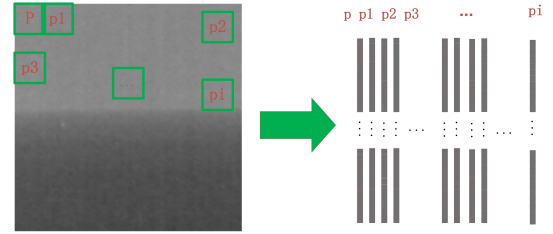

Fig. 3. Constructing similar patch images for any exemplar patch. The green boxes on the left are the samples of the original image, and on the right are the previously converted patches

However, there is a problem with this approach. If there are too many clouds in the background, some background information will be treated as foreground targets, resulting in false alarms. This is also one aspect of the improvement in our proposed algorithm, and the details are described in section 3.

B. Detection Based on Entropy Method

The main challenge of infrared small target detection is lacking sufficient information about targets. However, due to the long imaging distance, the size of the infrared target that we observe is small and without any shape or texture features. Because the small infrared targets continue moving, their size(in pixels) may vary from 2×2 to 8×8 (in Fig. 4). However, compared to the whole image, the infrared target image (f_T) can be regarded as a sparse matrix. Therefore, the detection of infrared small targets can be considered as a recovery task of a sparse matrix. The $l_0 - norm$ represents the number of non-zero entries and thus can represent the sparsity of the matrix. Specifically,

$$||f_T||_0 < T \qquad (2)$$

where $||*||_0$ denotes the $l_0 - norm$, and T is a positive number and satisfies $T \ll m \times n$ ($m \times n$ is the size of the target image). However, the solution of $l_0 - norm$ is a NP-hard problem. Fortunately, this problem has been effectively solved via convex relaxation optimization, using the $l_1 - norm$ instead of the $l_0 - norm$. Specifically,

$$||f_T||_1 < T \qquad (3)$$

where $||*||_1$ denotes the $l_1 - norm$. In practice, due to camera and weather restrictions, pictures taken contain noise, which affects the final detection performance. In the absence of the spatial-temporal information and shape of the target, the characteristics between the target and the surrounding area of the
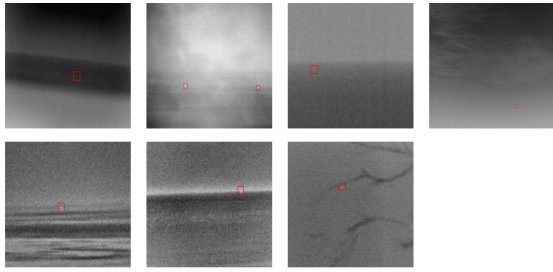
4

Fig. 4. The seven representative real infrared images. The background includes the sky, sea level and desert, also possessing interference by heavy fog, rain and clouds.
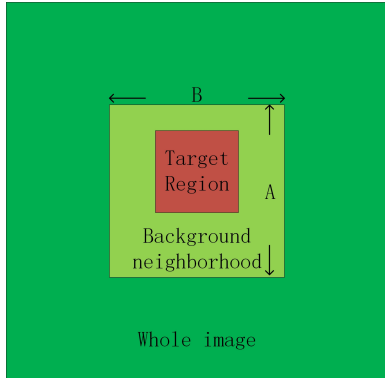

Fig. 5. The target local window consists of the target area and adjacent background areas.

target are important priors of infrared small target detection. After observing a large number of infrared image(Fig. 4), we observe that the small infrared targets are mostly located in homogeneous and bright areas. How to use these priors to design a model that can segment noise and targets is our next problem to be solved.

Information entropy reflects how much information is contained in the information source, which can represent the global characteristics of the information source. The most famous image entropy is based on image histograms, which is a method used to effectively illustrate the complexity of gray value distributions in whole images. For an image (I), its image entropy formulation can be written as

$$H(I) = \sum_{i=0}^{s-1} q(x_i) [-\log(q(x_i))] \tag{4}$$

$$q(x_i) = m_i/M \tag{5}$$

where s is the gray level of the image, x_i is the ith gray value, m_i is the number of pixels of the ith gray value, and M is the total number of image pixels.

III. THE PROPOSED METHOD FOR INFRARED TARGET DETECTION

In this section, we first introduce the non-local low-rank matrix recovery and saliency filtering regularization. Then, we discuss the infrared small target detection model by optimizing the recovery task of the low-rank and sparse matrices. Based on this, we propose a new model via non-local low-rank and local entropy regularization for target detection. The proposed

model consists of two components: patch recombination for the self-similarity of low-rank background images and the sparse constraint of the target image local entropy. We consider the infrared small target detection task as the optimization problem of recovering the sparse and low-rank matrix. Furthermore, the corresponding algorithm is explained in detail.

A. The Non-local Low-Rank Matrix Recovery

In general, the background is slowly transitional; thus, we suspect that there is a large correlation between different local patches in the infrared image regardless of the distance between the pixels, as illustrated in the first line of Fig. 2. Although patches $P1$, $P2$, and $P3$ are located in different areas of the infrared image, their structures are extremely similar. Such an assumption implies that a sufficient number of similar patches can be found for any example patch of size $\sqrt{m} \times \sqrt{m}$ at position i denoted by $x_i \in C^m$. For each exemplar patch x_i, we only look for all similar patches that satisfy Eq. (6) in the K neighborhood of its local window, i.e.,

$$S_i = \{i_j | \|x_i - x_{ij}\|^2 \le Q\} \tag{6}$$

where S_i denotes the set of those similar patches and Q is a threshold that determines the size of the set. We arrange these similar patches into column vectors and reconstitute a new matrix $X_i = [x_{i_0}, x_{i_1}, x_{i2}, \ldots, x_{i_{n-1}}]$, where n denotes the number of similar patches, as illustrated in Fig. 3. To verify our guess, we performed an experiment. In the second line of Fig. 2, the singular values of the background patch image (f_B) quickly drop to zero, which implies that the background patch image is an intrinsically low-rank matrix. Specifically,

$$rank(f_B) < k \tag{7}$$

where k is a constant determined by background, k constrains the complexity of the background image, and the value of k is larger for a complex background than for an uniform background. Actually, X_i may contain noise, which could break the low-rank of X_i. One possible solution is to divide the matrix X_i into $X_i = W_i + L_i$, where W_i and L_i represent a Gaussian noise matrix and a low-rank matrix, respectively. However, classic nuclear norm minimization (NNM) tends to over-shrink the rank components and treats all singular values equally, limiting its applicability for real applications [30], [45]. Therefore, we consider a smooth but non-convex surrogate of the rank [43], [46] rather than the nuclear norm. Specifically,

$$M(Y,\mu) = \log det(Y + \mu I) \tag{8}$$

where μ is a small constant value. Fig. 6 shows a comparison of a non-convex substitution function, rank and nuclear norm in the scalar case, which confirms that the function $M(Y,\mu)$ can be closer to the rank minimization optimization than the nuclear norm. For a general matrix L_i, that is neither square nor positive semi-definite, we slightly modify Eq. (8) as follows:

$$L(L_i,\mu) = \log det\left(\left(L_i L_i^T\right)^{1/2} + \mu I\right) \tag{9}$$

Then, the low-rank approximation problem can be solved as follows:

$$L_i = \arg \min L\left(L_i, \mu\right) \quad \text{s.t.} \quad \left\|X_i - L_i\right\|_F^2 \leq \zeta^2 \tag{10}$$

where ζ^2 denotes the variance of additive Gaussian noise. Eq. (11) can be solved in its Lagrangian form, e.g., as follows:

$$L_i = \arg \min \left\|X_i - L_i\right\|_F^2 + \lambda L\left(L_i, \mu\right) \tag{11}$$

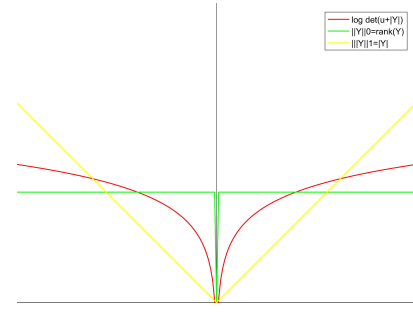
Fig. 6. Comparison of $L(Y, u)$, $rank(Y)$ and the nuclear norm in the case of a scalar, where red represents $\log det(u + |Y|)$, green represents $\|Y\|_0 = rank(Y)$ and yellow represents $\|Y\|_1 = Y$.

