

Infrared small target detection via self-regularized weighted sparse model



Tianfang Zhang ^a, Zhenming Peng ^{a,*}, Hao Wu ^a, Yanmin He ^a, Chaohai Li ^b, Chunping Yang ^{a,*}

^a The Laboratory of Imaging Detection and Intelligent Perception, School of Information and Communication Engineering, University of Electronic Science and Technology of China, 610054 Chengdu, China

^b School of Electronic Science and Engineering, University of Electronic Science and Technology of China, 610054 Chengdu, China

ARTICLE INFO

Article history:

Received 22 April 2020

Revised 29 July 2020

Accepted 28 August 2020

Available online 11 September 2020

Communicated by Ma Jiayi

Keywords:

Self-regularize

Subspace cluster

Low rank representation

Sparse constraint

Infrared small target detection

ABSTRACT

Infrared search and track (IRST) system is widely used in many fields, however, it's still a challenging task to detect infrared small targets in complex background. This paper proposed a novel detection method called self-regularized weighted sparse (SRWS) model. The algorithm is designed for the hypothesis that data may come from multi-subspaces. And the overlapping edge information (OEI), which can detect the background structure information, is applied to constrain the sparse item and enhance the accuracy. Furthermore, the self-regularization item is applied to mine the potential information in background, and extract clutter from multi-subspaces. Therefore, the infrared small target detection problem is transformed into an optimization problem. By combining the optimization function with alternating direction method of multipliers (ADMM), we explained the solution method of SRWS and optimized its iterative convergence condition. A series of experimental results show that the proposed method outperforms state-of-the-art baselines.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Infrared search and track (IRST) system uses thermal radiation characteristics of objects for passive detection, which has the advantages of difficult interference, small size and clear imaging. Also, as the main embodiment of IRST performance, infrared small target detection technology has also attracted much attention [1–3]. And it is widely used in military, industrial, traffic, security monitoring, meteorology, medical and other industries. However, infrared images lack color information, and background usually has a lot of clutter and random noise. Moreover, due to the long imaging distance, the target occupies a small area and lacks shape and texture information, which makes this task challenging. Therefore, infrared small target detection is becoming more and more popular in information science and has attracted the attention of many scholars [4–6].

In the past decades, the field of infrared small targets detection has developed rapidly, and many detection methods have been proposed successively. In general, infrared small target detection algorithms can be roughly divided into two categories: detect

before track (DBT) and track before detect (TBD). The idea of TBD method is to combine multi-frame information to calculate the motion trajectory of the target [7–9]. It is inevitable to make assumptions about the movement of the target. However, in the actual situation, due to the influence of detector jitter, relative displacement and other factors, this assumption cannot be perfectly satisfied and the detection effect will be reduced. Moreover, because TBD methods consume massive computing resources, their detection efficiency is also lower. In contrast, DBT methods pay more attention to the characteristics of the background and target, and use them to detect the target on a single frame. In general, DBT methods have the advantages of fast operation speed, high detection precision and robust to noise. Therefore, it has developed rapidly in recent years.

DBT methods can be divided into three categories: the background suppression (BS)-based methods, the human visual system (HVS)-based methods and the optimization-based methods which are the most representative. The optimization-based methods consider that an infrared image X with a target is the linear superposition of three parts: background image B , target image T and clutter image E [10]. As shown in Eq. (1).

$$X = B + T + E \quad (1)$$

* Corresponding authors.

E-mail addresses: sparkcarleton@gmail.com (T. Zhang), zmpeng@uestc.edu.cn (Z. Peng), haowu_cn@163.com (H. Wu), heyamin@uestc.edu.cn (Y. He), lich@uestc.edu.cn (C. Li), yangchunping@uestc.edu.cn (C. Yang).

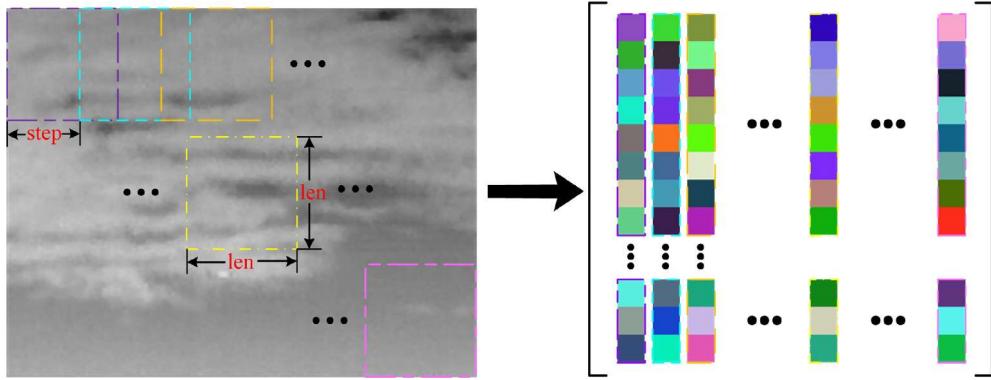


Fig. 1. Illustration of infrared image conversion into patch image.

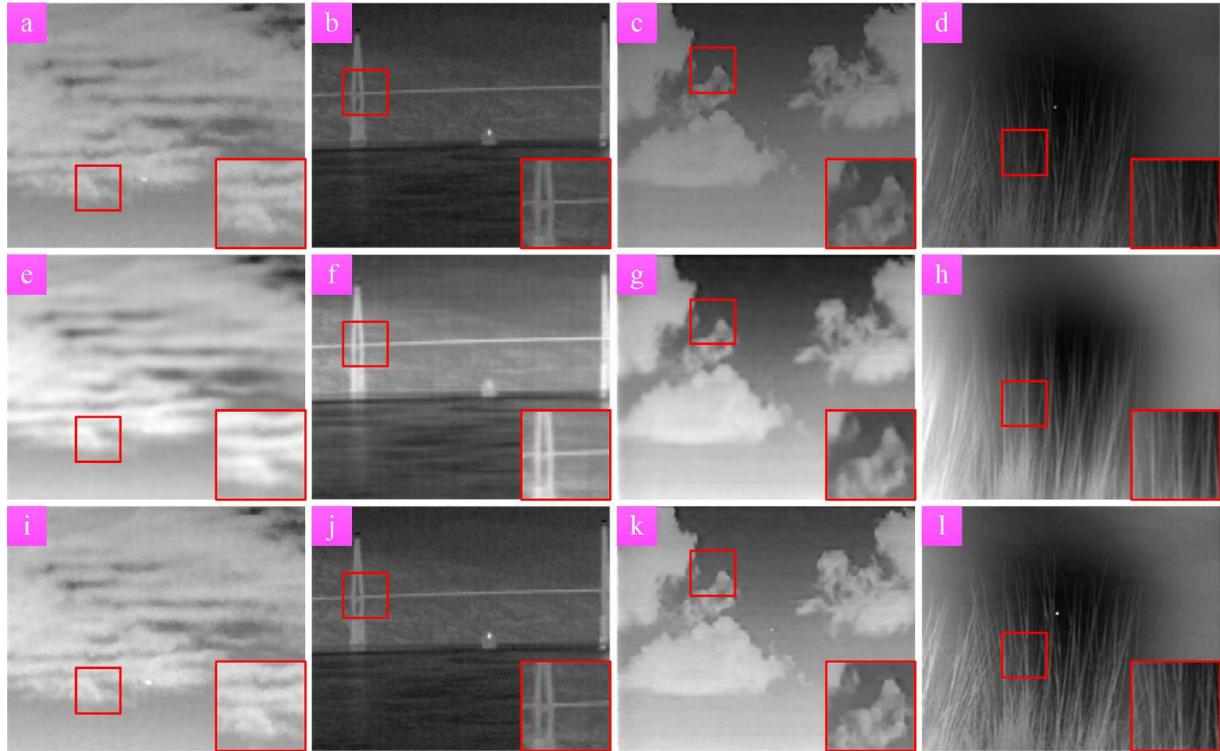


Fig. 2. Illustration of background estimation capability. (a)–(d) are the original images; (e)–(h) are the background estimated by IPI, a single subspace method, and (i)–(l) are the background estimated by the multi-subspaces method SRWS proposed in this paper.

Then, slide the sliding window as shown in Fig. 1 to convert the image into a patch image as shown in Eq. (2). The sliding window slides from the upper left to the lower right. Each sliding window is reshaped into a column vector and arranged backwards.

$$\mathbf{X} = \mathbf{B} + \mathbf{T} + \mathbf{E} \quad (2)$$

The background patch image contains a large number of repeated elements, which can be regarded as a low-rank matrix. In the target patch image, only the target has response, and the rest is all zero, so it can be regarded as a sparse matrix. Therefore, the infrared small target detection problem is transformed into a sparse matrix reconstruction problem. The robust principal component analysis (RPCA) problem [11] shown as Eq. (3) can be solved by applying constraints on the low-rank and sparse items respectively.

$$\min_{\mathbf{B}, \mathbf{T}} \text{rank}(\mathbf{B}) + \|\mathbf{T}\|_0 \quad \text{s.t. } \mathbf{X} = \mathbf{B} + \mathbf{T} \quad (3)$$

where $\text{rank}(\cdot)$ is the rank function of the matrix, and $\|\cdot\|_0$ represents the L0 norm of the matrix, namely the number of non-zero elements in the matrix.

The optimization-based methods are developing very fast, and many excellent methods are proposed. But it is worth noting that these methods implicitly assume that all elements in the background patch image come from the same subspace. However, this assumption is not always accurate for the real data of complex scenes, and the specific description can be seen in Section 2.

In this paper, we present a novel infrared small target detection method called self-regularized weighted sparse (SRWS) model. This method uses overlapping edge information (OEI) to detect the background edge and structure information, and then applies it to constrain the sparse item to enhance the accuracy of the detection algorithm from the perspective of target. Meanwhile, from the perspective of background, we apply the self-regularization item to mine the potential information in

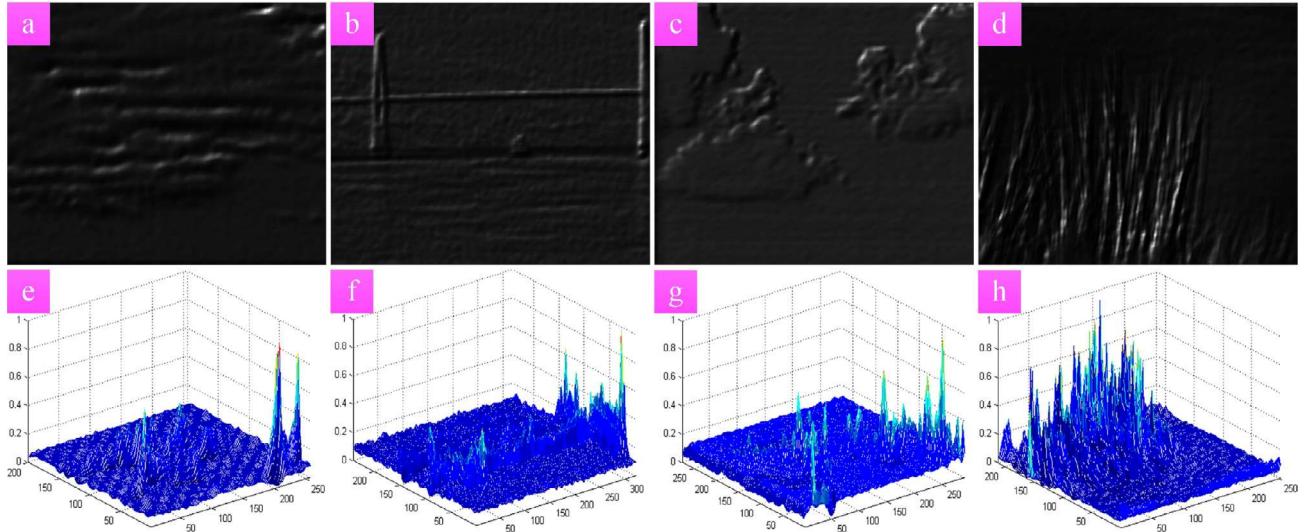


Fig. 3. Illustration of OEI. (a)–(d) are OEI for four different scenes, (e)–(h) are their 3D display.

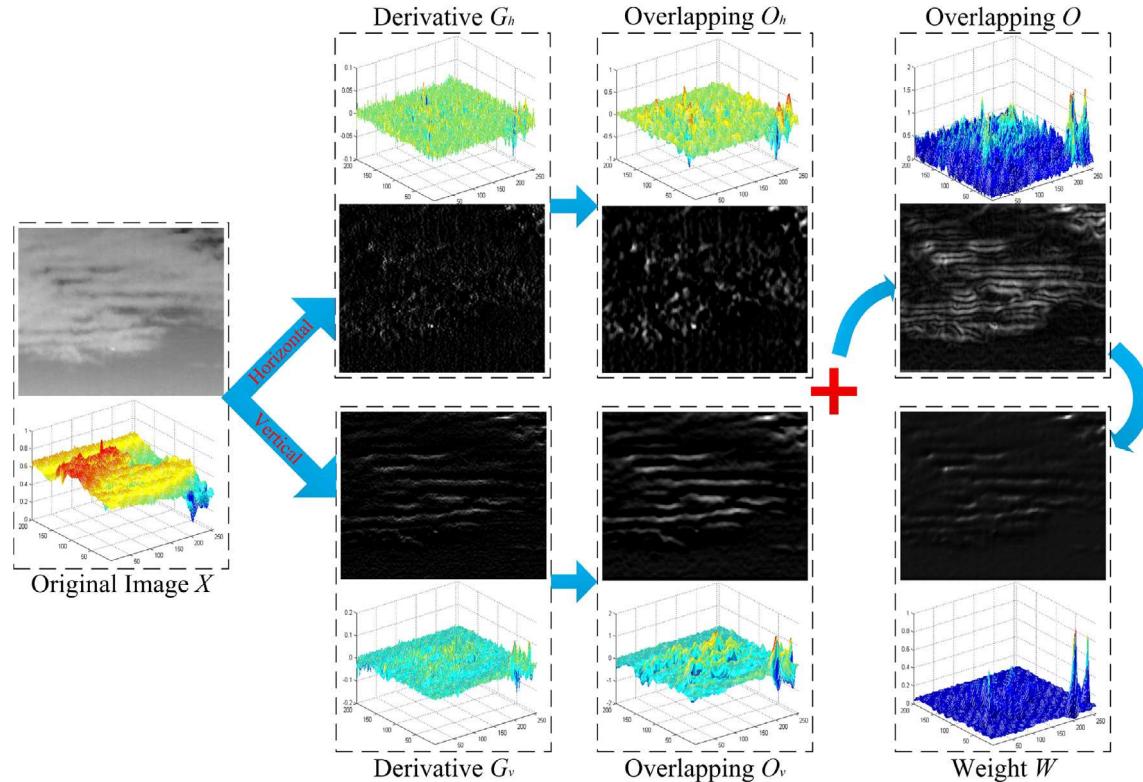


Fig. 4. Flow chart for calculating OEI.

background, and extract clutter from subspaces different from pure background. Therefore, the algorithm has a stronger interpretation ability to random factors, and can also better suppress the background and enhance the robustness.

The main contributions of this paper can be summarized as follows:

- (1) First, we proposed a target detection algorithm named SRWS model based on the reasonable assumption that background elements may come from multi-subspaces. And the infrared

small target detection problem is transformed into a matrix decomposition problem, which can be solved efficiently by an optimization method;

- (2) Secondly, by applying OEI constraint sparse item, the detection accuracy is enhanced from the aspect of target. Meanwhile, the self-regularization item is used to mine the potential information in the background, so as to improve the background suppression ability of the algorithm, and the experiments also proved the superiority of the proposed method;

(3) Thirdly, combining the optimization problem with alternating direction method of multipliers (ADMM), we proposed an efficient solver of SRWS model. And the number of iterations can be reduced by improving the convergence strategy of the algorithm. Furthermore, the solver can be applied to other models in similar application fields.

The remaining of this paper is organized as follows: Section 2 describes research work in related fields; Section 3 describes the methodology applied in this paper as well as the SRWS model and solution method; Section 4 shows the experimental part of this paper; Section 5 discusses the similarities and differences between SRWS and baselines; the conclusion is given in Section 6.

2. Related work

2.1. Infrared small target detection

As mentioned above, according to the principle of the algorithm, infrared small target detection methods can be divided into three categories: BS-based methods, HVS-based methods and optimization-based methods.