B. Saliency Filtering based on Image Entropy

Image entropy is a global statistic that does not characterize the spatial texture or frequency information of an image. The local entropy operator can measure the amount of information contained in the local window, which can represent the texture and frequency information of the image to some extent (Fig. 4). In this paper, the local entropy operator is defined as

$$H(I) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} f(i,j) \log(f(i,j)) \tag{12}$$

where $M \times N$ is the local window size, and $f(i,j)$ is the gray value at the point (i,j). Therefore, the local entropy value obtained by substituting Eq. (12) is consistent with sparsity (Fig. 7). Its sparse constraint operator is as follows:

$$\left\|f_T \odot \log f_T\right\|_1 \tag{13}$$

where $\odot$ denotes the point-wise product (Hadamard product) of the two matrices and f_T denotes the infrared target image.

In this paper, we assume that image noise is Gaussian noise that is disorderly distributed. Thus, we have

$$\left\|f_D - f_B - f_T\right\|_F^2 < \zeta^2 \tag{14}$$

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

Under this assumption, from the definition of local entropy, the entropy value is greater in bright regions than in dark regions. Compared with the noise distribution, the pixel value distribution of multi-pixel targets located in uniform and bright areas is relatively close, and a larger entropy value is obtained (Fig. 7). To some extent, local entropy operators can be employed to enhance infrared small targets for complex backgrounds.

C. The Formulation of Small Target Detection

Generally, the low-rank and sparse matrix recovery model for target detection can be written as

$$\min_{f_B, f_T} rank\left(f_B\right) + \gamma \left\|f_T\right\|_0, \quad \text{s.t.} f_B + f_T = f_D \tag{15}$$

where f_D, f_B, f_T and γ are the original input infrared image, the infrared background image, the infrared target image, and a positive weight constant, respectively. The model in Eq. (15) can be transformed as the following convex optimization problem by PCA:

$$\min_{f_B, f_T} \left\|f_B\right\|_* + \gamma \left\|f_T\right\|_1, \quad \text{s.t.} \quad f_B + f_T = f_D \tag{16}$$
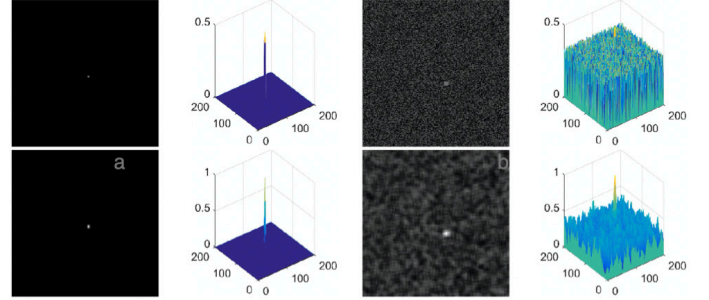

Fig. 7. The first column represents the target image (upper) and entropy map (under), while the 3D gray distributions are shown in the second column. The third column represents the target image with Gaussian noise (upper) and entropy map (under), while the 3D gray distributions are shown in the fourth column.

where $\|*\|_*$ denotes the nuclear norm of the matrix.

Based on the above model, to more effectively approximate the non-local low-rank component of the background and reduce the impact of random noise on the detection performance, we propose a novel model:

$$\min_{f_T, f_B, L_i} \eta \sum_i \left\{ \left\|\tilde{R}_i f_B - L_i\right\|_F^2 + \lambda L\left(L_i, \mu\right) \right\} \\ + \alpha \left\|f_T \odot \log f_T\right\|_1 + \beta \left\|f_D - f_B - f_T\right\|_F^2 \tag{17}$$

where $\tilde{R}_i f_B$ denotes the matrix formed by the set of similar patches for every exemplar patch x_i. The proposed non-local low-rank regularization can utilize the non-convexity of the similar patch image minimization; the local entropy regularity utilizes the difference between the noise and the target. The experimental results for different original infrared images show that the proposed method achieves a better detection performance than some baseline methods.

The Fig. 1 shows the entire method of small target detection proposed in this paper. First, the patch image is constructed from the input original infrared image. Second, for the recovery of the background image, we enforce the regularization of the low-rank property on the non-local low-rank patch sets for each exemplar patch along with the constraint of being linear. Third, for the recovery of the target image, we enforce the regularization of the sparse property on the local entropy along with the constraint of linear measurements. In this paper, our model can be solved effectively via ADMM. The algorithm is detailed in Algorithm 1.
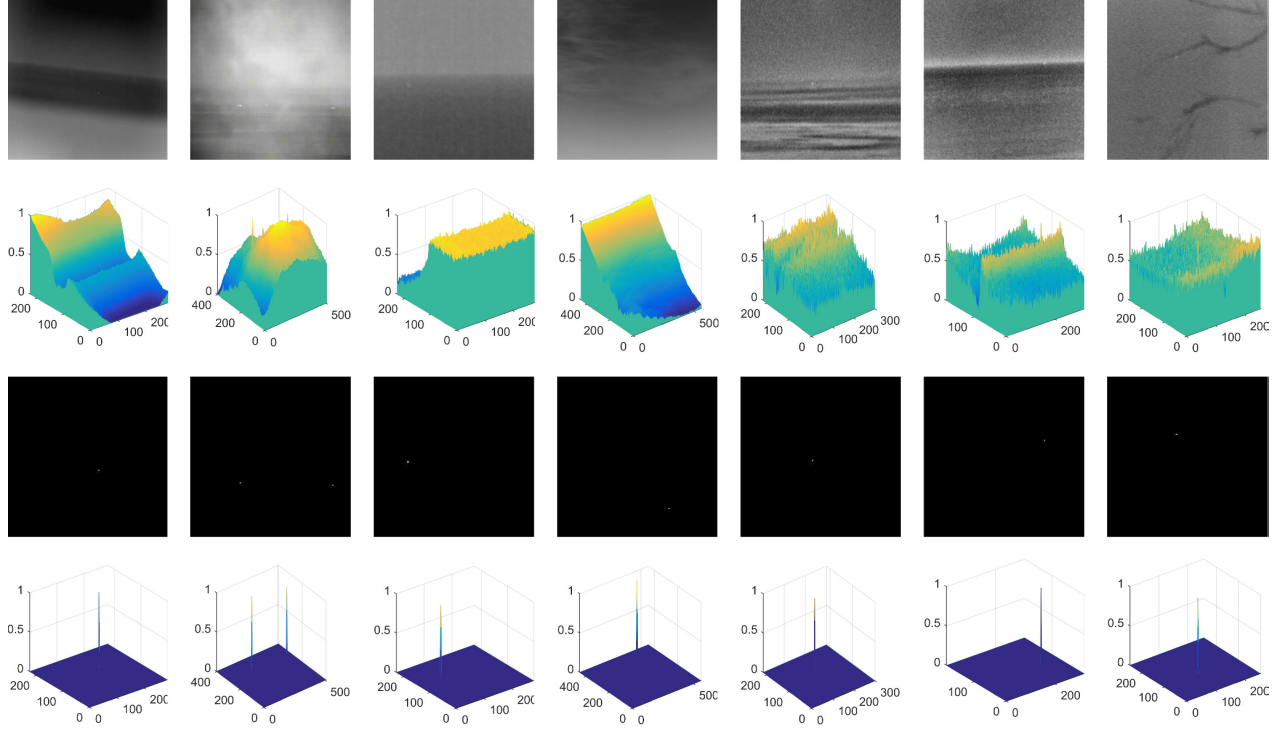
Fig. 8. The experimental results for target enhancement. The first line is the original image sequence, the second line is the gray histogram of the original image, the third line is the image when the background is removed, and the fourth line is the corresponding histogram.

D. The Solution of Low-Rank Matrix L_i

For the input infrared image, we first extract exemplar patches x_i every l pixels along each direction. Then, we can obtain a matrix X_i for every x_i by grouping a set of similar patches, as described in Section II. The solution of step 4 in Algorithm 1 is described below:

$$L_i = \underset{L_i}{\arg\min} \eta \left\| \tilde{R}_i f_B - L_i \right\|_F^2 + \lambda L (L_i, \mu) \qquad (18)$$

Because $L(L_i, \mu)$ is approximately the sum of the logarithm of singular values, Eq. (18) can be rewritten as

$$L_i = \underset{L_i}{\arg\min} \|X_i - L_i\|_F^2 + \frac{\lambda}{\eta} \sum_{j=1}^{l} \log(\sigma_j(L_i) + \mu) \qquad (19)$$

where $\tilde{R}_i B = X_i$, $l = min(m, n)$ and $\sigma_j(L_i)$ denote the jth singular value of L_j. Eq. (19) can be solved by iteratively solving

$$L_i^{k+1} = \underset{L_i}{\arg\min} \|X_i - L_i\|_F^2 + \frac{\lambda}{\eta} \sum_{j=1}^{l} \frac{\sigma_j}{\sigma_j^k + \varepsilon} \qquad (20)$$

For convenience, we can rewrite Eq. (20) as follows:

$$L_i^{k+1} = \underset{L_i}{\arg\min} \frac{1}{2} \|X_i - L_i\|_F^2 + \tau\varphi\left(L_i, w^{(k)}\right) \qquad (21)$$

where $\tau = \frac{\lambda}{2\eta}$ and $\varphi\left(L_i, w^{(k)}\right) = \sum_{j=1}^{l} w_j^{(k)} \sigma_j$ denote the weighted nuclear norm with weights $w_j^{(k)} = \frac{1}{\sigma_j^k + \varepsilon}$. Note that because the singular values σ_j are ordered in descending

order, the weights are ascending. Then, the solution of the reconstructed matrix in the $(k+1)th$ iteration is obtained by

$$L_i^{k+1} = U\left(\tilde{\Sigma} - \tau \operatorname{diag}\left(w^{(k)}\right)\right)_+ V^T \qquad (22)$$

where $\tilde{\Sigma}V^T$ is the SVD of X_i, and $(a)_+ = max(a, 0)$. The detailed proof is given in the next section.

TABLE I
ALGORITHM 1: TARGET DETECTION VIA LOW-RANK REGULARIZATION AND LOCAL ENTROPY

Step 1: Input an infrared image f_D.
Step 2: Initialization
a) Setting parameters $\eta = 0.35, \lambda = 0.3$.
b) Grouping a set of similar patches S_i for each exemplar patch x_i.
Step 3: While not converged do
Step 4: For L_i: $\underset{L_i}{\arg\min} \eta \left\| \tilde{R}_i f_B - L_i \right\|_F^2 + \lambda L(L_i, \mu)$.
Step 5: For f_B: $f_B = \underset{f_B}{\arg\min} \eta \sum_i \left\| \tilde{R}_i f_B - L_i \right\|_F^2 + \beta \|f_D - f_B - f_T\|_F^2$.
Step 6: For f_T: $f_T = \underset{f_T}{\arg\min} \alpha \|f_T \odot \log f_T\|_1 + \beta \|f_D - f_B - f_T\|_F^2$.
Step 7: Setting $N = f_T \odot \log f_T$.
Step 8: For N: $N = \underset{N}{\arg\min} \alpha \|N\|_1 + \frac{\gamma}{2}\|f_N - f_T \odot \log f_T\|_F^2$.
Step 9: End While.
Step 10: Output target image f_T and background image f_B.

E. The Solution of Background Image f_B

After solving for each L_i, we can reconstruct the background image by solving the following minimization problem:
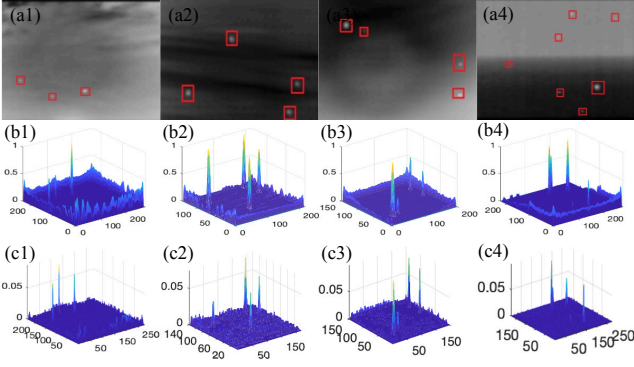
Fig. 9. Comparison of model logarithmic specification and nuclear specification. (a1)-(a4) are the original image sequences, (b1)-(b4) are the test conducted by the nuclear specification, and (c1)-(c4) are the logarithmic test.

$$\arg\min_{f_B}\eta\sum_i\left\|\tilde{R}_if_B-L_i\right\|_F^2+\beta\left\|f_D-f_B-f_T\right\|_F^2 \quad (23)$$

For fixed L_i and f_T, f_B admits a closed-form solution:

$$f_B=\left(\eta\sum_i\tilde{R}_i^T\tilde{R}_i+\beta I\right)^{-1}\left(\eta\sum_i\tilde{R}_i^TL_i+\beta\left(f_D-f_T\right)\right) \quad (24)$$

where $\tilde{R}_i^T\tilde{R}_i=\sum_{r=0}^nR_r^TR_r$, and $\tilde{R}_i^TL_i=\sum_{r=0}^{n-1}R_r^Tx_{ir}$. Note that the term $\tilde{R}_i^T\tilde{R}_i$ is a diagonal matrix. Each of the entries in the diagonal matrix corresponds to an image pixel location, and its value is the number of overlapping patches that cover the pixel location. The term $\tilde{R}_i^T\tilde{L}_i$ denotes the patch average result, averaging over all collected similar patches for each exemplar patch. Therefore, Eq. (24) can easily computed in one step.

F. The Solution of Target Image f_T

The solution of step 6 in algorithm 1 is described below:

$$\arg\min_{f_T}\alpha\left\|f_T\odot\log f_T\right\|_1+\beta\left\|f_D-f_B-f_T\right\|_F^2 \quad (25)$$

To facilitate the calculation of the optimal solution of Eq. (25), we introduce a variable N and make $N=f_T\odot\log f_T$. Thus, Eq. (25) can be written as

$$\arg\min_{f_r}\alpha\|N\|_1+\beta\left\|f_D-f_B-f_T\right\|_F^2+\frac{\gamma}{2} \quad (26)$$
$$s.t.N=f_T\odot\log f_T$$

Eq. (26) can be transformed into an unconstrained optimization problem:

$$\arg\min_{f_T}\alpha\|N\|_1+\beta\left\|f_D-f_B-f_T\right\|_F^2+\frac{\gamma}{2}\left\|N-f\odot\log f\right\|_F^2 \quad (27)$$

Then, the problem of solving f_T is transformed into solving two sub-problems.

- For N sub-problem,

$$\arg\min_N\alpha\|N\|_1+\frac{\gamma}{2}\left\|N-f_T\odot\log f_T\right\|_F^2 \quad (28)$$

The minimizer is as follows:

$$N=\text{shrinkage}\left(f_T\odot\log f_T,\frac{\alpha}{\gamma}\right) \quad (29)$$

where $shrinkage(a,b)=sign(a)\odot\max\{abs(a)-b,0\}$.