BS-based methods assume the continuity of background in the infrared image. They consider that the background will not mutate and design the algorithm to detect targets. Typical methods are Tophat method [12], the two-dimensional adaptive least mean square (TDLMS) algorithm [13] and Max-Median method [14]. Wang [15] applied facet model to infrared small target detection. Guo proposed the phase spectrum of Fourier transform (PFT)

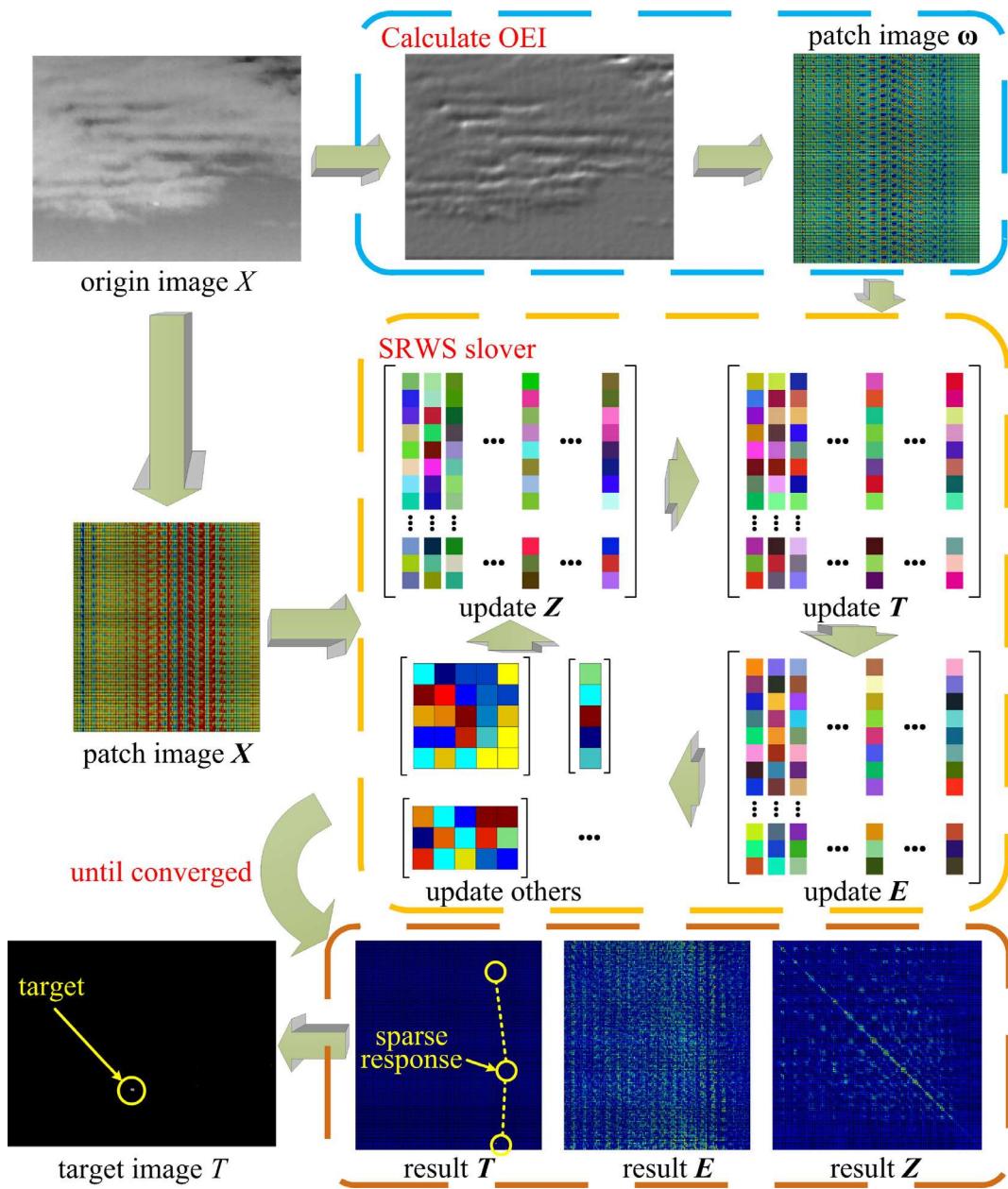


Fig. 5. Detection procedure of SRWS model.

method [16]. Han considers the directivity of clutter, and proposed improved difference of Gabor (IDoGb) [17], Qi applied facet model and simplified it into facet kernel, and finally proposed fast saliency method [18]. Li applied local steering kernel (LSK) [19] to suppress the background and proposed a method called local adaptive contrast measure based on LSK reconstruction (LACM-LSK) [20]. Due to the relatively simple assumption of such methods, the algorithm is not ideal when clutter, noise and other discontinuous backgrounds appear.

In recent years, approaches based on HVS have become popular. HVS believes that the difference between the target and background causes human to directly capture the target's location. Chen applied this idea into the field of infrared small target detection and proposed local contrast method (LCM) [21]. Wang used difference of Gaussian (DoG) [22] filter to detect targets. Then Han [23] proposed improved local contrast method (ILCM) by increasing the sliding-window step size to improve the detection speed. Subsequently, several entropy-based methods have also been proposed [24–26]. Wei proposed multiscale patch-based contrast measure (MPCM) [27], which simplified the operation steps of the algorithm and greatly improved the computational efficiency. Shi proposed a novel method called high-boost-based multiscale local contrast measure (HBMLCM) [28,29]. This method starts from the aspect of target and emphasizes the difference between the target and background, but it is difficult to deal with the stripe target, and the running time of multi-scale filtering operation cannot be guaranteed.

Optimization-based methods consider the image is a linear superposition of the target, background and noise, and attempts to solve the target image iteratively according to related theory of optimization such as sparse coding and RPCA problem. In this way, the detection problem is transformed into an optimization problem. Gao uses this idea and proposed infrared patch-image (IPI) model [10]. Subsequently, optimization-based methods developed vigorously. Dai weighted each column in the patch image to obtain a better background suppression effect and proposed the weighted infrared patch-image (WIPI) model [30]. Then he proposed a method called non-negative infrared patch-image model based on partial sum minimization of singular values (NIPPS) [31,32]. The total variation regularization and principal component pursuit (TV-PCP) method [33] proposed by Wang used total variation constraint optimization formula and effectively suppress the edge information in the image. Zhang used L21-norm to constrain clutter and proposed non-convex rank approximation method

(NRAM) [34]. Zhang introduced Lp-norm, better constrained sparse target, and proposed non-convex optimization with Lp-norm constraint (NOLC) [35]. In addition, Dai used the concept of tensor and combines it with local structural tensor [36] to propose the reweighted infrared patch-tensor (RIPT) model [37], and introduced infrared small target detection into the field of tensor. Subsequently, a tensor-based nonconvex low-rank constraint named partial sum of tensor nuclear norm (PSTNN) method [38] is proposed. In contrast, the assumption of such methods is closest to reality, and their detection effects achieve the state-of-the-art performance with the improvement of models and optimization methods.

2.2. Subspace clustering

For a long time, subspace clustering has been paid attention by many researchers. The purpose is to use the sparsity of data in a specific space to represent more effectively or reveal the essential characteristics of data. Elhamifar [39] applied compressed sensing and put forward sparse subspace clustering (SSC) by using L0-norm constraint affine matrix. Subsequently, low-rank representation (LRR) method [40] was proposed, which used rank function constraint affine matrix and nuclear norm approximation to obtain better results in corrupted data. Liu [41] proposed latent low-rank representation (LatLRR) by using hidden data to construct the dictionary. Zhuang [42] used both L0-norm and nuclear norm for constraining and presented non-negative low-rank and sparse representation (NNLRSR). Lu [43] proposed graph-regularized low-rank representation (GLRR) and used it to destriping of hyperspectral images. Subsequently, other improvement methods [44–46] of LRR were proposed.

$$\min_{Z, P, Q} \|Z - PQ\|_F^2 \quad s.t. X = DZ, \vec{1}_n^T Z = \vec{1}_n^T \quad (4)$$

Another effective improved form of LRR is fixed-rank representation (FRR) [47], which can obtain more realistic subspace composition under insufficient observation. The objective function of FRR is given by Eq. (4), where Z is the coefficient matrix of over-complete dictionary D , $\vec{1}_n^T$ is a vector of length n with all values are 1, P and Q are the LR decomposition of \tilde{Z} which means $\tilde{Z} = PQ$. The purpose of FRR is to make Z and \tilde{Z} as close as possible to find the sparsest representation of Z . Li added sparse constraint to FRR and proposed sparse FRR method [48]. Subsequently, Wei

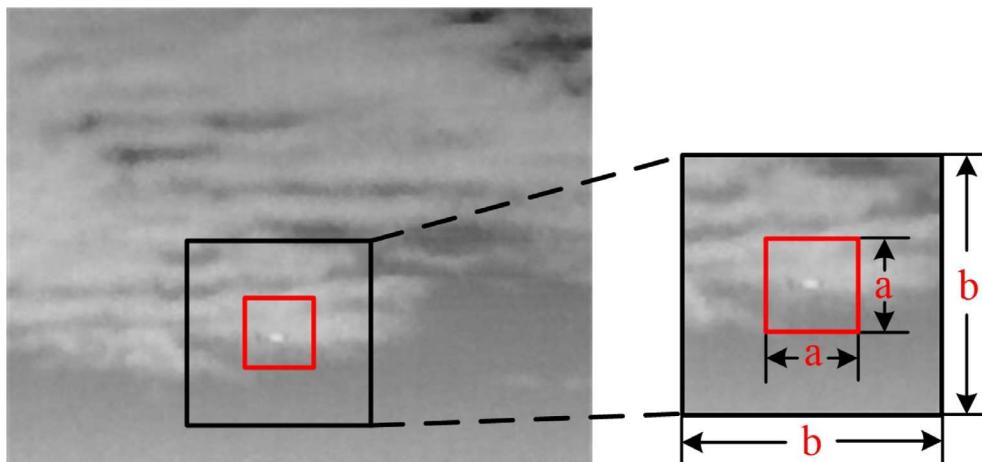


Fig. 6. Illustration of background and target region.

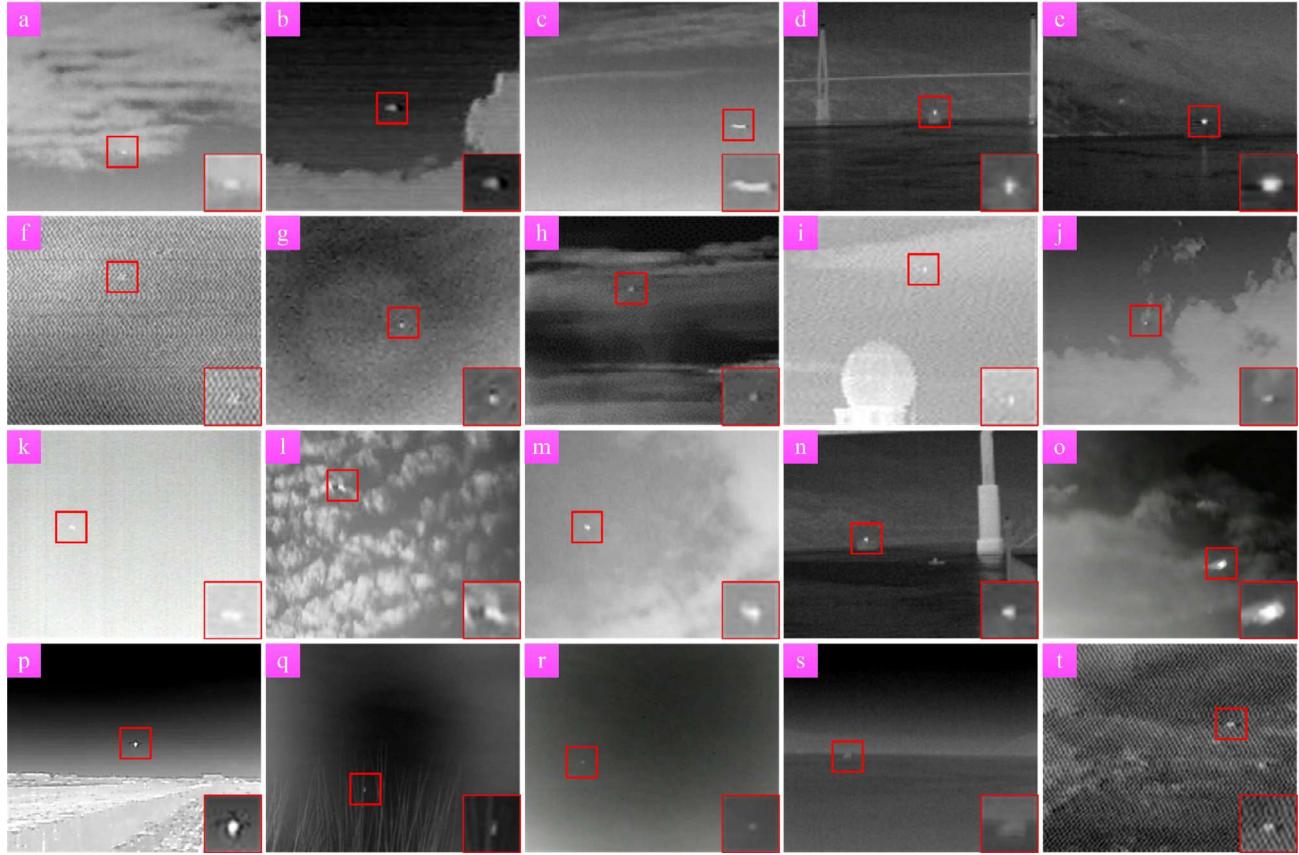


Fig. 7. Illustrations of 20 complex scenes.

[49] constructed the Laplace regularizer using affine matrix to constrain FRR and proposed self-regularized fixed rank representation (SRFRR). The effect of this method is obviously better than other algorithms [49]. Our model also benefits from these advances.

2.3. Subspace clustering for target detection

Recently, there have been a few studies of using subspace clustering for infrared small targets detection. By using LRR [50,51] and constructing over-complete dictionary, He proposed a low-rank and sparse representation (LRSR) model [52]. Then, Wang proposed a multi-subspaces learning (SMSL) method [53] by analyzing the structure of the background data.

In terms of algorithm optimization, when we use self-expression method, which means use original data as the over complete dictionary, if the coefficient is identity matrix, multi-subspaces methods are the same as that of single subspace. So, we can say that the latter is a special case of the former. Furthermore, the multi-subspaces methods are more flexible in handling complex data.

In terms of background estimation capacity, single subspace methods like IPI [10] can well deal with the homogeneous background, while it is difficult to estimate the heterogeneous background with large interference such as background clutter and noise, which limits the application of single subspace methods. In contrast, thanks to the assumption that data may come from multi-subspaces, the multi-subspaces methods are good estimators of both homogeneous and heterogeneous background. In Fig. 2, we can clearly see that the background estimated by a single subspace method called IPI is very fuzzy, while the

multi-subspaces method called SRWS proposed in this paper is little different from the original image.

3. Methodology

In this section, we first describe the origin and construction of OEI. Then, the objective function and optimization method of SRWS are given. In this paper, for convenience of expression, we summarize the notations and norm definitions used in Table 1.

3.1. Overlapping edge information

All this time, the total variation (TV) model [54] has been widely concerned. However, the TV model will produce unexpected staircase effect in the application process [55–58]. Therefore, in recent years, the concept of overlapping group sparse

Table 1
Notations and norm definitions.

Name	Definition
Scalar	a
Vector	\vec{a}
Original Image	A
Patch Image	\mathbf{A}
Matrix Element	$A(i,j)$
A Column In Matrix	$A(:,j)$
Vector 2 Norm	$\ \vec{a}\ _2$
L1 Norm	$\ A\ _1 = \sum_i \sum_j A(i,j) $
Frobenius Norm	$\ A\ _F = \sqrt{\sum_i \sum_j A(i,j)^2}$
L21 Norm	$\ A\ _{2,1} = \sum_i \sqrt{\sum_j A(i,j)^2}$

Table 2

Description of sequential image.

Sequence	Image Size	Number	Description of Target	Description of Background
Seq. 1	200 × 256	30	Size: 4 × 6 Move on the bottom	Structural cloud Target passed through cloud
Seq. 2	200 × 256	40	Size: 4 × 10 Move on the right	Long strip cloud Low signal-to-noise ratio
Seq. 3	240 × 320	409	Size: 6 × 9 Large movement	Structural cloud Background changes quickly
Seq. 4	240 × 320	67	Size: 6 × 8 Fast motion	Low signal-to-noise ratio Longitudinal detector stripes
Seq. 5	172 × 256	400	Size: 3 × 2 – 9 × 10 Size varies greatly	Massive clouds Clouds change quickly
Seq. 6	213 × 252	185	Size: 2 × 3 Brightness changes	Grass interference Detector fade

(OGS) has received extensive attention [59,60]. Using OGS to suppress the staircase effect also got very good effect. The group sparse optimization solution using ADMM [61] is provided first [62]. Selesnick then used majorization-minimization (MM) algorithm [63] to use OGS for image denoising [64]. Soon after, OGS was used in image reconstruction, improved denoising frame and hyperspectral image (HSI) denoising [65].