- For f_T sub-problem,

$$\arg\min_{f_T}\beta\left\|f_D-f_B-f_T\right\|_F^2+\frac{\gamma}{2}\left\|N-f_T\odot\log f_T\right\|_F^2 \quad (30)$$

To obtain the optimal solution of Eq. (30), the corresponding gradient can be obtained as follows:

$$\beta\frac{\partial}{\partial f_T}\left(\|f_D-f_B-f_T\|_F^2\right)+\frac{\gamma}{2}\frac{\partial}{\partial f_T}\left(\|N-f_T\odot\log f_T\|_F^2\right)=0 \quad (31)$$

The problem of solving the Frobenius norm of the matrix can be transformed into determining the trace of the matrix. The general form is as follows:

$$F(X)=\alpha\|X-Y\|_F^2=\alpha(\sqrt{\text{tr}\{(X-Y)^H(X-Y)\}})^2$$
$$=\alpha\,\text{tr}\left(X^HX\right)-\text{tr}\left(Y^HX\right)-\text{tr}\left(X^HY\right)+\text{tr}\left(Y^HY\right) \quad (32)$$

and

$$\frac{\partial F(X)}{\partial X}=\frac{\partial\,\text{tr}\left(X^HX\right)-\text{tr}\left(Y^HX\right)-\text{tr}\left(X^HY\right)+\text{tr}\left(Y^HY\right)}{\partial X}$$
$$=\alpha(X+X-Y-Y)=2\alpha(X-Y) \quad (33)$$

The corresponding iteration can be written as follows:

$$f_T^{(n+1)}=-\frac{\gamma}{2\beta}\left(f_T^{(n)}\odot\log f_T^{(n)}-N^{(n+1)}\right)\odot\left(1+\log f_T^{(n)}\right)$$
$$+f_D^{(n+1)}-f_B^{(n+1)} \quad (34)$$

First, in Algorithm 1, after obtaining an improved estimate of the background image, the low-rank matrices L_i can be updated by solving Eq. (20). Second, the combination of the updated L_i and target image is then used to improve the estimate of the background image by solving Eq. (22). Third, the estimate of the target image can be obtained by using an updated background image to solve Eq. (32). Such a process is iterated until convergence. The overall procedure is described below as Algorithm 1.

IV. EXPERIMENTAL RESULTS

In this section, we use seven real infrared images to illustrate the performance of the proposed method. First, we introduce the evaluation metrics, the baseline methods and information of the datasets in this paper. Next, we discuss the target enhancement performance and the effects of the parameters by performing real experiments. Finally, we perform experiments on the infrared images with different basic methods and confirm the superiority of the proposed method.

A. Evaluation Metrics, Baseline Methods and Datasets

Evaluation Metrics. The signal-to-clutter ratio gain and the background suppression factor are commonly used indicators to measure the performance of small target detection methods. The formulation can be defined as follows:

$$SCRG=\frac{(S/C)_{out}}{(S/C)_{in}} \quad (35)$$

$$BSF=\frac{C_{in}}{C_{out}} \quad (36)$$

where S is the signal amplitude, C is the clutter standard deviation, and C_{in} and C_{out} are the clutter standard deviation of the input image and the output image, respectively. From the
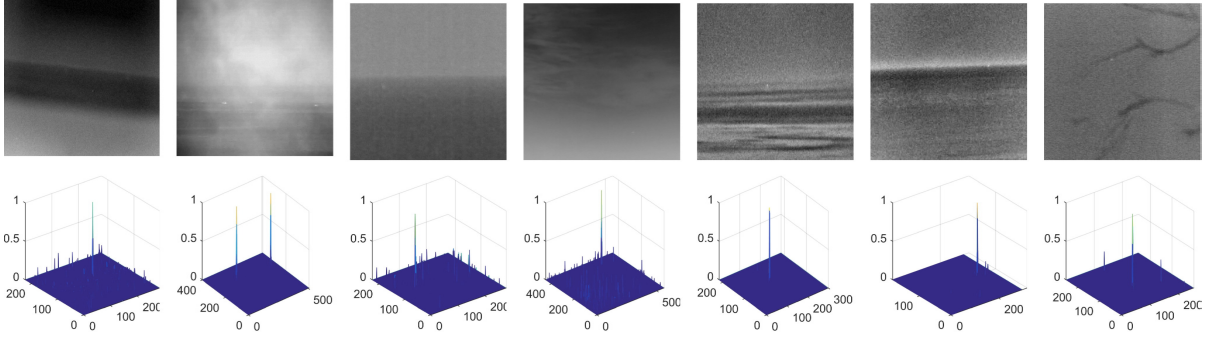
8

Fig. 10. The target enhancement experimental results with Gaussian noise. The first line is the enhanced image, and the second line is the corresponding gray histogram.
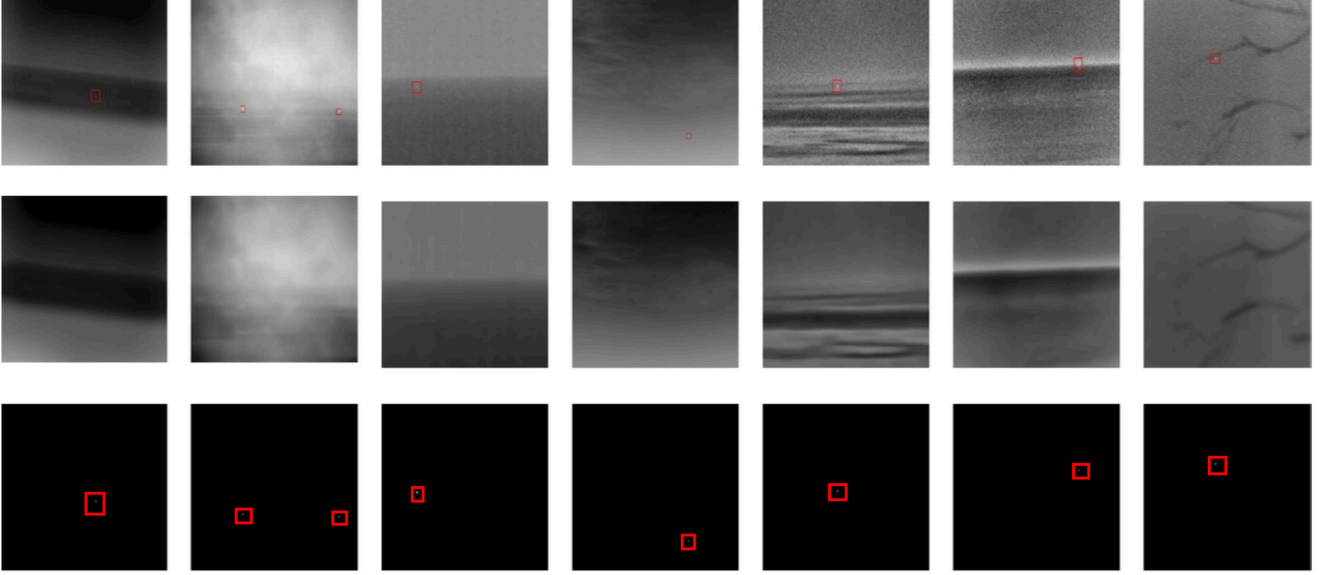

Fig. 11. The result of the separation of the background images and the target images from the real infrared images. The first line is the original image, the second line is the predicted background image, and the third line is the target image.

above definition, we can see that the larger $SCRG$ and BSF values indicate better small target enhancement and background suppression, respectively. The probability of detection (Pd) and false alarm rate (Fa) are important indicators for evaluating the performance of infrared small target detection methods. Pd indicates the probability that the real target detected in the infrared image exists, while Fa indicates the probability that the target is detected in an infrared image in which the target is not present. These can be defined as follows:

$$Pd = \frac{number\ of\ true\ infrared\ targets}{number\ of\ actual\ infrared\ targets} \quad (37)$$

$$Fa = \frac{number\ of\ pixels\ in\ false\ infrared\ targets}{number\ of\ pixels\ in\ all\ test\ infrared\ images} \quad (38)$$

The above two indicators range between 0 and 1. From the above definition, we find that the greater the value of Pd is, the more complete the targets, whereas a smaller value of Fa indicates more target detection.

Baseline Methods. To prove the superiority and robustness of the proposed method, we use different baseline methods for comparison.

- Local adaptive contrast operation based on regularized feature reconstruction(**LACRFR**) [44]: This method uses the regularization of reconstructed features to construct adaptive operators.
- Infrared patch-image model(**IPI**) [27]: IPI exploits the low-rank of non-local backgrounds and the sparseness of targets to translate target detection problems into low-rank and sparse recovery problems.
- Local contrast measure(**LCM**) [11]: LCM is a method of using the local entropy around the target.
- Multi-scale patch-based contrast measure(**MPCM**) [19]: MPCM highlights the target by comparing different sizes of image blocks.
- Novel weighted image entropy(**NWIE**) [16]: NWIE represents a new way to measure the operator weights.

Datasets. To more efficiently reflect the robustness of our proposed method, we use seven real infrared images with a complex background to compare the proposed method with
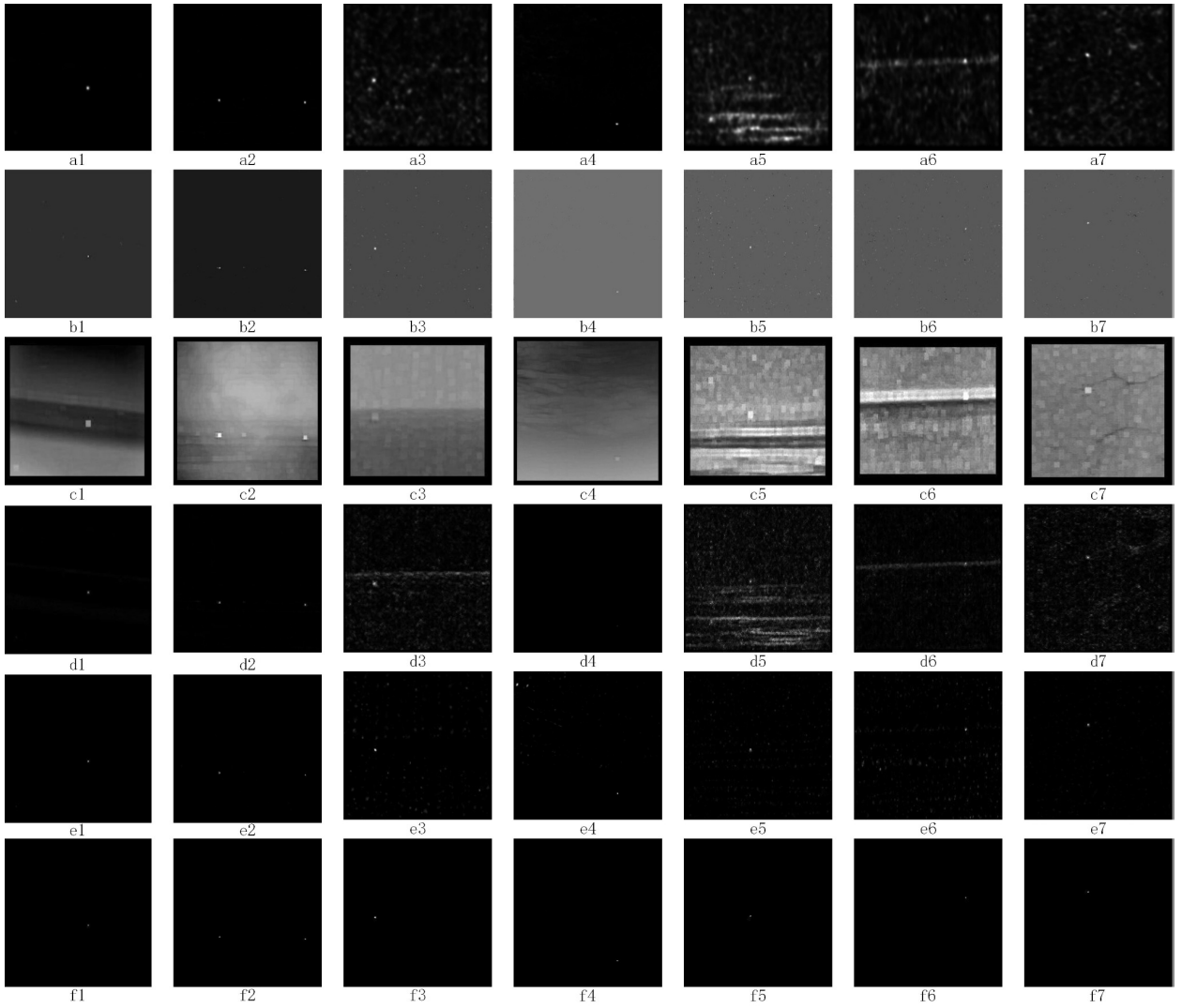
Fig. 12. The filtered results by different methods. $(a1)$-$(a7)$, $(b1)$-$(b7)$, $(c1)$-$(d7)$, $(e1)$-$(e7)$ and $(f1)$-$(f7)$ are the enhanced results obtained through the LACRFR, IPI, LCM, MPCM, NWIE and TNLRS.

the baseline methods. Fig. 4 shows representative images of seven infrared images.

B. Quantitative Analysis of Target Enhancement

The purpose of infrared target detection is to efficiently enhance the infrared targets and suppress the background and noise. If TNLRS can better separate the background and noise from the input image, it will be easier to detect small targets and achieve higher accuracy. To more effectively prove the performance of the proposed detection method for target enhancement, we conducted experiments on seven infrared images with complex backgrounds. The experimental results are shown in Fig. 8.

The first row of Fig. 8 denotes seven representative infrared images randomly derived from different complex scenes. The second row of Fig. 8 indicates the corresponding 3D gray distributions of infrared images. We found it difficult to determine the position and shape of the targets due to cluttered

background and noise. The third row of Fig. 8 shows the position and shape of the targets after applying the proposed method, and the corresponding 3D gray distributions of the targets are shown in the fourth row of Fig. 8. From Fig. 8, we can clearly see that TNLRS effectively suppresses the background and noise while effectively enhancing the target.