Inspired by this, we use OEI to characterize the background features of infrared images. As shown in Fig. 3, the edge part of OEI is relatively large, and we can use this feature to constraint the sparse item and highlight the target in heterogeneous background. In order to obtain the OEI of image X, we calculate the overlap matrix of the horizontal and vertical derivatives, and then combine them to get the matrix O of the following equation.

$$O(i,j) = |O_h(i,j)| + |O_v(i,j)| \quad (5)$$

where

$O_h(i,j) = \sum_{i=-m_1}^{m_2} \sum_{j=-m_1}^{m_2} G_h(i,j)$, $O_v(i,j) = \sum_{i=-m_1}^{m_2} \sum_{j=-m_1}^{m_2} G_v(i,j)$, G_h and G_v represent the first derivatives of the original image X in the horizontal and vertical directions respectively; in addition, $m_1 = \lceil \frac{-1}{2} \rceil$, $m_2 = \lfloor \frac{1}{2} \rfloor$, the operator $\lceil a \rceil$ represents the largest integer

not greater than the number a , l is the number of overlapping groups. Then convert the matrix O to patch image \mathbf{O} . Finally, we define the OEI of the original image X as the following.

$$\omega = \exp \left(\alpha * \frac{\mathbf{O} - \mathbf{O}_{\min}}{\mathbf{O}_{\max} - \mathbf{O}_{\min}} \right) \quad (6)$$

where, \mathbf{O}_{\max} and \mathbf{O}_{\min} are respectively the maximum and minimum value of matrix \mathbf{O} , α is the stretching coefficient, and $\exp(\cdot)$ is the exponential. The specific steps for calculating OEI are given in Fig. 4.

Therefore, by combining the linear hypothesis in Eq. (2) and the definition of FRR in Eq. (4), we can obtain a detection model of infrared small targets as follows. λ and γ are penalty factors and the symbol \bullet represents Hadamard product.

$$\begin{aligned} & \min_{Z, P, Q, E, T} \|Z - PQ\|_F^2 + \lambda \|E\|_{2,1} + \gamma \|\omega \bullet T\|_1 \text{ s.t. } X \\ &= DZ + T + E, \vec{1}_n^T Z = \vec{1}_n^T \end{aligned} \quad (7)$$

3.2. Self-regularized weighted sparse model

Considering that FRR cannot suppress discrete points well, in order to mine the potential information in the background and

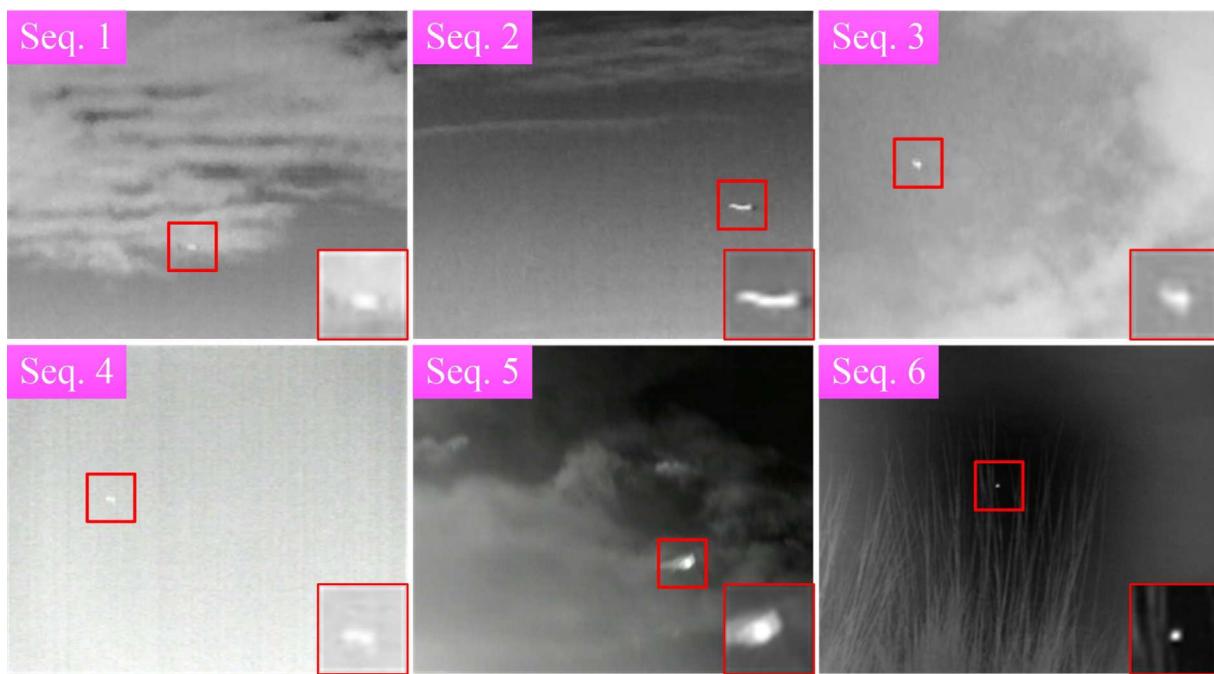


Fig. 8. Illustrations of six infrared image sequences.

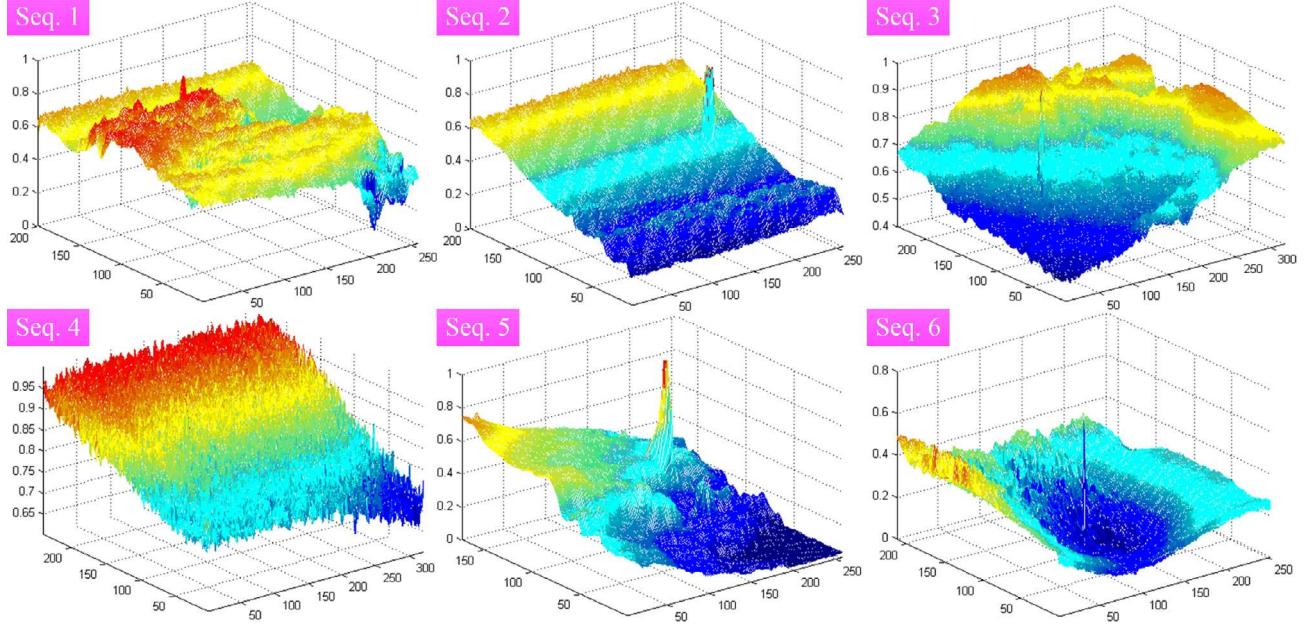


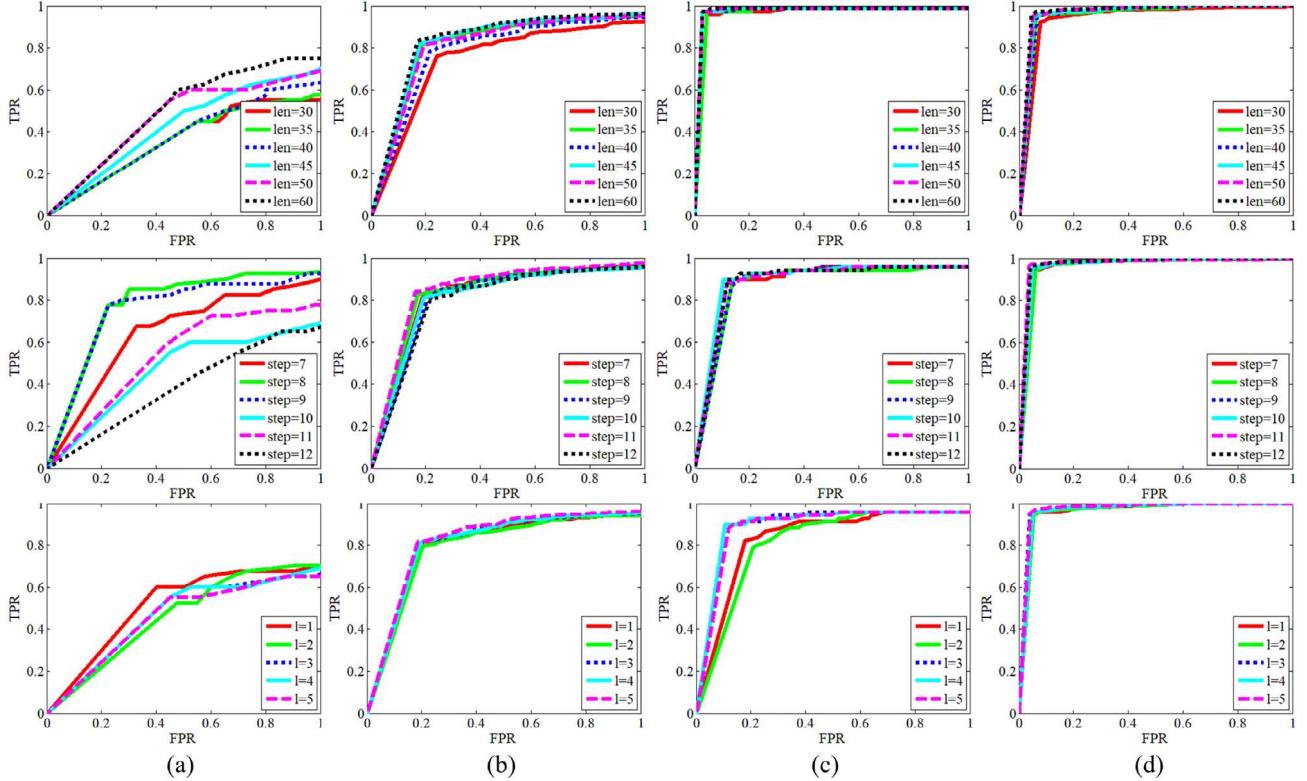
Fig. 9. 3D display of six infrared image sequences.

extract noise and clutter more accurately, we apply the self-regularization item [49] $R(\mathbf{Q}, \mathbf{Z})$ which is formed by the Laplace matrix transformation.

We first define the matrix $\mathbf{W} = (\mathbf{Z} + \mathbf{Z}^T)/2$ for a given matrix \mathbf{Z} . Then calculate matrix \mathbf{V} according to $\mathbf{V}(i, i) = \sum_j \mathbf{W}(i, j)$. And then we can get the Laplace matrix of \mathbf{Z} as follows $\mathbf{L} = \mathbf{V} - \mathbf{W}$. Thereafter we define the self-regularization term $R(\mathbf{Q}, \mathbf{Z})$ in the following Equation.

$$R(\mathbf{Q}, \mathbf{Z}) = \text{tr}(\mathbf{Q} \mathbf{L} \mathbf{Q}^T) = \frac{1}{2} \sum_{i,j} \|\mathbf{Q}(:, i) - \mathbf{Q}(:, j)\|_2^2 \bullet \mathbf{W} = \frac{1}{2} \|\mathbf{M} \bullet \mathbf{Z}\|_1 \quad (8)$$

where $\text{tr}(A) = \sum_i A(i, i)$ represents the trace of the matrix, and for matrix \mathbf{M} , $\mathbf{M}(i, j) = \|\mathbf{Q}(:, i) - \mathbf{Q}(:, j)\|_2^2$. Since $\frac{1}{2} \|\mathbf{M} \bullet \mathbf{Z}\|_1$ is a function only related to the matrix \mathbf{Q} and \mathbf{Z} , so R is defined as $R(\mathbf{Q}, \mathbf{Z})$.

Fig. 10. The ROC of parameters len , $step$ and l were compared.

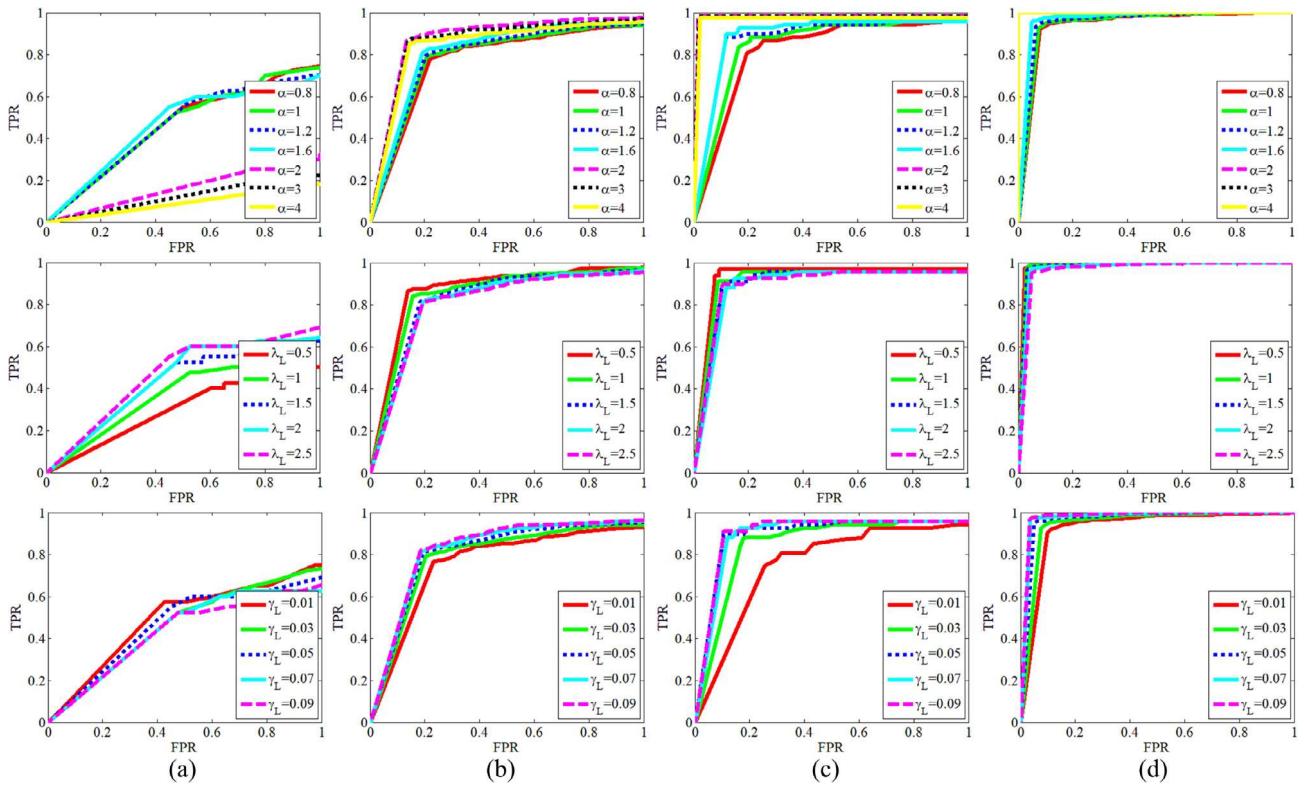


Fig. 11. The ROC of parameters α , λ_L and γ_L were compared.

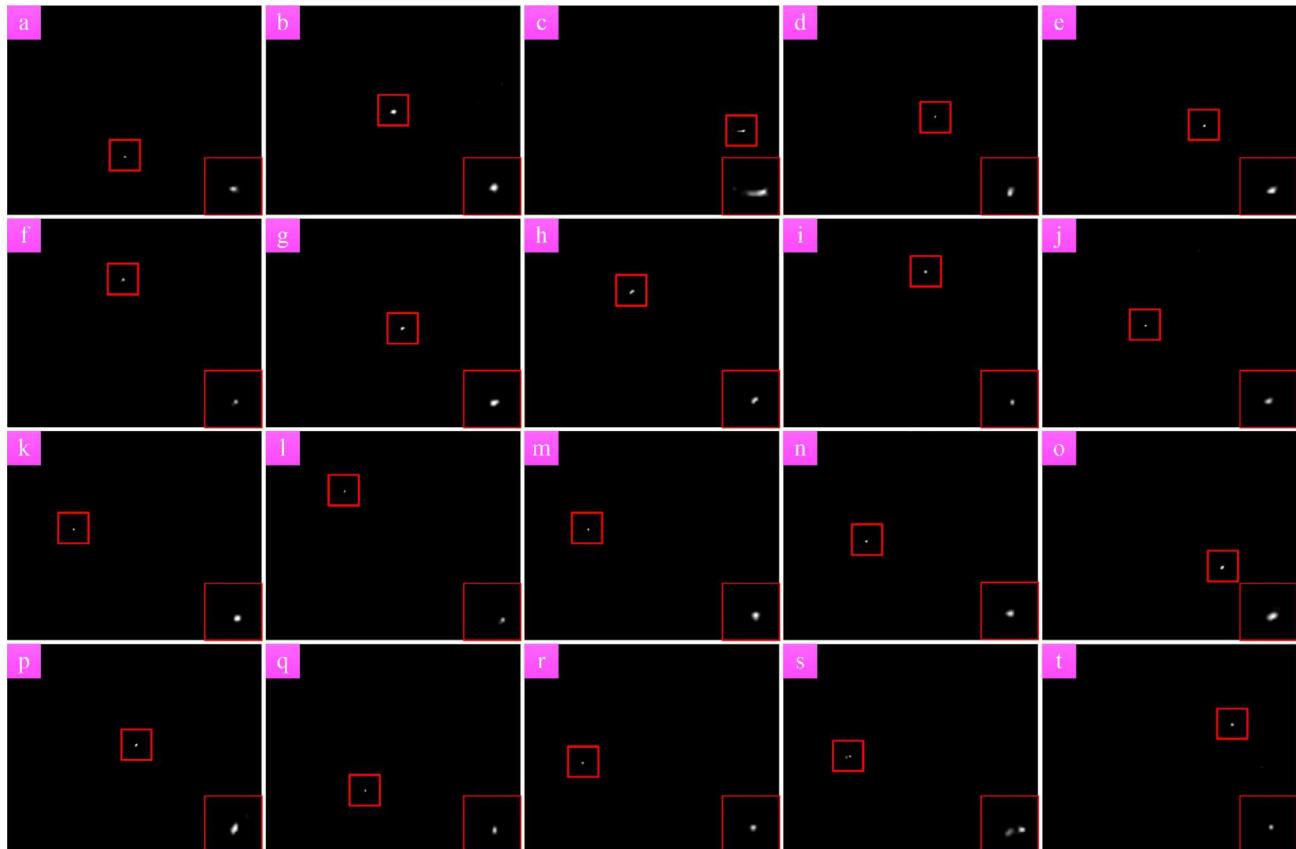


Fig. 12. Illustration of SRWS model detecting complex background images.