Some small target detection methods based on image entropy operations are sensitive to noise. To more effectively prove the robustness to the target enhancement performance, we add Gaussian white noise based on the original input image, as shown in the first row of Fig. 10. However, from the description of Table I, we found that most of the input images contain heavy noise. Therefore, we have made appropriate adjustments to the noise variance of special images. The corresponding 3D gray distributions of targets after TNLRS are shown in the second row of Fig. 10. After preprocessing the input images, it is no longer possible to observe whether there are small targets with the naked eye, and then we start to

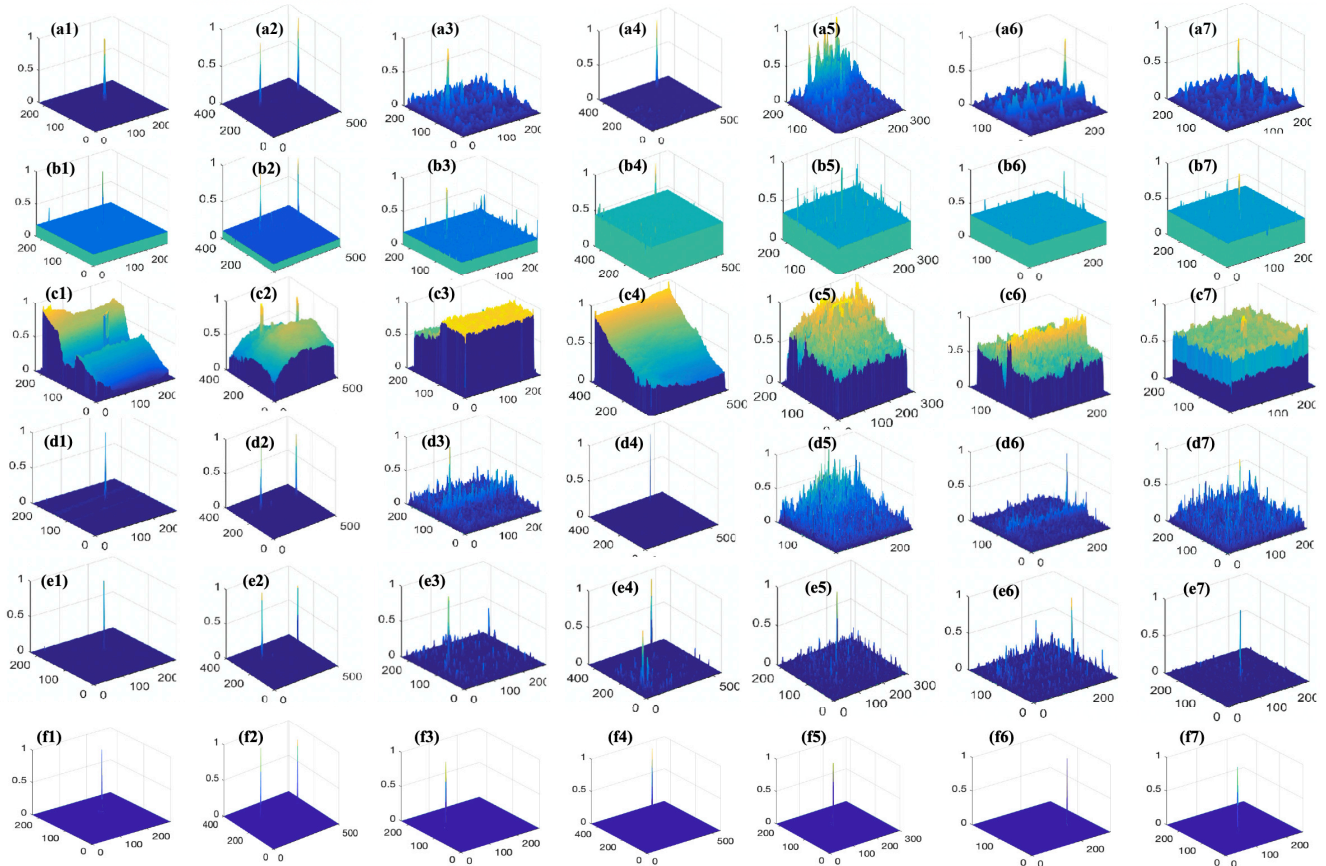Fig. 13. The 3D gray distributions by some baseline methods and our proposed method, where $(a1)$-$(a7)$, $(b1)$-$(b7)$, $(c1)$-$(d7)$, $(e1)$-$(e7)$ and $(f1)$-$(f7)$ are the results obtained through the LACRFR, IPI, LCM, MPCM, NWIE and TNLRS methods.
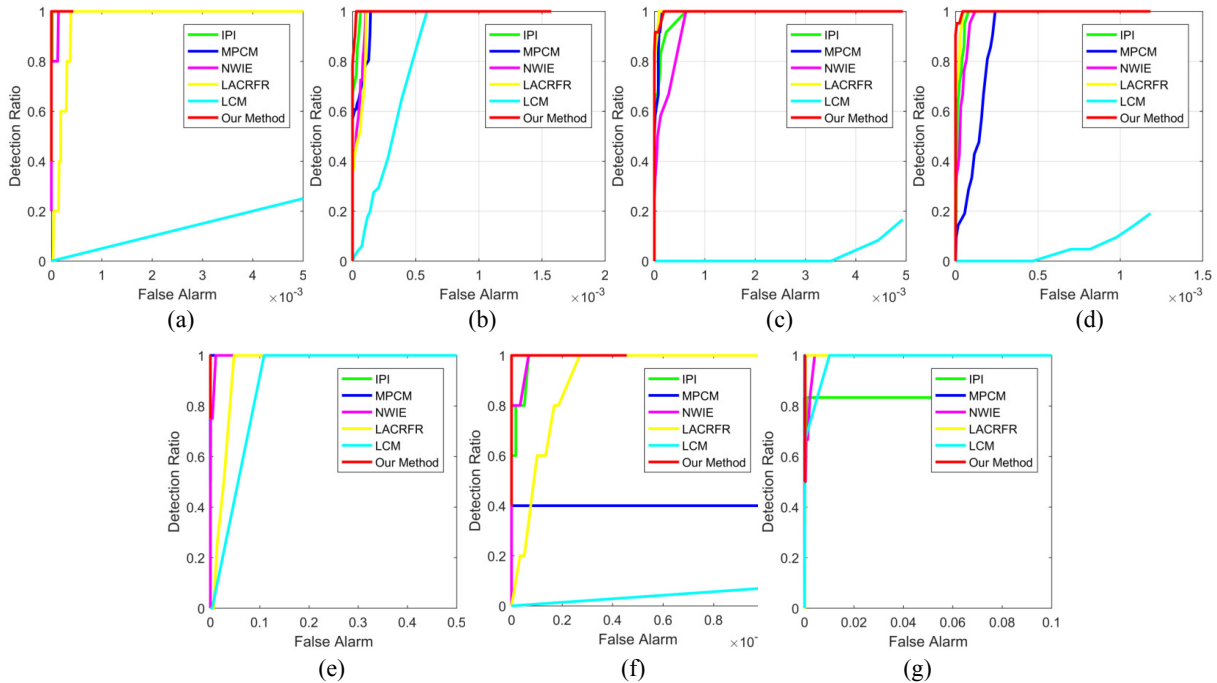

Fig. 14. The characteristic (Roc) curves of six methods for the seven real images, where (a)-(g) correspond to the seven sets of actual image sequences.

TABLE II
THE VALUES OF SCRG AND BSF OBTAINED THROUGH DIFFERENT METHODS

	LACRFR		IPI		LCM		NWIE		MPCM		TNLRS	
	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
1	90.914	22.649	104.9	33.736	1.160	0.975	80.87	403.9	105.5	80.86	**174.69**	**568.5**
2	**28.314**	20.029	24.52	22.509	1.211	0.601	22.48	**208.8**	27.70	21.97	26.393	130.1
3	12.955	9.640	18.61	9.346	0.556	0.508	5.428	445.0	12.98	62.38	**31.861**	**529.3**
4	38.914	30.701	43.62	24.461	0.849	0.844	11.08	281.5	19.77	**444.75**	54.349	12.50
5	10.033	1.588	7.060	5.094	0.997	0.335	1.440	7.460	8.591	443.6	**24.153**	**636.8**
6	3.073	2.683	8.978	7.957	0.653	0.359	3.581	9.579	5.338	**482.7**	**34.337**	198.6
7	4.882	3.067	7.721	4.420	0.347	0.197	1.722	16.14	8.103	**520.2**	**14.966**	125.7

TABLE III
THE VALUES OF SCRG AND BSF OBTAINED FOR DIFFERENT VALUES OF λ

	$\lambda = 0.1$		$\lambda = 0.2$		$\lambda = 0.3$		$\lambda = 0.4$		$\lambda = 0.5$		$\lambda = 0.6$	
	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
1	150.23	500.71	171.57	512.23	**174.69**	**568.5**	169.43	532.6	155.5	550.86	166.34	548.83
2	23.231	118.96	25.673	129.33	**26.393**	**130.1**	24.32	128.53	23.26	126.27	22.13	124.1
3	29.32	498.51	29.88	511.61	**31.861**	**529.3**	30.11	509.51	29.05	510.35	26.74	511.89
4	49.53	11.12	51.25	11.89	**54.349**	**12.50**	53.79	12.22	52.48	11.28	43.18	11.59
5	22.174	22.814	23.834	629.32	**24.153**	**636.8**	23.171	612.17	22.96	604.83	22.99	611.67
6	31.62	189.23	**35.75**	**199.73**	34.337	198.6	34.521	199.51	33.56	190.65	33.16	188.84
7	13.64	123.13	**15.16**	125.5	14.966	**125.7**	14.72	124.73	14.68	124.89	14.21	121.67

experiment with our proposed method and various comparison methods.

C. The Overall Performance Comparison of Target Detection

To more effectively prove the performance of our proposed method, we selected five different basic methods for comparison: LACRFR, IPI, LCM, MPCM, and NWIE.

First, the effects of matrix recovery are shown in Fig. 11. The red squares in the first row of Fig. 11 indicate the edges and location of the infrared target. The complex backgrounds can be recovered completely, as shown in the second row of Fig. 11. The corresponding target images are shown in the third row of Fig. 11. These results indicate that the task of low-rank and sparse matrix recovery can be solved efficiently by our proposed method.

Second, the filtered target results and corresponding 3D gray distributions of targets under various methods on complex backgrounds are shown in Fig. 12 and Fig. 13. From Fig. 13 a1-f1 and a2-f2, we observe that some dim small targets are well enhanced. However, there is still a heavy background and noise in Fig. 13 a3-e3, a5-e5, a6-e6 and a7-e7. Comparing Fig. 13 f1-f7, we clearly observe that the results produced by our proposed method contain very little clutter and noise residuals and are suitable for various complex backgrounds, therein also demonstrating stable robustness.

Third, $SCRG$ and BSF are effective measures for the performance of infrared small target detection. The value of $SCRG$ indicates the performance of enhancing dim targets, while the value of BSF implies the ability to suppress complex backgrounds. For Fig. 4, the details of $SCRG$ and BSF are shown in Table II obtained using LACRFR, IPI, LCM, NWIE, MPCM and the proposed method. The details in Table II show that our proposed method is superior to some baseline methods. The experimental results suggest that TNLRS can effectively enhance the dim target while suppressing background clutter and noise.

Finally, for the seven infrared small target images, the Pd obtained using TNLRS are 1.00, 0.96, 0.91, 0.94, 1.00, 1.00 and 1.00, and the Fa are 0.0040, 0.0039, 0.0098, 0.0044, 0.0024, 0.0045 and 0.0073, respectively. The characteristic curve is a graphical plot of the probabilities of detection versus the false alarm rates. We provide ROC curves obtained using the baseline methods and proposed method for the seven infrared images in Fig. 14. We can see that TNLRS outperforms the baseline methods, which implies that TNLRS is more robust to various clutter and noisy backgrounds.

D. The Analysis of Computational Complexity

Here, we discuss the computational complexity of our model with a detailed specification of the running environment. As shown in Algorithm I, the algorithm complexity is composed of two main parts: saliency filtering regularization and low-rank model reconstruction. Here, we define that the image size is $M \times N$, m and n are the rows and columns of the input image.

For saliency filtering regularization, the computational complexity is mainly determined by the saliency filtering. We need to traverse the original image and find a pixel block similar to the one in the operation window; thus, the operation can be computed in $O(m \times n \times R^2)$. For the low-rank matrix construction, we need to convert the original matrix into a list of column vectors for low-rank transitions. The complexity is $O(m^2 \times n^2 \times p)$, where p is the number of patches.

Based on the analysis above, the complexity of our proposed algorithm is $O(m \times n \times R^2 + m^2 \times n^2 \times p)$. We perform the time consumption test using MATLAB R2016a on a laptop with an Intel Core i5-4210 CPU and 4 GB of RAM.

E. Convergence and Parameter Analysis

In this section, we discuss the parameters that we selected in the proposed method. As mentioned above, regularity measures the degree to which a function is smooth. The higher the
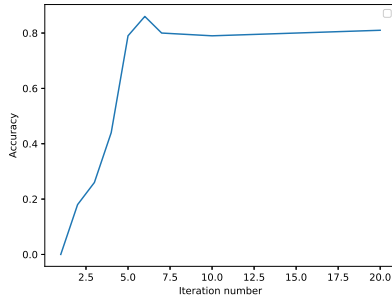
Fig. 15. The convergence of our proposed TNLRS.

regularity is, the smoother the function. Using the regularization method will automatically weaken the unimportant feature variables, automatically extracting important feature variables from many feature variables and reducing the magnitude of the feature variables.

To obtain the best detection results, we use different parameter settings for different images. We have listed the set of different η-values and conducted comparative tests with λ = 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6. The experimental results are shown in Table III. Here, we visually see the λ value corresponding to the highest value of the indicator for different image data. For image 1, the best value for η is 0.35. For image 6 and 7, the highest value of the index corresponds to η equal to 0.3. After weighting all the experimental results, we chose $\eta = 0.35$ in this paper. To demonstrate the convergence of our proposed TNLRS algorithm, we provide the convergency result in Fig. 15. We can observe that our proposed method TNLRS converges very quickly. In general, TNLRS only needs to iterate 5 to 8 times in our test.

V. CONCLUSION

This paper proposes a novel model for infrared small target detection based on the non-local low rank modeling and saliency filtering regularization, which can be efficiently solved via ADMM. Combining the non-local low-rank property of the background and the sparsity of the target local entropy image, the experiments have demonstrated that our proposed method have a superior performance compared with the other methods. From the quantitative analysis, we can observe that our proposed algorithm not only significantly increases the SCRG and BSF values of the image but also achieves high accuracy and low false alarm rates.

REFERENCES

[1] X. Bai, Z. Chen, Y. Zhang, Z. Liu, and Y. Lu, "Spatial information based fcm for infrared ship target segmentation," in *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2014, pp. 5127–5131.

[2] S. Kim and J. Lee, "Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track," *Pattern Recognition*, vol. 45, no. 1, pp. 393–406, 2012.

[3] Y. Dai and Y. Wu, "Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 10, no. 8, pp. 3752–3767, 2017.

[4] I. S. Reed, R. M. Gagliardi, and L. B. Stotts, "Optical moving target detection with 3-d matched filtering," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 4, pp. 327–336, 1988.

[5] K. A. Melendez and J. W. Modestino, "Spatiotemporal multiscan adaptive matched filtering," in *Signal and Data Processing of Small Targets 1995*, vol. 2561. International Society for Optics and Photonics, 1995, pp. 51–65.

[6] N. T. Thanh, H. Sahli, and D. N. Hao, "Infrared thermography for buried landmine detection: Inverse problem setting," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 12, pp. 3987–4004, 2008.

[7] C. Gao, T. Zhang, and Q. Li, "Small infrared target detection using sparse ring representation," *IEEE Aerospace and Electronic Systems Magazine*, vol. 27, no. 3, pp. 21–30, 2012.

[8] Y. Gu, C. Wang, B. Liu, and Y. Zhang, "A kernel-based nonparametric regression method for clutter removal in infrared small-target detection applications," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 3, pp. 469–473, 2010.

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

[10] G. Wang, T. Zhang, L. Wei, and N. Sang, "Efficient method for multiscale small target detection from a natural scene," *OptEn*, vol. 35, pp. 761–768, 1996.

[11] C. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 574–581, 2013.