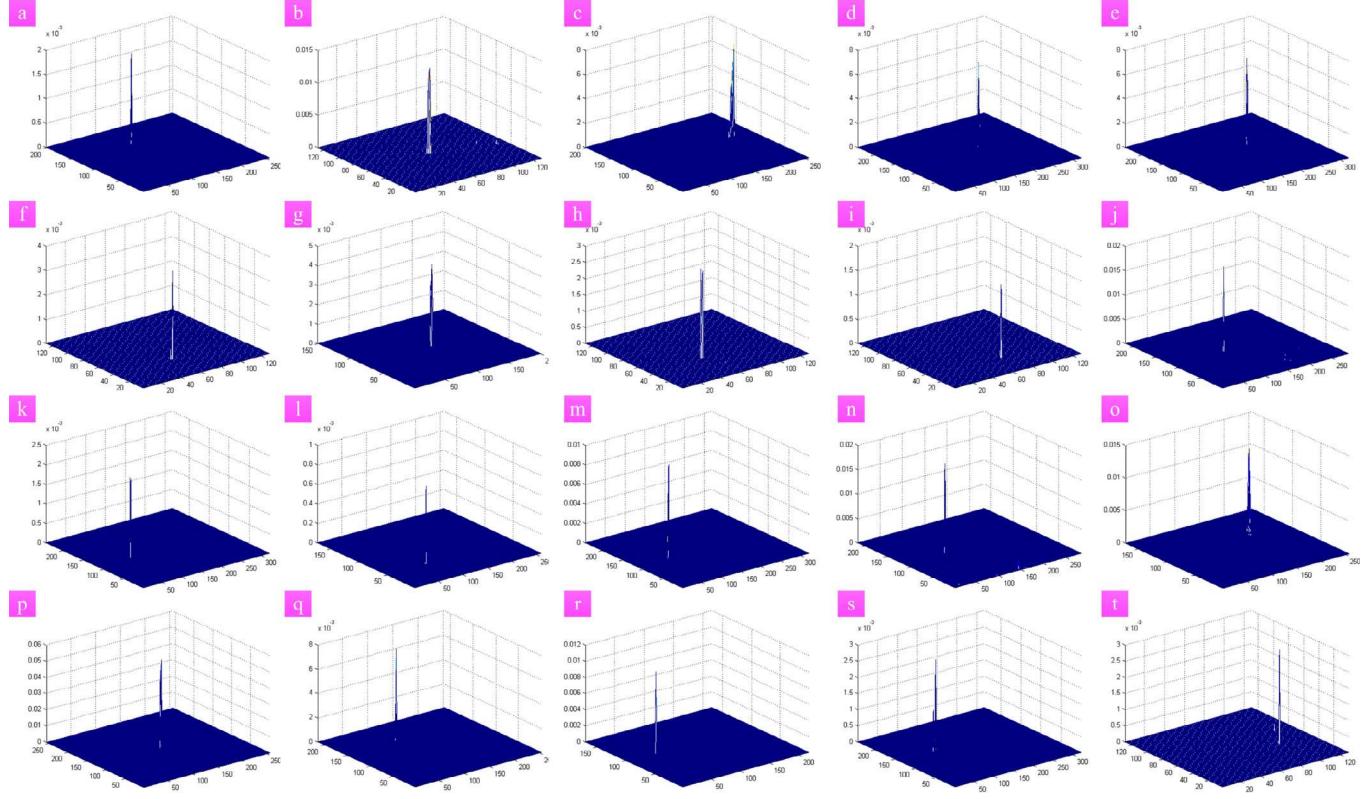


Fig. 13. 3D display of SRWS model detecting complex background images.

Table 3
Parameter setting of baselines.

Abbreviation	Parameter Setting
Tophat	Structure shape: disk, Structure size: 5×5
FastSaliency	Filter size: 5×5 Standard deviations: 1.5
DoG	Filter size: 5×5
MPCM	Patch radius: 1, 2, 3, 4
HBMCLM	Inner window radius = 1, 2, 3, 4 Outer window radius = 7
IPI	Patch size: 50×50 , Patch step: 10, $L = 1$, $\lambda = L/\sqrt{\min(m, n)}$
NRAM	Patch size: 50×50 , Patch step: 10, $L = 1$, $\lambda = L/\sqrt{\min(m, n)}$
PSTNN	Patch size: 40×40 , Patch step: 40, $\lambda_L = 0.7$, $\lambda = \lambda_L/\sqrt{\min(m, n)}$
NOLC	Patch size: 40×40 , Patch step: 10, $\lambda = 1/\sqrt{\min(m, n)}$, $p = 0.5$
LRSR	Patch size: 50×50 , Patch step: 10, $\beta = 1/\sqrt{\min(m, n)}$, $\lambda = 1/\sqrt{\min(m, n)}$
SMSL	Patch size: 30×30 , Patch step: 30, $\lambda = 2/\sqrt{\min(m, n)}$
SRWS	Patch size: 50×50 , Patch step: 10, $\beta = 1/\sqrt{\min(m, n)}$, $\lambda = \lambda_L/\sqrt{\min(m, n)}$, $\gamma = \gamma_L/\sqrt{\min(m, n)}$

In practical terms, we know that nearby data samples should have similar representations. The use of $R(\mathbf{Q}, \mathbf{Z})$ is a good way to preserve the local geometry structure between the data [49,66]. In addition, $R(\mathbf{Q}, \mathbf{Z})$ can represent data without the additional dependency of an over complete dictionary, so it is called self-regularization item. So we can get the objective function of SRWS model in this paper as follow, where β is penalty factor:

$$\begin{aligned} & \min_{\mathbf{Z}, \mathbf{P}, \mathbf{Q}, \mathbf{E}, \mathbf{T}} \|\mathbf{Z} - \mathbf{PQ}\|_F^2 + \beta R(\mathbf{Q}, \mathbf{Z}) + \lambda \|\mathbf{E}\|_{2,1} + \gamma \|\mathbf{w} \bullet \mathbf{T}\|_1 \text{ s.t. } \mathbf{X} \\ &= \mathbf{DZ} + \mathbf{E}, \bar{\mathbf{1}}_n^T \mathbf{Z} = \bar{\mathbf{1}}_n^T \end{aligned} \quad (9)$$

3.3. Optimization

In this section, we will introduce the solution method of Eq. (9) in detail. As for the construction of dictionary matrix \mathbf{D} , we adopt the self-representation method, that is, we use the original data as the dictionary. This approximation is widely used in subspace clustering [39] and can get good results. Therefore, the optimization objective is updated as follows:

$$\begin{aligned} & \min_{\mathbf{Z}, \mathbf{P}, \mathbf{Q}, \mathbf{E}, \mathbf{T}} \|\mathbf{Z} - \mathbf{PQ}\|_F^2 + \beta R(\mathbf{Q}, \mathbf{Z}) + \lambda \|\mathbf{E}\|_{2,1} + \gamma \|\mathbf{w} \bullet \mathbf{T}\|_1 \text{ s.t. } \mathbf{X} \\ &= \mathbf{XZ} + \mathbf{E}, \bar{\mathbf{1}}_n^T \mathbf{Z} = \bar{\mathbf{1}}_n^T, \mathbf{Z} = \mathbf{J} \end{aligned} \quad (10)$$

We then use ADMM to construct the following Lagrange function for this objective function. And we update each parameter to optimize the Lagrange function while holding the other parameters fixed.

$$\begin{aligned} L(\mathbf{Z}, \mathbf{E}, \mathbf{T}) = & \|\mathbf{Z} - \mathbf{PQ}\|_F^2 + \beta R(\mathbf{Q}, \mathbf{Z}) + \lambda \|\mathbf{E}\|_{2,1} + \gamma \|\mathbf{w} \bullet \mathbf{T}\|_1 \\ & + \langle Y_1, \mathbf{X} - \mathbf{XZ} - \mathbf{T} - \mathbf{E} \rangle + \langle Y_2, \bar{\mathbf{1}}_n^T \mathbf{Z} - \bar{\mathbf{1}}_n^T \rangle \\ & + \langle Y_3, \mathbf{Z} - \mathbf{J} \rangle + \frac{\mu}{2} \|\mathbf{X} - \mathbf{XZ} - \mathbf{T} - \mathbf{E}\|_F^2 \\ & + \frac{\mu}{2} \|\bar{\mathbf{1}}_n^T \mathbf{Z} - \bar{\mathbf{1}}_n^T\|_F^2 + \frac{\mu}{2} \|\mathbf{Z} - \mathbf{J}\|_F^2 \end{aligned} \quad (11)$$

where, $\langle \cdot, \cdot \rangle$ represents the inner product of matrix; Y_1 , Y_2 and Y_3 are Lagrange multipliers; μ represents a non-negative penalty factor. We can get the target patch image by iteratively optimizing the above Lagrange function. In this process, we need to solve three key problems. We will explain their solution method separately.

a. *The First Problem* Given $\mu > 0$, matrix $A, B \in R^{m \times n}$ and the following optimization function.

$$B^* = \arg \min_B \frac{1}{2} \|B\|_1 + \frac{1}{2} \|B - A\|_F^2 \quad (12)$$

This convex optimization problem can be solved by soft thresholding operator [67]. The solution of this function is defined as follows:

$$\mathcal{S}_\varepsilon[A] = \begin{cases} A - \varepsilon & \text{if } A > \varepsilon \\ A + \varepsilon & \text{if } A < -\varepsilon \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

b. *The Second Problem* Let $A = [a_1, a_2, \dots, a_n]$ be a given matrix, $\|\cdot\|_F$ be the Frobenius norm and $\|\cdot\|_{2,1}$ be the L21-norm. For the following optimization function:

$$B^* = \arg \min_B \lambda \|B\|_{2,1} + \frac{1}{2} \|B - A\|_F^2 \quad (14)$$

Then we define an operator $\mathcal{T}_\lambda[A] = B^*$. The i th column of B^* is

$$B^*(:, i) = \begin{cases} \frac{\|A(:, i)\|_2 - \lambda}{\|A(:, i)\|_2} A(:, i) & \text{if } \lambda < \|A(:, i)\|_2 \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

c. *The Third Problem* For a Sylvester equation [68] like the follow, we can use the function `lyap()` to solve in MATLAB. So the optimal solution is $X^* = \text{lyap}(A, B, C)$.

$$AX + XB + C = 0 \quad (16)$$

According to the three key problems mentioned above, we obtain the optimal solution for each updating equation by fixing the other parameters. The specific steps of the optimization function in an iterative process is shown in **Algorithm 1**.

Algorithm 1 Calculation in One Iteration.

Note: When we update the parameters, hold the others fixed.

1: **Update Q^k :**

$$Q^{k+1} = \text{lyap}\left(2(P^k)^T P^k, \beta(L + L^T), -2(P^k)^T Z^k\right)$$

2: **Update P^k :** $P^{k+1} = Z^k(Q^{k+1})^{-1}$

3: **Update J^k :** $J^{k+1} = \mathcal{S}_{\frac{\mu k}{2\mu k}}[Z^k + Y_3^k/\mu^k]$

4: **Update Z^k :** $Z^{k+1} = H^{-1}U$, Here

$$H = 2I + \mu^k(X^T X + \vec{I}_n^T \vec{I}_n + I)$$

$$U = 2P^{k+1}Q^{k+1} + \mu^k X^T (X - T^k - E^k + Y_1^k/\mu^k) \\ + \mu^k \vec{I}_n^T (\vec{I}_n - Y_2^k/\mu^k) + \mu^k J^{k+1} - Y_3^k$$

5: **Update E^k :** $E^{k+1} = \mathcal{T}_{\lambda/\mu^k}[X - XZ^{k+1} - T^k + Y_1^k/\mu^k]$

6: **Update T^k :** $T^{k+1} = \mathcal{S}_{\frac{\mu k}{\mu^k}}[X - XZ^{k+1} - E^{k+1} + Y_1^k/\mu^k]$

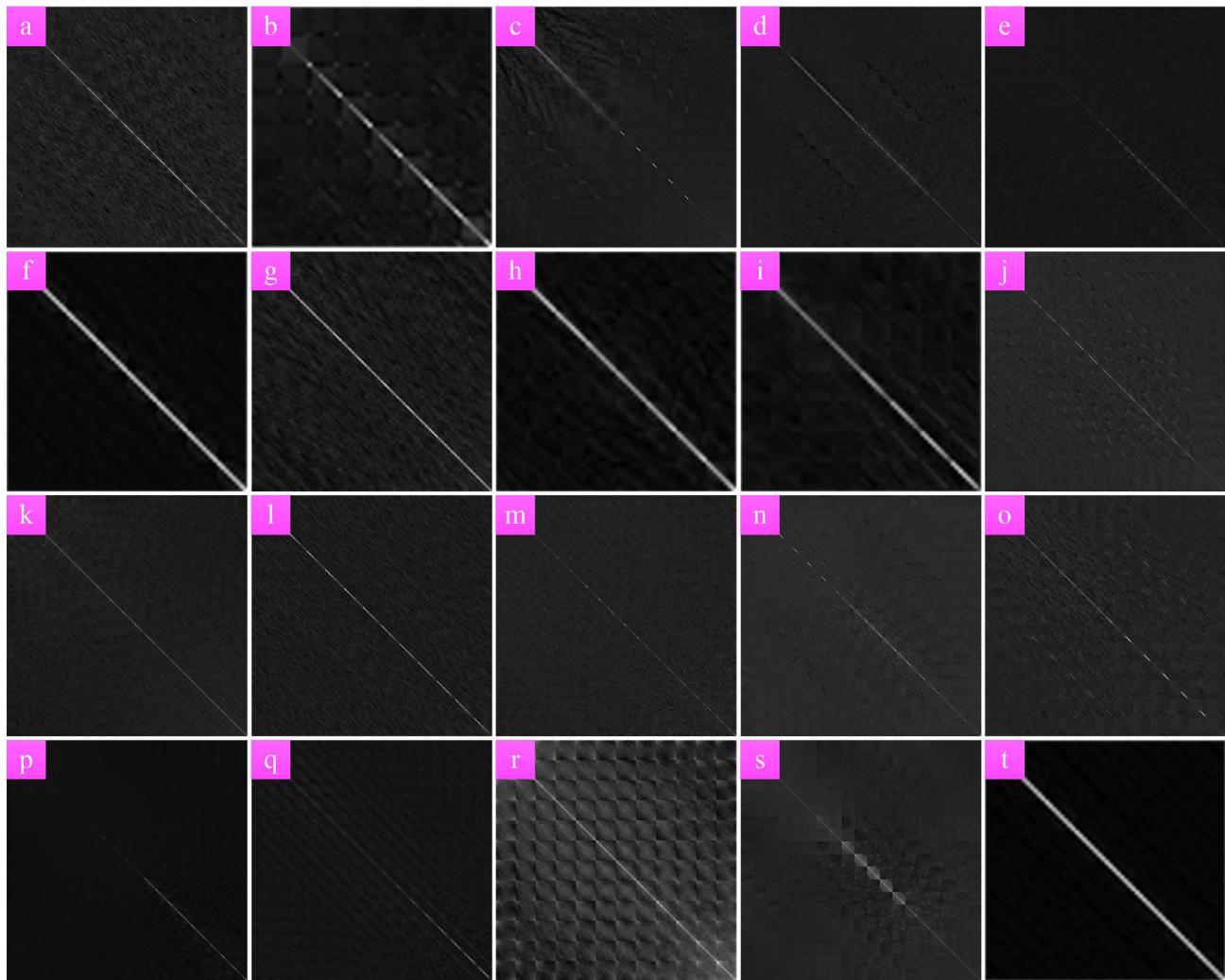


Fig. 14. Illustration of SRWS coefficient matrix Z in complex background images.