[12] M. Zeng, J. Li, and Z. Peng, "The design of top-hat morphological filter and application to infrared target detection," *Infrared Physics & Technology*, vol. 48, no. 1, pp. 67–76, 2006.

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

[14] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," in *Signal and Data Processing of Small Targets 1999*, vol. 3809. International Society for Optics and Photonics, 1999, pp. 74–83.

[15] B. Zhou, Y.-z. WANG, L.-h. SUN, and Y.-q. HE, "Study on local entropy used in small target detection," *Acta Photonica Sinica*, vol. 37, no. 2, p. 381, 2008.

[16] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Small infrared target detection based on weighted local difference measure," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 7, pp. 4204–4214, 2016.

[17] X. Qu, H. Chen, and G. Peng, "Novel detection method for infrared small targets using weighted information entropy," *Journal of Systems engineering and electronics*, vol. 23, no. 6, pp. 838–842, 2012.

[18] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," in *2007 IEEE Conference on computer vision and pattern recognition*. Ieee, 2007, pp. 1–8.

[19] Y. Wei, X. You, and H. Li, "Multiscale patch-based contrast measure for small infrared target detection," *Pattern Recognition*, vol. 58, pp. 216–226, 2016.

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

[21] X. Bai and Y. Bi, "Derivative entropy-based contrast measure for infrared small-target detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 4, pp. 2452–2466, 2018.

[22] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Entropy-based window selection for detecting dim and small infrared targets," *Pattern Recognition*, vol. 61, pp. 66–77, 2017.

[23] B. Du, L. Ru, C. Wu, and L. Zhang, "Unsupervised deep slow feature analysis for change detection in multi-temporal remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 9976–9992, 2019.

[24] B. Du, Y. Wang, C. Wu, and L. Zhang, "Unsupervised scene change detection via latent dirichlet allocation and multivariate alteration detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 12, pp. 4676–4689, 2018.

[25] B. Du, M. Zhang, L. Zhang, R. Hu, and D. Tao, "Pltd: Patch-based low-rank tensor decomposition for hyperspectral images," *IEEE Transactions on Multimedia*, vol. 19, no. 1, pp. 67–79, 2016.

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

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

[28] H. Wang, F. Yang, C. Zhang, and M. Ren, "Infrared small target detection based on patch image model with local and global analysis," *International Journal of Image and Graphics*, vol. 18, no. 01, p. 1850002, 2018.

[29] Y. He, M. Li, J. Zhang, and Q. An, "Small infrared target detection based on low-rank and sparse representation," *Infrared Physics & Technology*, vol. 68, pp. 98–109, 2015.

[30] Y. Xie, S. Gu, Y. Liu, W. Zuo, W. Zhang, and L. Zhang, "Weighted schatten p-norm minimization for image denoising and background subtraction," *IEEE transactions on image processing*, vol. 25, no. 10, pp. 4842–4857, 2016.

[31] V. Rao, A. Sandu, M. Ng, and E. D. Nino-Ruiz, "Robust data assimilation using l_1 and huber norms," *SIAM Journal on Scientific Computing*, vol. 39, no. 3, pp. B548–B570, 2017.

[32] W. He, H. Zhang, H. Shen, and L. Zhang, "Hyperspectral image denoising using local low-rank matrix recovery and global spatial–spectral total variation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 3, pp. 713–729, 2018.

[33] Z. Luo, A. Mishra, A. Achkar, J. Eichel, S. Li, and P.-M. Jodoin, "Non-local deep features for salient object detection," in *Proceedings of the IEEE Conference on computer vision and pattern recognition*, 2017, pp. 6609–6617.

[34] 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," *Infrared Physics & Technology*, vol. 81, pp. 182–194, 2017.

[35] J. D. Victor and M. M. Conte, "Local image statistics: maximum-entropy constructions and perceptual salience," *JOSA A*, vol. 29, no. 7, pp. 1313–1345, 2012.

[36] M. Xu, H. Jiang, and S. Watcharawittayakul, "A local detection approach for named entity recognition and mention detection," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2017, pp. 1237–1247.

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

[38] R. Achanta, F. Estrada, P. Wils, and S. Süsstrunk, "Salient region detection and segmentation," in *International conference on computer vision systems*. Springer, 2008, pp. 66–75.

[39] H. Deng, J. Liu, and Z. Chen, "Infrared small target detection based on modified local entropy and emd," *Chinese Optics Letters*, vol. 8, no. 1, pp. 24–28, 2010.

[40] Y. Xu, Y. Zhao, C. Jin, Z. Qu, L. Liu, and X. Sun, "Salient target detection based on pseudo-wigner-ville distribution and rényi entropy," *Optics letters*, vol. 35, no. 4, pp. 475–477, 2010.

[41] Y. Li, Y. Zhang, J.-G. Yu, Y. Tan, J. Tian, and J. Ma, "A novel spatio-temporal saliency approach for robust dim moving target detection from airborne infrared image sequences," *information sciences*, vol. 369, pp. 548–563, 2016.

[42] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang, "Saliency detection via graph-based manifold ranking," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 3166–3173.

[43] M. Fazel, H. Hindi, and S. P. Boyd, "Log-det heuristic for matrix rank minimization with applications to hankel and euclidean distance matrices," in *Proceedings of the 2003 American Control Conference, 2003.*, vol. 3. IEEE, 2003, pp. 2156–2162.

[44] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, "Compressive sensing via nonlocal low-rank regularization," *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3618–3632, 2014.

[45] X. Zhang, J. Chi, J. Hu, L. Liu, and Y. Xing, "Infrared small target detection using modified order morphology and weighted local entropy," in *2nd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2017)*. Atlantis Press, 2016.

[46] D.-Y. Tsai, Y. Lee, and E. Matsuyama, "Information entropy measure for evaluation of image quality," *Journal of digital imaging*, vol. 21, no. 3, pp. 338–347, 2008.

Hu Zhu received the B.S. degree in mathematics and applied mathematics from Huaibei Coal Industry Teachers College, Huaibei, China, in 2007, and the M.S. and Ph.D. degrees in computational mathematics and pattern recognition and intelligent systems from Huazhong University of Science and Technology, Wuhan, China, in 2009 and 2013, respectively. In 2013, he joined the Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests include pattern recognition, image processing, and computer vision. E-mail: peter.hu.zhu@gmail.com

Haopeng Ni received B.S. degree in radio and television engineering from Nanjing University of Posts and Telecommunications, Nanjing, China, in 2019. He is now pursuing his master degree in signal and information processing in Nanjing University of Posts and Telecommunications. His current research interest is object detection. E-mail: nihaopeng1997@126.com

Shiming Liu received B.S. degree in communication engineering from Henan Polytechnic University, JiaoZuo, China, in 2017. He is now pursuing his master degree in electronic and communication in Nanjing University of Posts and Telecommunications. His current research interest is computer vision. E-mail: never66liu@163.com

Guoxia Xu received the B.S. degree in information and computer science from Yancheng Teachers University, Jiangsu Yancheng, China in 2015, and the M.S. degree in computer science and technology from Hohai University, Nanjing, China in 2018. He was a research assistant in City University of Hong Kong and Chinese University of Hong Kong. Now, he is pursing his Ph.D. degree in Department of Computer Science, Norwegian University of Science and Technology, Gjovik Norway. His research interest includes pattern recognition, image processing, and computer vision. E-mail: gxxu.re@gmail.com

Lizhen Deng received the B.S. degree in electronic information science and technology from Huaibei Coal Industry Teachers College, Huaibei, China, in 2007, and the M.S. degree in communication and information systems from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2010. She received her Ph.D. degree in electrical engineering from Huazhong University of Science and Technology, China, in 2014. In 2014, she joined the Nanjing University of Posts and Telecommunications, Nanjing, China. Her current research interests include image processing, computer vision, pattern recognition, and spectral data processing. E-mail: alicedenglzh@gmail.com