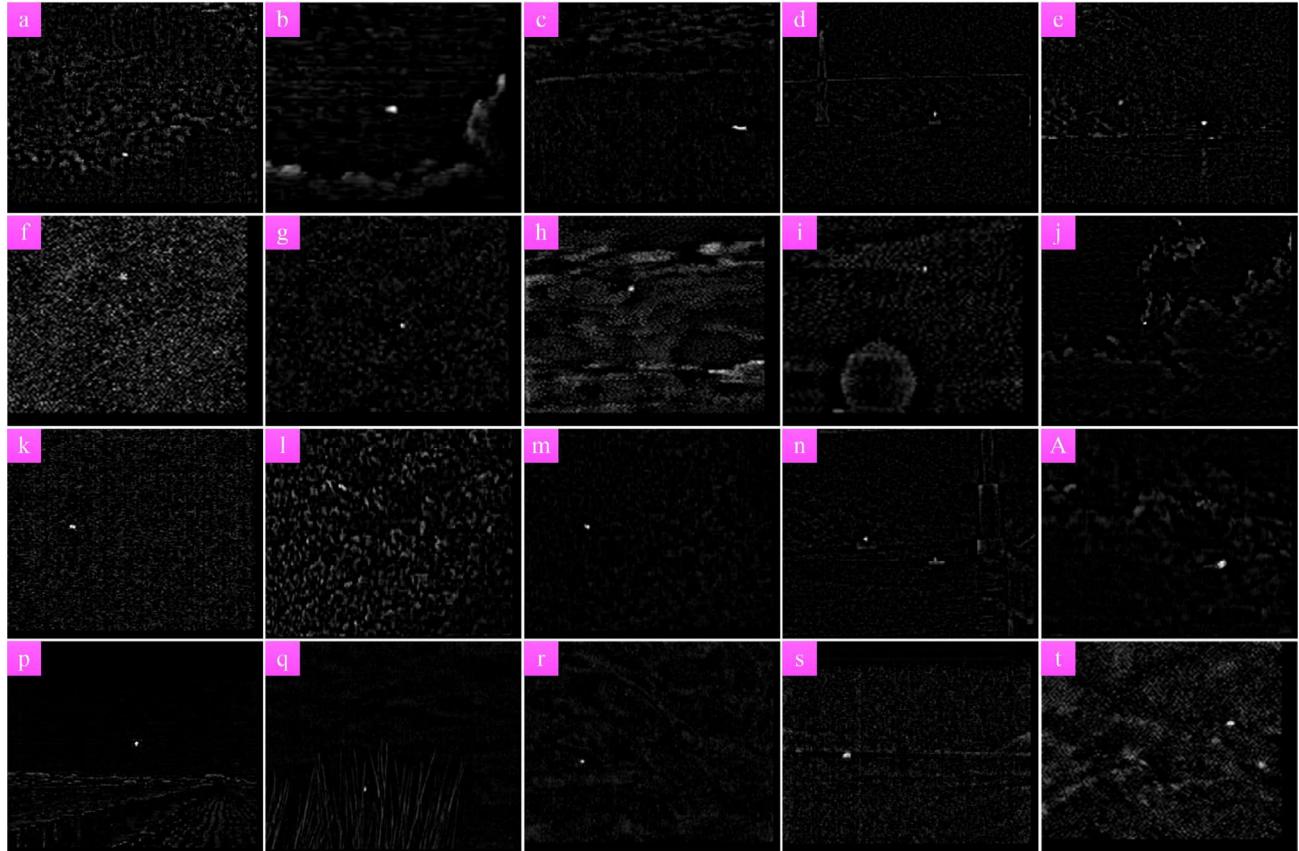


Fig. 15. Illustration of SRWS noise matrix E in complex background images.

7: Update others:

$$Y_1^{k+1} = Y_1^k + \mu^k (\mathbf{X} - \mathbf{XZ}^{k+1} - \mathbf{T}^{k+1} - \mathbf{E}^{k+1})$$

$$Y_2^{k+1} = Y_2^k + \mu^k (\tilde{\mathbf{I}}_n^T \mathbf{Z} - \tilde{\mathbf{I}}_n^T)$$

$$Y_3^{k+1} = Y_3^k + \mu^k (\mathbf{Z}^{k+1} - \mathbf{J}^{k+1})$$

$$\mu^{k+1} = 1.5\mu^k$$

8: Update \mathbf{k} : $k = k + 1$.

The specific detection steps of the SRWS model are shown in **Algorithm 2**. And Fig. 5 shows the SRWS model detection procedure. The values of specific parameters are given in the experiment section.

Algorithm 2 Iterative method of SRWS.

Note: When we update the parameters, hold the others fixed.

- 1: **Input:** $\mathbf{X}, \omega, \beta, \lambda, \gamma$;
- 2: **Initialization:** $\mu^0 = 1/(5 * \text{std}(\mathbf{X}))$ and others are set to 0;
- 3: **While** not converged;
- 4: %Update parameters: According to **Algorithm 1**;
- 5: %Check the convergence condition
- 6: $\frac{\|\mathbf{X} - \mathbf{XZ} - \mathbf{Z} - \mathbf{T} - \mathbf{E}\|_F}{\|\mathbf{X}\|_F} < 10^{-7}$ or $\left\| \|\mathbf{T}^{k+1}\|_0 - \|\mathbf{T}^k\|_0 \right\| \leq 1$
- 7: **end**
- 8: **Output:** $\mathbf{Z} = \mathbf{Z}^{k+1}, \mathbf{E} = \mathbf{E}^{k+1}, \mathbf{T} = \mathbf{T}^{k+1}$.

About the convergence condition, we found that at the later stage of the optimization process, the number of non-zero elements in the target patch image increased slowly, but had little impact on the detection results. Therefore, the number of non-zero elements in two iterations is not more than 1, which is also taken as one of the criterion.

4. Experiments and analysis

In this section, we introduce the experimental part of this paper. Firstly, the quantitative evaluation metrics is introduced, and then the experiments setup is explained, including test data and comparison methods. After that, the effectiveness of SRWS model was demonstrated qualitatively, and parameters in SRWS were determined through experimental methods. Finally, SRWS was compared quantitatively with baselines.

4.1. Evaluation metrics

In the field of infrared small targets detection, background suppression factor (BSF), signal-to-clutter ratio gain (SCRG) and receiver operating characteristic (ROC) are usually used as quantitative evaluation metrics. In addition, the computational efficiency of the optimization-based methods is particularly concerned because it needs to be solved by iterative methods. Therefore, running time and iteration times are also used as quantitative evaluation metrics.

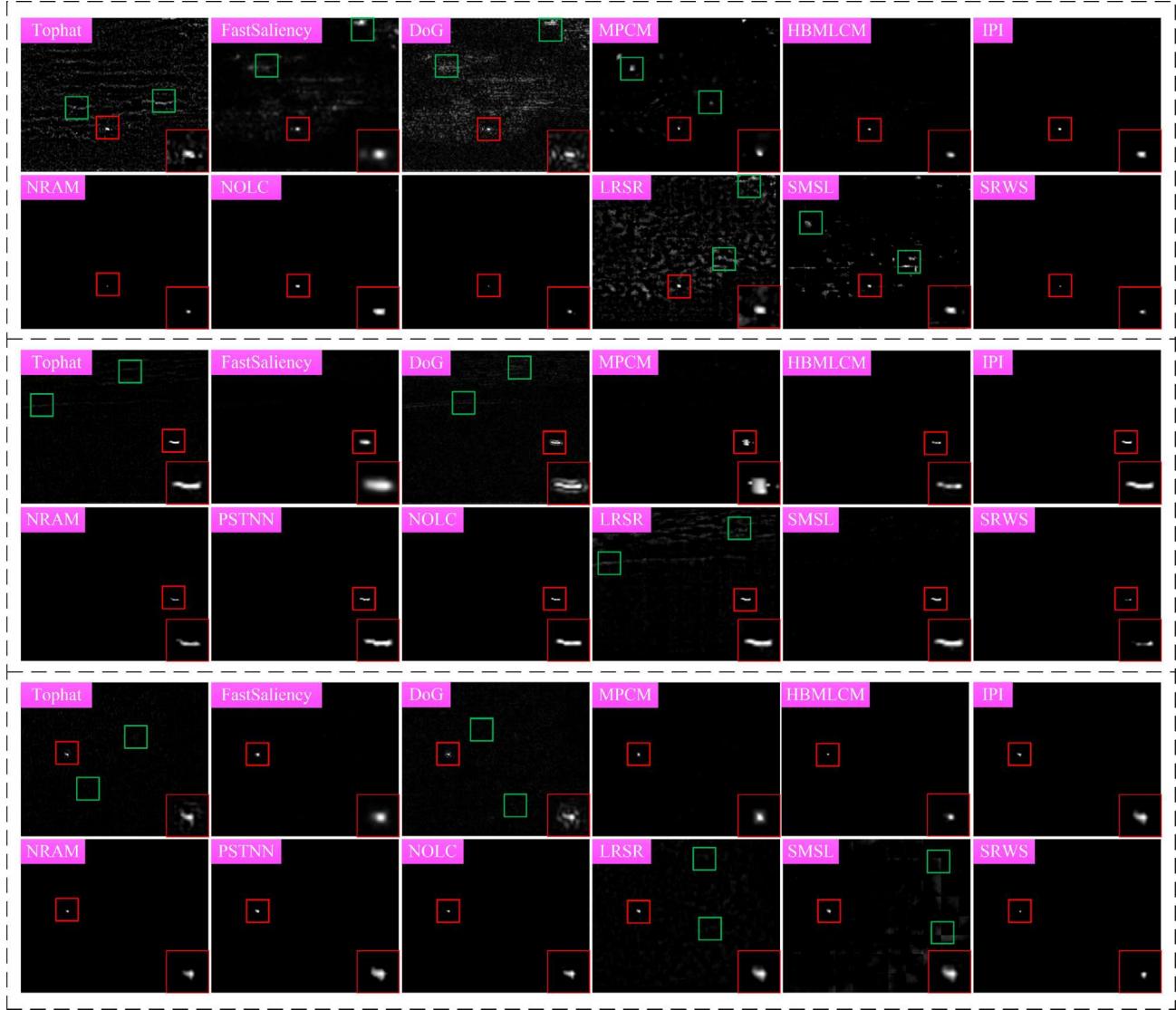


Fig. 16. Comparison results of baselines in Seq. 1–3.

4.1.1. BSF and SCRG

As quantitative evaluation metrics, BSF and SCRG [16,18] can well describe the suppression effect of background and highlight degree of target. They are defined as follow:

$$\text{SCRG} = \frac{S_{\text{out}}/C_{\text{out}}}{S_{\text{in}}/C_{\text{in}}}, \quad \text{BSF} = \frac{C_{\text{in}}}{C_{\text{out}}} \quad (17)$$

where S represents the target amplitude and C represents the standard deviation of the background region; in and out are the input original infrared image and the output detection result, respectively.

In the previous definition of SCRG and BSF, the results often had a lot of infinite value which is shown as Inf [69], which greatly interfered with the quantitative evaluation of the algorithm. We modified the definition of background region to avoid this problem. Fig. 6 shows the definition of the background region and target region. Where, the red square represents target region with diameter a ; the black square represents background region with diameter b . Here, we should make the target region completely frame targets. Considering the size change of targets, we set $a = 11$ and $b = 81$.

From the definition, we can see that the larger the value of BSF and SCRG, the stronger the algorithm's ability to suppress the background and highlight the target.

4.1.2. ROC

Another evaluation method is ROC, which is widely used in signal detection. And it also has a good development in the field of machine learning and data mining. The abscissa of ROC is false positive rate (FPR), and the ordinate is true positive rate (TPR). It uses line to show the dynamic relationship between FPR and TPR. The definition is as follows:

$$\text{FPR} = \frac{\#\text{detected background area}}{\#\text{real targets}} \quad (18)$$

$$\text{TPR} = \frac{\#\text{detected targets}}{\#\text{real targets}} \quad (19)$$

We think the closer the curve is to the top left corner, the better the algorithm works. For quantitative purposes, there is another parameter: area under the curve (AUC). In comparison, AUC can be used as a criterion for the merits of the algorithm. The larger the AUC, the more accurate the algorithm detection.

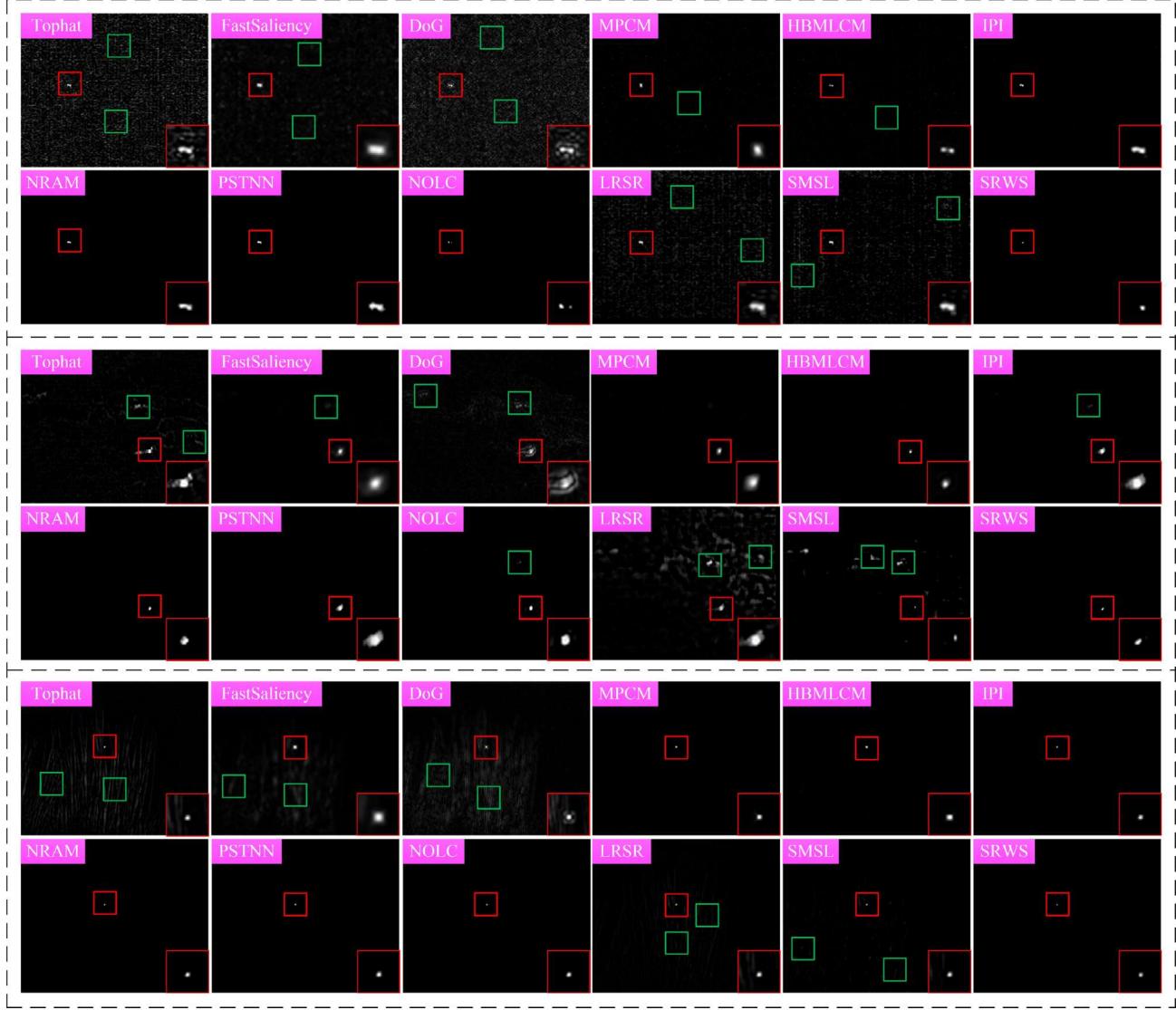


Fig. 17. Comparison results of baselines in Seq. 4–6.

4.1.3. Running time and iteration times

The optimization-based infrared small target detection methods cannot avoid using the optimization method to solve the model iteratively, so their computational efficiency has been widely concerned. Among them, because of the different of model construction, the choice of iterative solution model and the termination conditions, the running time of the methods are different. Therefore, the running time of the algorithm becomes an interesting metric. Meanwhile the iteration times of the algorithm is directly related to the running time. We also regard the iteration times as an important quantitative evaluation metrics.

4.2. Experiments setup

In this section, we introduce the experimental setup in this paper. In order to fully demonstrate the effectiveness of the SRWS method, images of 20 complex scenes as shown in the Fig. 7 were used for visual comparison. For the convenience to show, we reshape the image into the same size. The upper left corner of each image is labeled from (a) to (l). The target area in the image is circled in red, and the magnification is placed in the lower right corner of the image for better observation.

As we can see, these images include cloud background, highlighted clutter, sea level, structural interference of the detector, bridges, grass and other interference factors. Some images have complex background, and some have very low signal-to-noise ratios that make it difficult to identify the target. Target shape has dot shape, strip shape. Targets vary in size, some as small as two or three pixels. Being able to detect all of these targets is a difficult challenge.

For quantitative evaluation, we used six infrared target image sequences. Table 2 shows the details and Fig. 8 shows representative images of the six sequences, more specifically, their 3D display is given in Fig. 9. There are target size variation, large amount of background clutter and low signal-to-noise ratio in sequential images. We carried out parameter setting experiments and quantitative evaluation with the baselines in these six sequences.

In this paper, we demonstrate the effectiveness of SRWS by comparing with eleven different baselines. Among the BS-based methods, Tophat [12] is the most classical detection method and fast saliency [18] used facet model to detect the target. DoG [22], MPCM [27] and HBMLCM [28] are methods based on human vision system. Among the optimization-based methods, IPI method [10] is the first of its kind. NRAM [34], PSTNN [38] and NOLC [35] are methods based on the assumption of a single subspace. In addition,

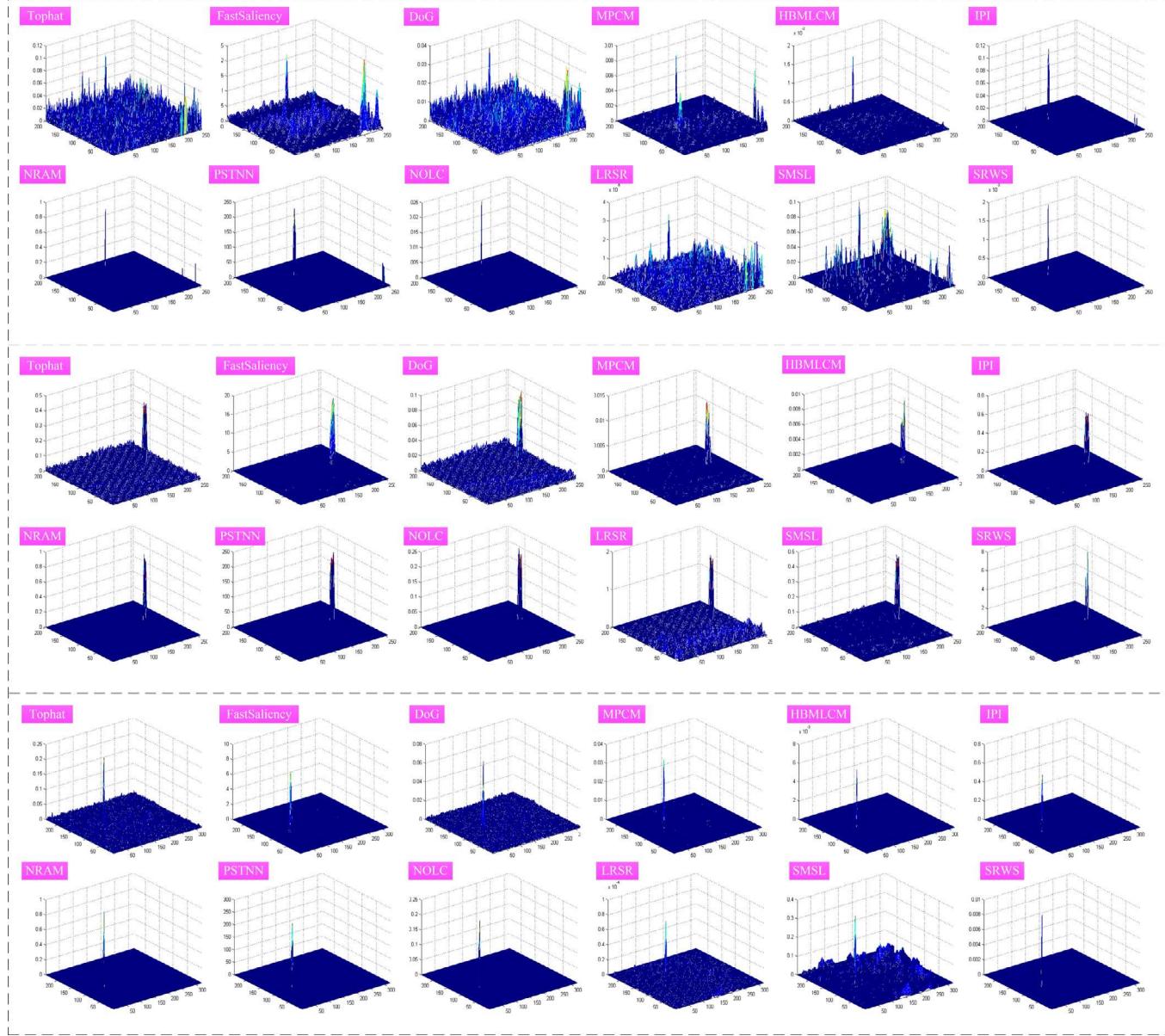


Fig. 18. 3D display of baselines comparison results in Seq. 1–3.

we chose two methods based on the multi-subspaces hypothesis, LRSR [52] and SMSL [33], to verify that the SRWS method is still effective in the similar detection methods. The specific parameter settings for these methods are shown in Table 3.

4.3. Parameter setting

In this section, we experimentally determine the values of key parameters in SRWS model. There are six key parameters: patch size len ; step of patch; when constructing OEI, the number of overlapping groups l and the stretching coefficient α ; the coefficients λ_L and γ_L are used to get the penalty factors λ and γ , here we tested the Seq. 2–5.

The results of ROC are shown in Fig. 10, where (a)–(d) represent the Seq. 2–5, and the top to bottom are the comparison of len , step and l , respectively. len determines the size of patch image. The larger the len value, the more information will be contained in a patch. However, if the value is too large, the patch image sampling will be too small, which is not conducive to the detection of

targets. On the contrary, the smaller len is, the less information is contained in a patch, but too little will lead to too many patches and greatly increase the running time. As shown in the figure, we strike a balance between the two, and the value of len is suggested to 50. step also determines the size of patch image, similar to len . Too large will lead to insufficient sampling, and too little will lead to too large patch image size and increased running time. Based on the observations in Fig. 10, we recommend the step value as 10.

l is the number of overlapping groups, which determines the amount of information contained in the OEI. From the third row of Fig. 10, we can see that when l is 4, better results can be obtained in each sequence. α determines how different the elements in OEI are, and the greater the difference, the better the ability to display structural features, and the better the suppression of the background. The first row in Fig. 11 shows the α comparison data. The α is recommended to 2.

The arrangement of Fig. 11 is the same as that of Fig. 10. (a)–(d) is the sequence 1–4, and the comparison of parameters α , λ_L and γ_L is shown from top to bottom. λ_L and γ_L are penalty factors. The

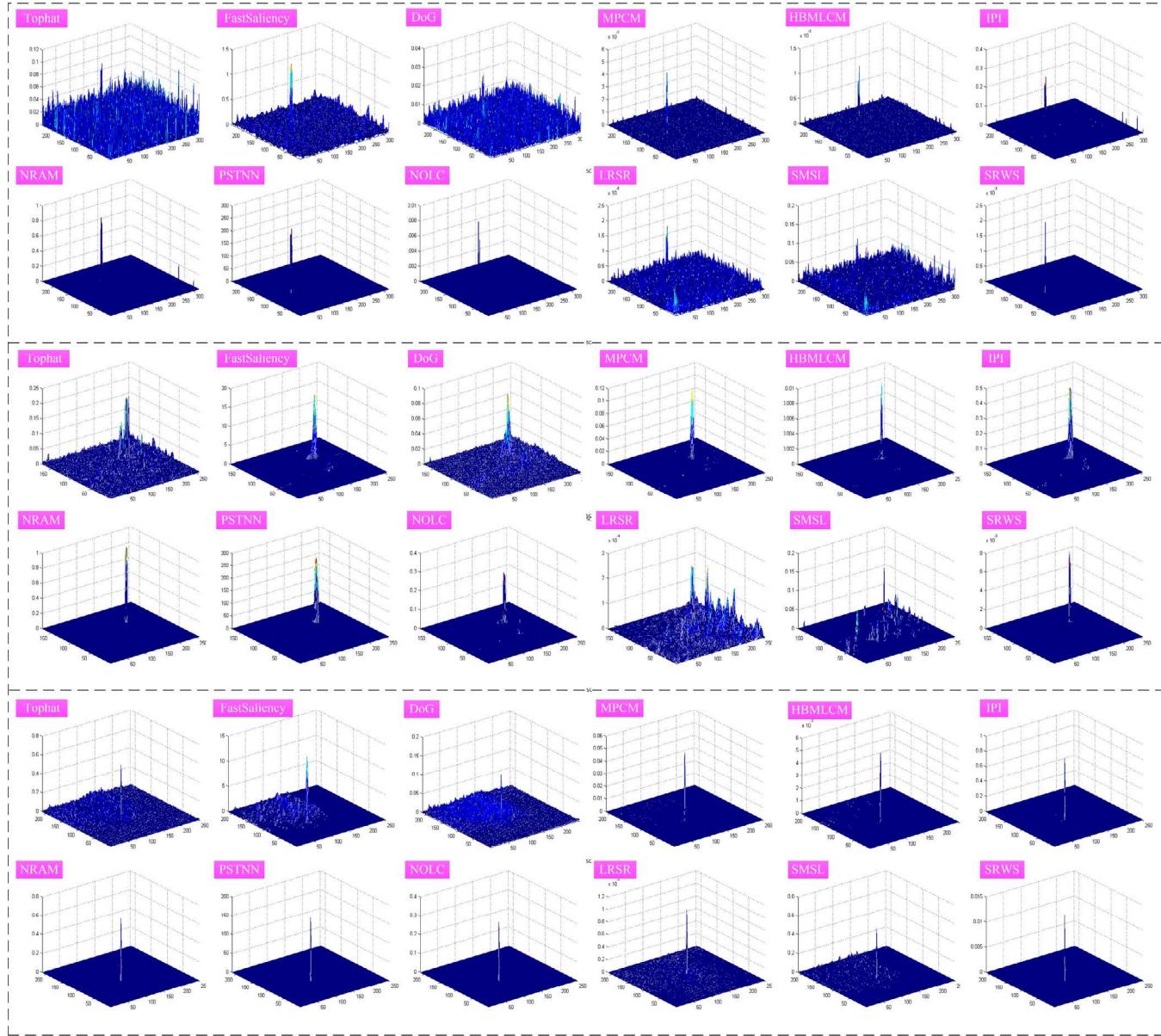


Fig. 19. 3D display of baselines comparison results in Seq. 4–6.

experiment in the second and third rows of Fig. 11 shows the difference between different values. Considering the robustness of the algorithm, we suggest that λ_L is 1 and γ_L is 0.09.

4.4. Visual evaluation

In this section, we visually evaluate the SRWS to verify its multi-scene adaptability, noise adaptability, and excellent performance against the baselines, respectively.

4.4.1. Multi-scene adaptability

We used the SRWS model to detect the infrared small target images in twenty complex scenes. Fig. 12 shows the specific experimental results, in which the upper left corner is the image number, and the detected target position in the image is framed in red, and placed in the lower right corner. Fig. 13 shows 3D display of the detection results. We can observe that SRWS can accurately

detect targets when dealing with complex scenes, which ensures the robustness of the algorithm.

In addition, as shown in Figs. 14 and 15, the coefficient matrix Z and noise matrix E obtained by SRWS model are shown. It can be clearly seen that Z in Fig. 14 is not the identity matrix, and has a significant response in the off-diagonal region, which indicates that the assumption in this paper that background elements may come from multi-subspaces is reasonable. Combined with Fig. 15 can isolate the obvious structural noise, also suggests that by OEI constrained sparse item is feasible.

4.4.2. Visual comparison with baselines

In addition, we used sequential images to visually compare the SRWS with the baselines. Figs. 16 and 17 show the comparison results of Seq. 1–3 and Seq. 4–6 with the baselines. Each dashed box represents a sequence. At the same time, Figs. 18 and 19 are 3D display of the detection results. The mark is the same as in

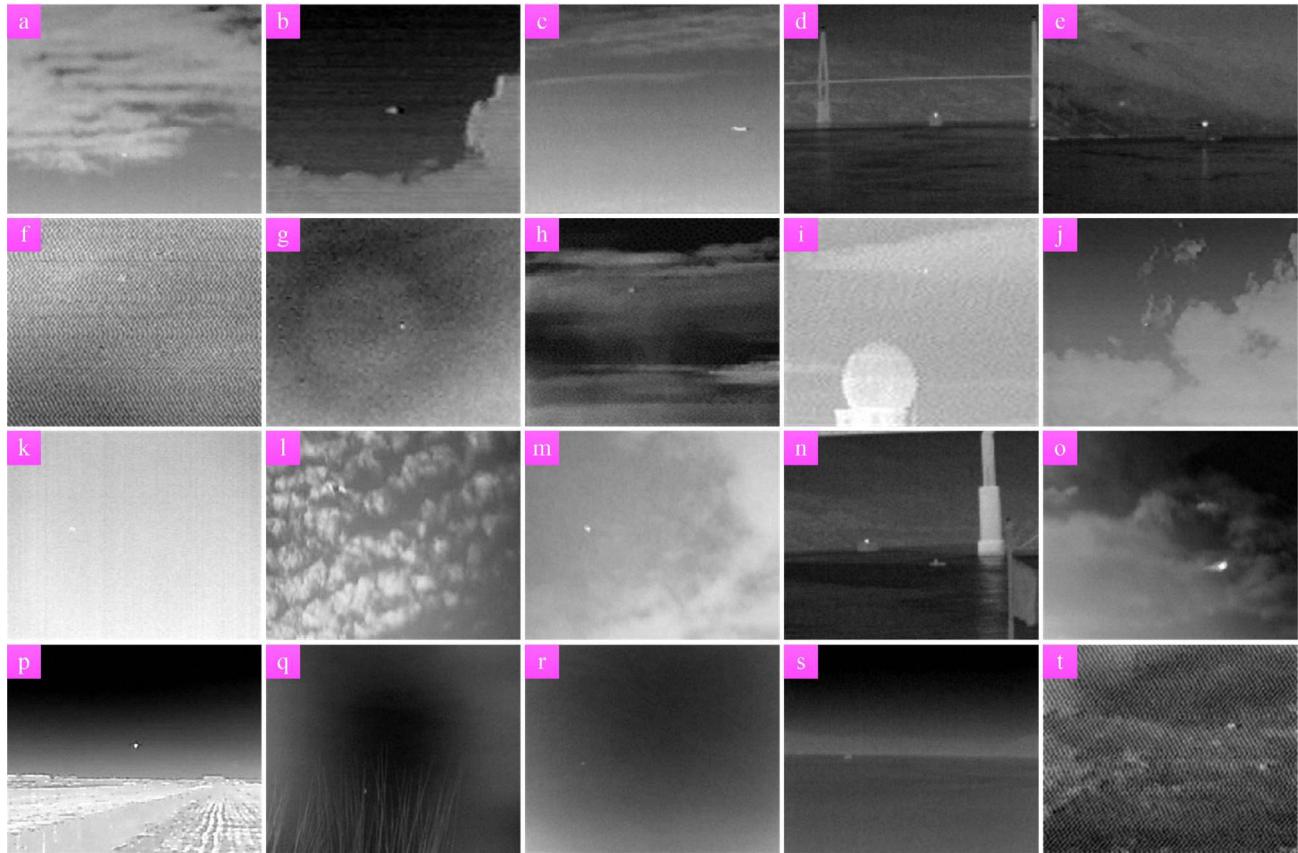


Fig. 20. Complex background images with variance of 0.01 noise was added.

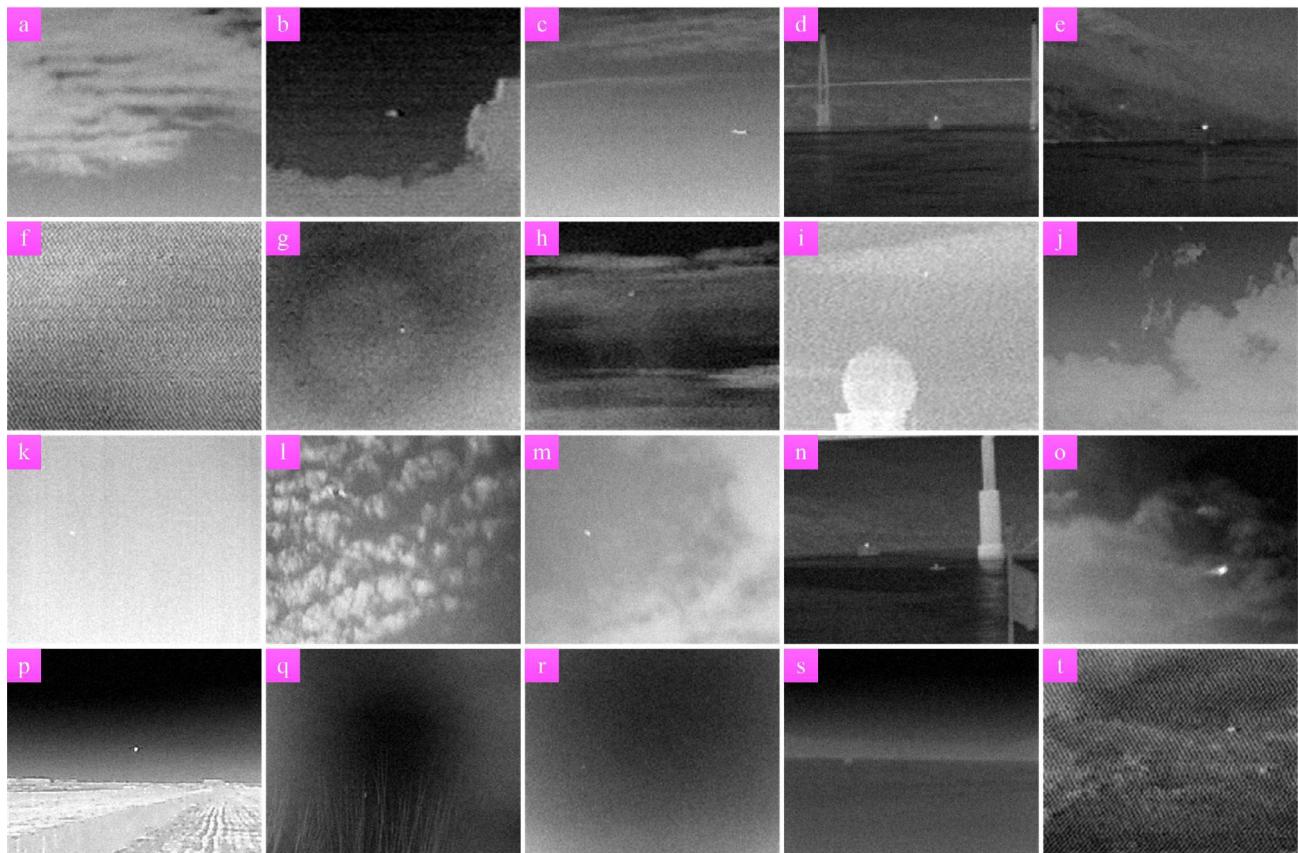


Fig. 21. Complex background images with variance of 0.03 noise was added.

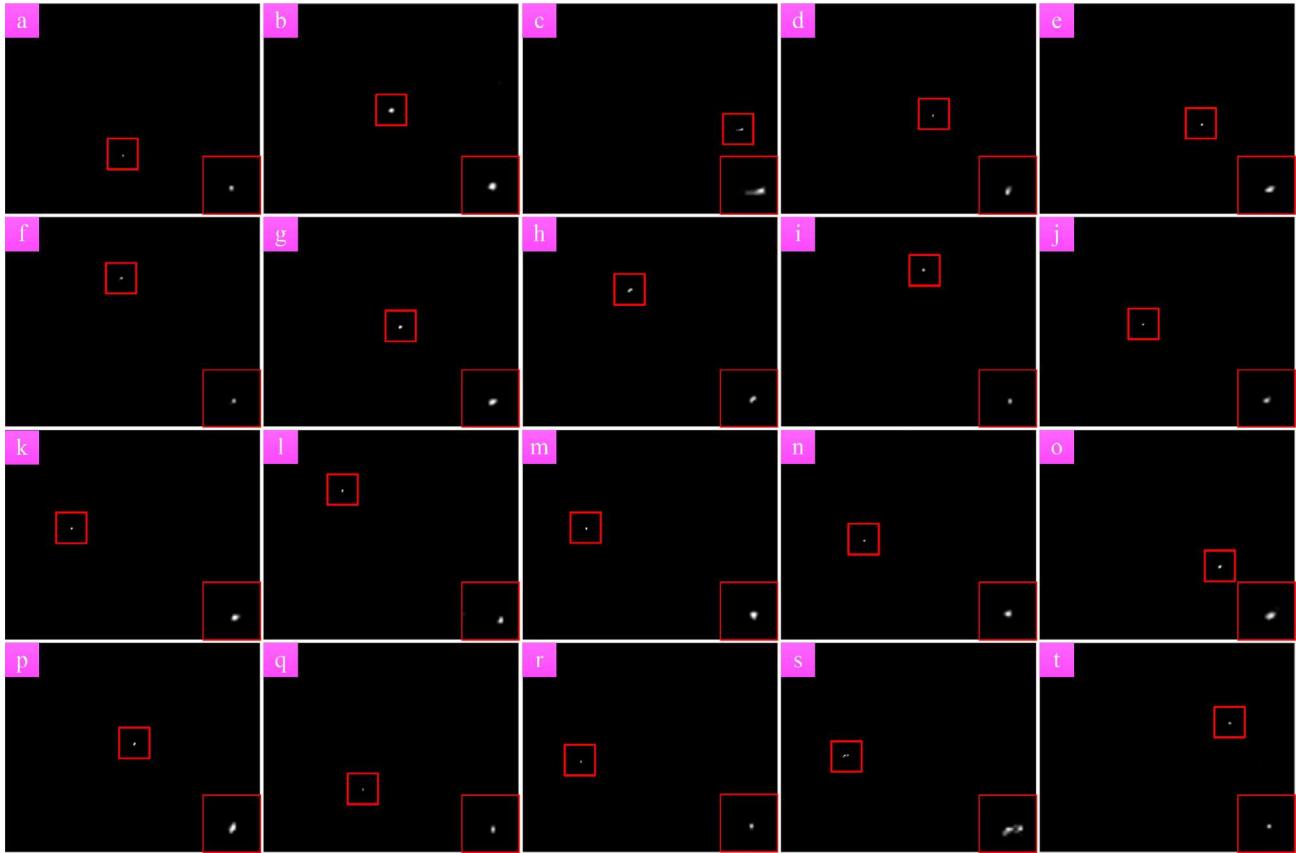


Fig. 22. Detection results of images with variance of 0.01.

the figure above. In some images, the clutter that is not easily observed is framed in green. (See Figs. 16 and 17).

In the figures, we can see that the detection results of Tophat, FastSaliency and DoG are very rough, and there are a lot of clutter in the results. As excellent HVS method, MPCM and HBMLCM can accurately detect the target, but there are still clutter in the background. IPI is a classical method based on optimization and the results are not ideal. NRAM, PSTNN and NOLC that proposed in recent years all have good detection effect. LRSR and SMSL based on the hypothesis of multi-subspaces have poor results. By contrast, the SRWS results are relatively pure in background and can yield better results. (See Figs. 18 and 19).

4.4.3. Noise adaptability

In addition, we also tested the adaptability of SRWS to noise. As shown in Figs. 20 and 21, we used twenty images with complex backgrounds, the upper left corner being the label, and added normal noises with mean value 0, variance 0.01 and 0.03 to them respectively.

Figs. 22 and 23 show the detection results of the SRWS model on the noise images. The target position is enlarged and placed in the lower right corner. Further, 3D display of their results is given in Figs. 24 and 25. Thanks to the estimation of the background by the self-regularization item and the suppression of the background by OEI, it can be seen that SRWS can always detect the target accurately when the noise becomes larger. Therefore, we can say that the proposed method is adaptive to noise.

4.5. Quantitative evaluation

In this section, we use quantitative evaluation to compare the advantages and disadvantages of each algorithm. Fig. 26 shows

ROC for Seq. 1–6, the upper left corner of the image is the sequence identifier. We can see that the dark red line represents the SRWS method, which is closer to the upper right corner of the curves in all six sequences. For more accurate evaluation, the AUC of each curve is summarized in Table 4. The maximum value of the AUC in each sequence is marked in bold red, and the second largest value in blue.

As can be seen from Table 4, the SRWS method obtained the largest AUC in five sequences, and the other sequence also obtained the second largest AUC. MPCM, HBMLCM and PSTNN, as excellent detection algorithms, also play an important role in comparison. AUC can directly reflect the detection accuracy of the algorithm, so we can see that compared with other methods, the detection accuracy of SRWS is higher, and the robustness of the algorithm is also guaranteed.

The other two key metrics are SCRG and BSF, and we have summarized their test data in Table 5. As in Table 4, the maximum value is marked with bold red and the second largest value is marked with blue.

As can be seen from Table 5, the Tophat, FastSaliency, DoG and LRSR methods only slightly highlight the target, but the results are not outstanding. MPCM, NRAM and NOLC can max out in one or two sequences, but not in all. HBMLCM gets a maximum value, five second maximum values, and does a good job of highlighting the target and suppressing the background. The SRWS model proposed in this paper has maximum value and second maximum value in almost all sequences. It shows that SRWS can suppress the background well, which fully reflects the excellent background suppression ability and robustness of this method. In particular, the effect of SRWS is much better than other algorithms in methods based on multi-subspaces assumptions.

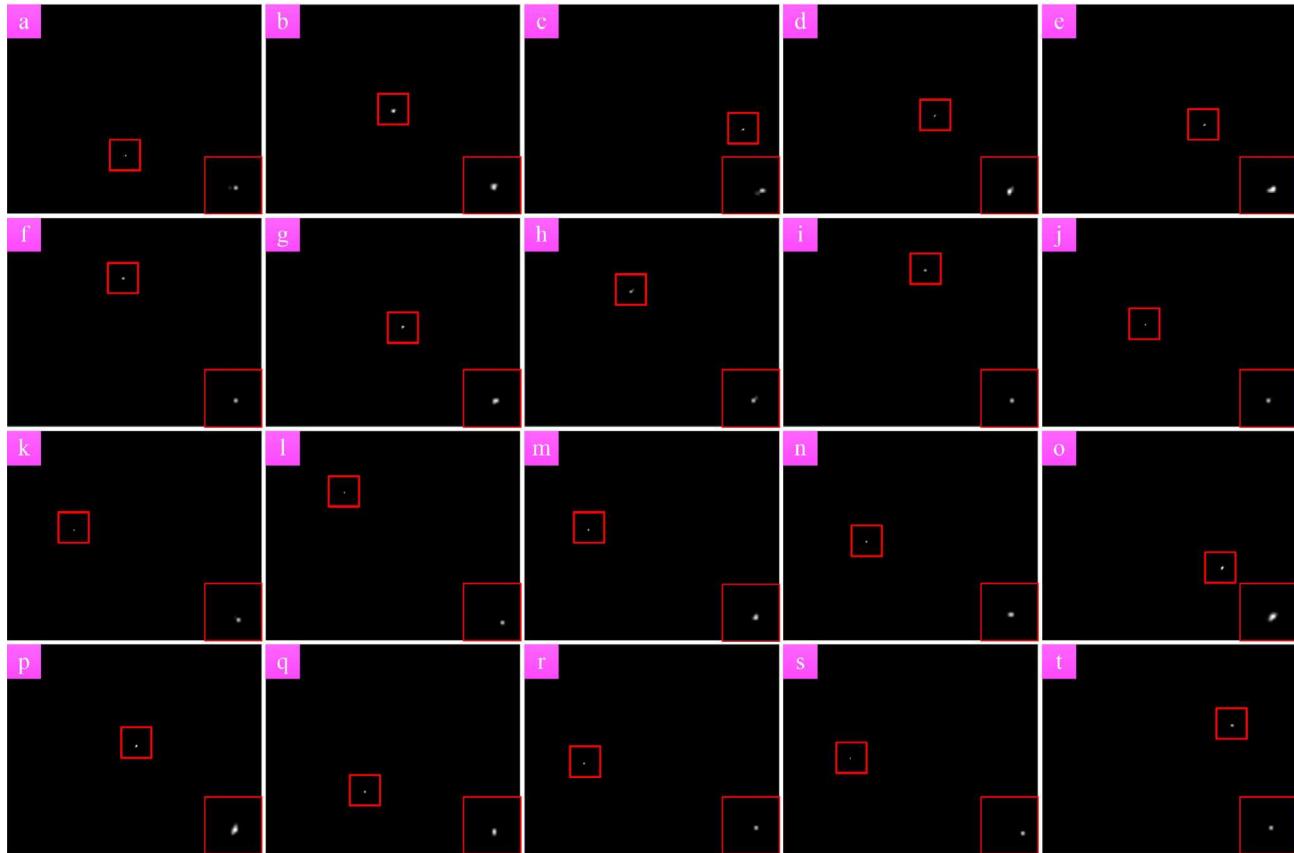


Fig. 23. Detection results of images with variance of 0.03.

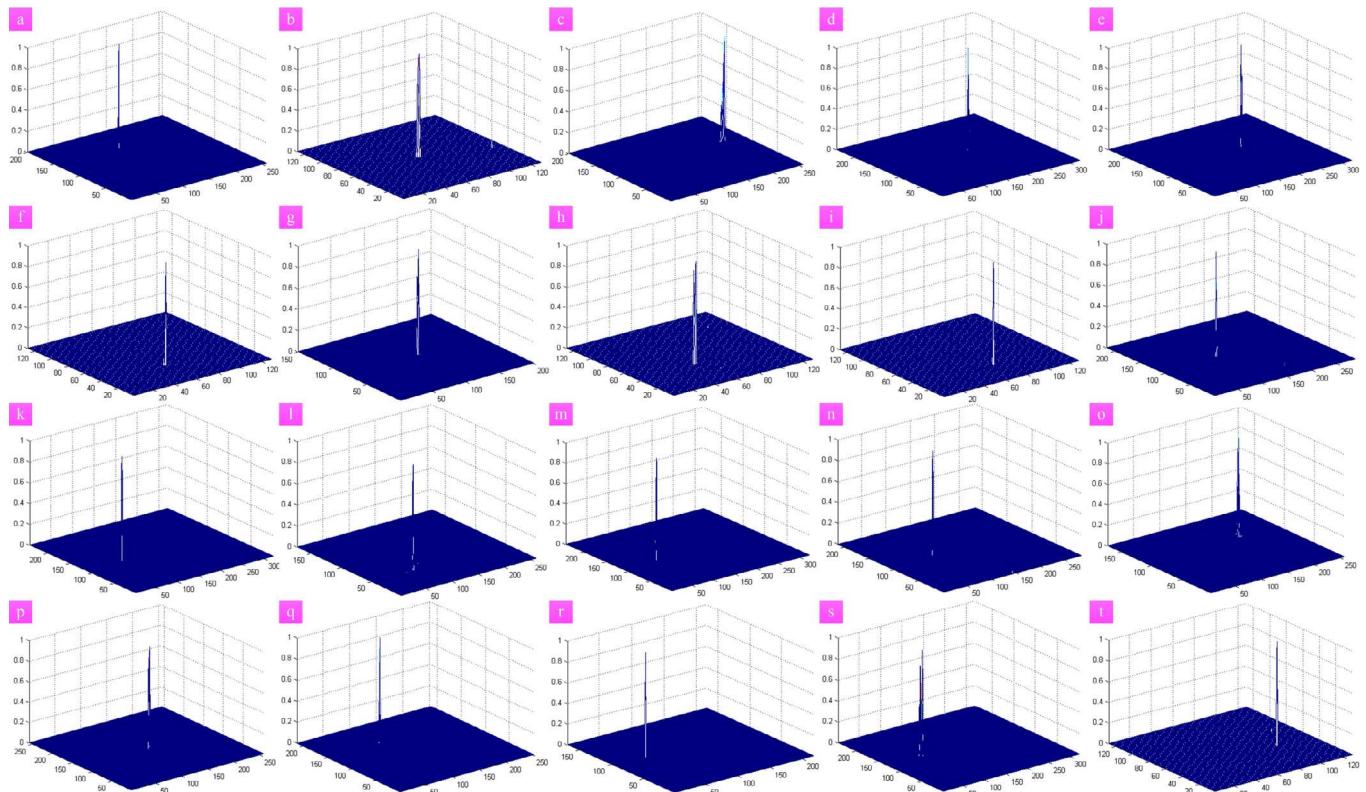


Fig. 24. 3D display of detection results with variance of 0.01.

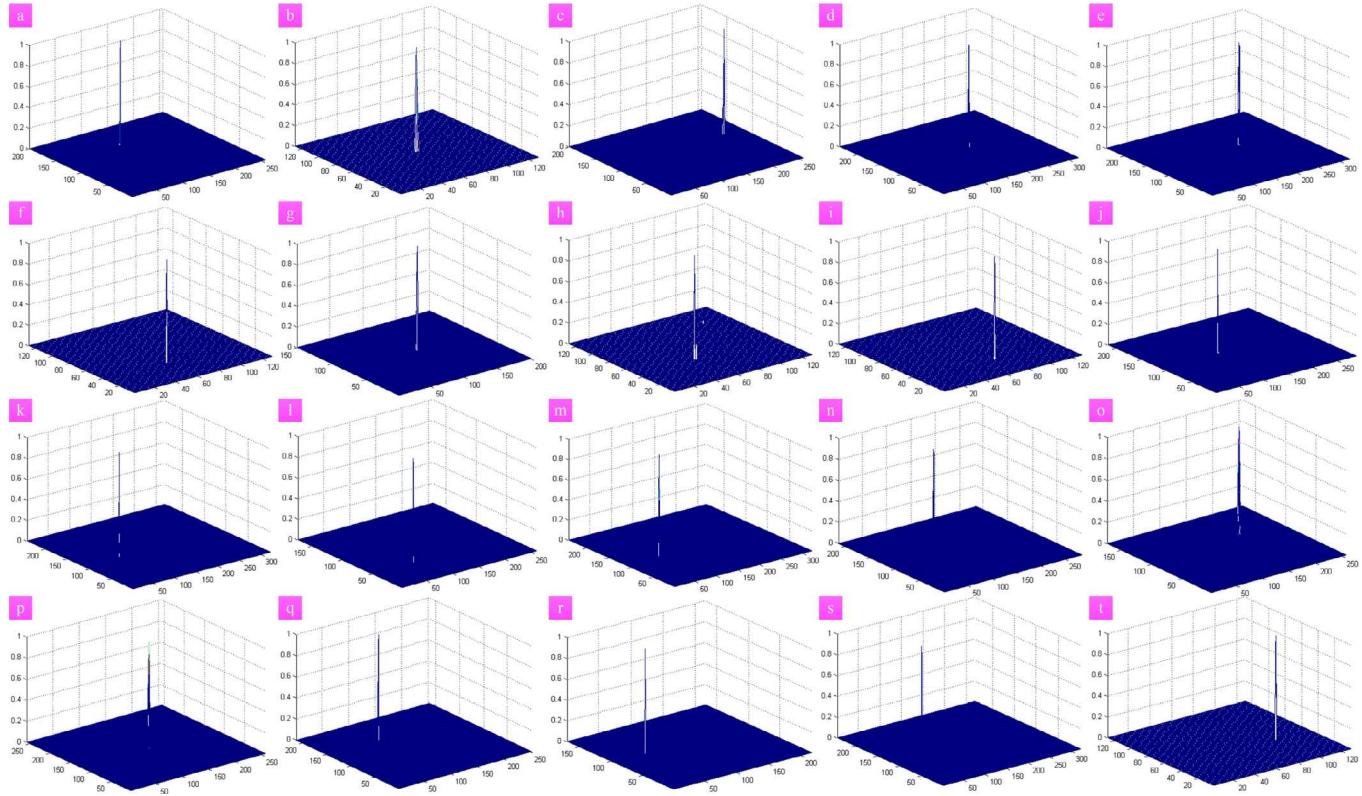


Fig. 25. 3D display of detection results with variance of 0.03.

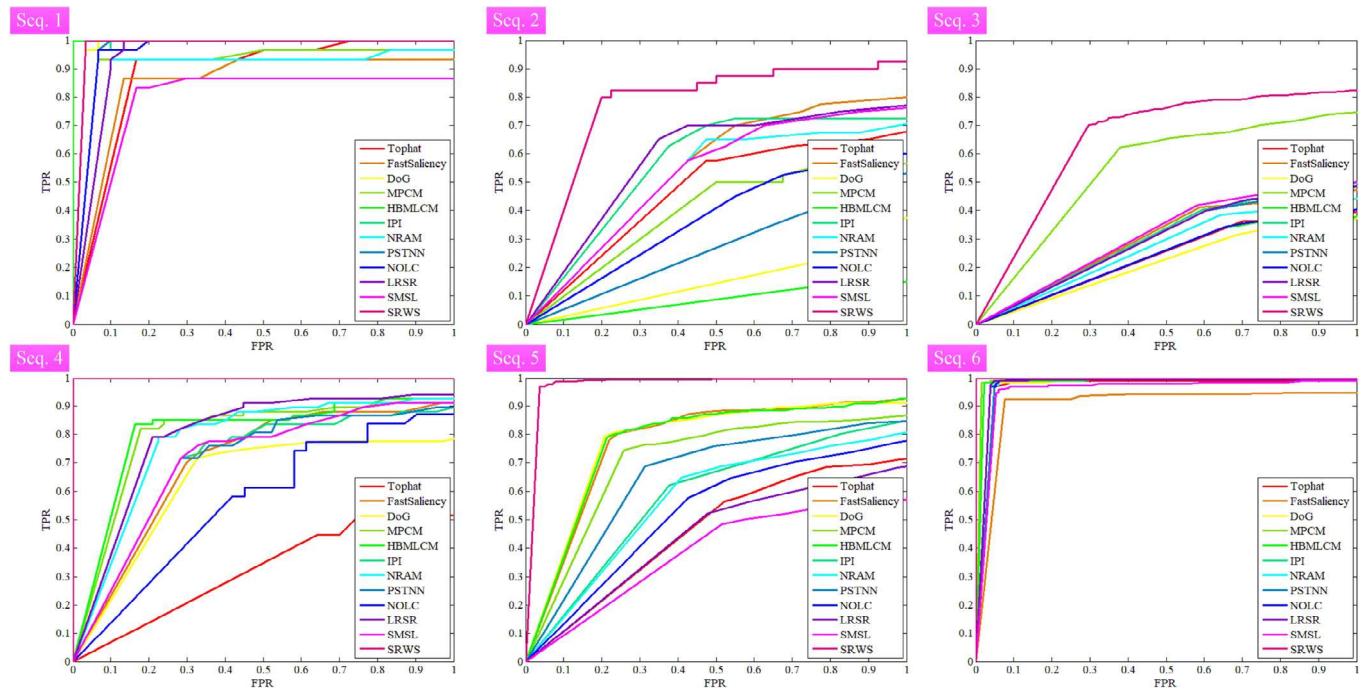


Fig. 26. ROC of comparison methods.

Table 4
Summary of AUC for comparison method.

	Seq. 1	Seq. 2	Seq. 3	Seq. 4	Seq. 5	Seq. 6
Tophat	0.8839	0.4648	0.2401	0.3211	0.4615	0.9632
FastSaliency	0.8500	0.5441	0.3005	0.6988	0.7714	0.9021
DoG	0.9817	0.1619	0.2168	0.6314	0.7755	0.9843
MPCM	0.9211	0.3909	0.5443	0.7984	0.7044	0.9900
HBMCLM	1.0000	0.0863	0.2376	0.8029	0.7724	0.9820
IPI	0.9633	0.5631	0.3035	0.6982	0.5797	0.9776
NRAM	0.8917	0.5041	0.2696	0.7733	0.5651	0.9769
PSTNN	0.9639	0.2682	0.2961	0.6993	0.6465	0.9722
NOLC	0.9606	0.3697	0.2403	0.5614	0.5215	0.9785
LRSR	0.9433	0.5832	0.2986	0.7948	0.4429	0.9745
SMSL	0.7889	0.5242	0.3161	0.7047	0.3850	0.9517
SRWS	0.9828	0.7763	0.6541	1.0000	0.9755	1.0000

Table 5
Summary of SCRG and BSF for comparison method.

	Seq. 1		Seq. 2		Seq. 3		Seq. 4		Seq. 5		Seq. 6	
Method	SCRG	BSF										
Tophat	6.3969	2.0782	3.4307	1.5590	2.8064	1.6040	2.2467	0.3975	3.7037	4.0821	2.2402	1.7165
FastSaliency	6.4033	2.3978	2.8326	1.8586	2.1891	6.3963	4.1472	1.1227	3.6363	6.2656	2.0563	2.2536
DoG	5.5417	1.7329	3.3879	1.5334	2.5953	1.5048	2.6060	0.4820	4.0097	4.2259	1.9759	1.3681
MPCM	12.226	5.3239	3.7264	3.1099	2.6134	49.936	7.2829	1.5212	5.3045	8.8853	6.2310	7.6776
HBMCLM	12.794	5.6544	5.5511	3.2769	3.8680	78.491	7.4254	2.3001	6.7690	22.098	4.9018	16.835
IPI	10.948	3.4274	3.5667	1.4685	1.8502	41.032	7.8733	1.5973	3.7158	3.1407	7.0574	3.9225
NRAM	14.125	3.8559	4.3498	1.5619	2.9677	1.2487	8.4500	0.9020	4.7645	3.0905	7.5243	4.4799
PSTNN	9.3872	2.2038	3.5697	1.2802	1.7317	0.7116	6.8179	0.7393	3.7057	2.4336	6.4887	3.8991
NOLC	13.328	5.6862	4.0131	1.7403	3.2480	4.1419	6.5204	7.2412	3.9781	3.7388	7.5074	5.4022
LRSR	5.6900	1.8641	3.3450	1.5437	2.3151	1.9226	3.3960	0.6121	3.6152	5.1400	3.4040	3.0723
SMSL	6.8350	2.1308	3.4681	2.7106	1.6762	1.1856	4.0529	0.6340	4.2973	27.468	4.5380	3.6937
SRWS	17.077	11.349	9.618	7.9536	4.4867	9.2146	7.9625	112.05	7.2124	22.540	9.1459	65.848

Table 6
Running time of the baselines (second).

	Seq. 1	Seq. 2	Seq. 3	Seq. 4	Seq. 5	Seq. 6
Tophat	0.0091	0.0030	0.0041	0.0034	0.0027	0.0030
FastSaliency	0.0048	0.0030	0.0036	0.0032	0.0030	0.0034
DoG	0.0057	0.0019	0.0018	0.0018	0.0017	0.0017
MPCM	0.0779	0.0767	0.0864	0.0865	0.0733	0.0764
HBMCLM	0.3805	0.3767	0.5634	0.5633	0.3257	0.3956
IPI	6.4182	5.9422	13.097	9.3940	5.1757	6.5198
NRAM	1.1281	1.3984	2.3363	2.2646	1.3659	1.5334
PSTNN	0.0723	0.0514	0.0677	0.0704	0.0622	0.0578
NOLC	3.2520	9.2496	8.1923	3.5626	13.092	3.3222
LRSR	0.8352	0.8136	1.3809	1.3851	0.7229	0.9396
SMSL	0.2604	0.2476	0.3509	0.5572	0.2457	0.5994
SRWS	1.3204	1.9830	3.8226	4.2434	1.4641	1.4916

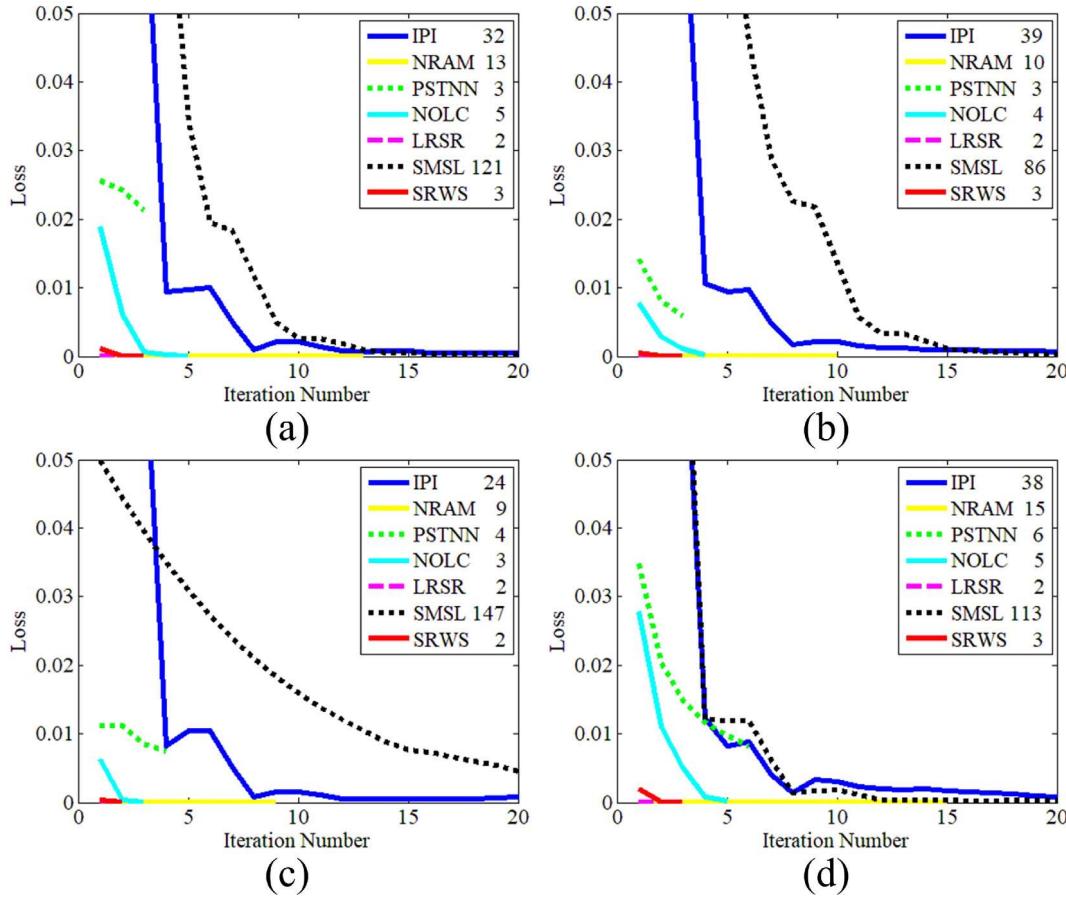


Fig. 27. Comparison of iteration times.

The other key metrics are running time and iteration times, which directly affects the detection efficiency of the algorithm. Table 6 shows a comparison of baselines running time, the data is in seconds. Since the BS-based methods and HVS-based methods are not suitable for optimal iterative solution, we only compare the iterative times of the following 7 methods. Fig. 27 shows a comparison curve of iteration times, where the number after the algorithm name represents the iteration number. The curves represent the relationship between the iteration number and the algorithm loss.

The BS-based and HVS-based methods are fast because it only operates in the spatial domain. In comparison, the optimization-based methods use iterative convergence and the overall running time are slow. Among them, IPI and SMSL use accelerated proximal gradient (APG) method [70] to solve, and their iteration times are relatively high. Other methods using ADMM framework are relatively more efficient, and NOLC, PSTNN and LRSR all have good performance. Because NOLC is non-convex optimization method, although the number of iterations is relatively small, the single iteration takes a long time, which limits the efficiency of the algorithm. The running time and iteration times of NRAM, NOLC and SRWS are at the same level. In comparison, SRWS has acceptable convergence speed, and the efficiency of the algorithm is also guaranteed.

5. Discussion

Both BS-based methods and HVS-based methods operate in spatial domain, and their assumptions of background and target are very simple, which leads to the difficulty in processing complex

background images, including Tophat, FastSaliency, DoG, MPCM and HBMLCM. Their advantage lies in short running time and high calculation efficiency.

In general, the optimization-based methods have better detection accuracy than other types of methods because of the more reasonable assumptions. However, iterative calculation is more time-consuming than spatial domain calculation, which makes the running time of optimization-based methods slightly inferior.

Among them, PSTNN considers the characteristics of target and uses the partial sum of tensor nuclear norm to seek solution in tensor field. However, the parameter setting limits its robustness. As for the single subspace methods, including IPI, NRAM and NOLC, they are improved in terms of target, noise or background. However, due to the limitation of their algorithm that it can only extract the information in single subspace, the estimation of complex background is not very accurate, which affects the detection accuracy.

In the case of multi-subspaces, LRSR uses to construct a complete dictionary itself to learn the target patch image, which makes it less robust. SMSL imposes constraints on the dictionary matrix but the efficiency cannot be guaranteed due to the APG solution.

The SRWS model presented in this paper has several characteristics. In terms of data sources, the hypothesis of multi-subspaces is applied to provide a theoretical basis for the complex background and noise that may come from multi-subspaces. In terms of target, the structure information of the background edge is detected by OEI, and the performance is improved by constraining sparse item. In terms of background, the self-regularization item is used to mine the potential information in background and clutter to extract a more perfect background. However, it still has some

limitations such as the running time of SRWS method is not short enough. And this issue can be solved by using more efficient solution framework or introducing more accurate prior information in future work.

In the experiments in this paper, we show the rationality of the SRWS assumption by showing the coefficient matrix and noise matrix and qualitatively demonstrate that SRWS is robustness to complex scene and noise. Then quantitatively evaluate it with the baselines to illustrate its target detection capability, background suppression capability, fast convergence characteristics and acceptable running time. The experimental results show that SRWS is competent for the detection of infrared small targets.

6. Conclusion

A novel infrared small target detection method called SRWS is proposed in this paper. The algorithm is designed for the hypothesis that data may come from multi-subspaces. Then in terms of target, the OEI, which can detect the background structure information, is applied to constrain the sparse item and enhance the accuracy. Furthermore, the self-regularization item is applied to mine the potential information in background, and extract clutter from multi-subspaces. Therefore, the infrared small target detection problem is transformed into an optimization problem. By combining the optimization function with ADMM, we explained the solution method of SRWS and optimized its iterative convergence condition. The experimental results qualitatively demonstrate that SRWS is robustness to complex scene and noise. In the quantitative evaluation, SRWS outperforms eleven baselines in terms of ROC, SCRG, BSF, running time and iteration times, which also means that it can conspicuously suppress background and detect target.

CRediT authorship contribution statement

Tianfang Zhang: Conceptualization, Methodology, Software, Validation, Writing – original draft, Visualization. **Zhenming Peng:** Formal analysis, Resources, Writing – review & editing, Project administration, Funding acquisition. **Hao Wu:** Methodology, Writing – review & editing. **Yanmin He:** Writing – review & editing, Supervision. **Chaohai Li:** Writing – review & editing, Supervision. **Chunping Yang:** Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by Sichuan Science and Technology Program (2019YJ0167), National Natural Science Foundation of China (61775030, 61571096) and Open Research Fund of Key Laboratory of Optical Engineering, Chinese Academy of Sciences (2017LBC003).

References

- [1] S. Huang, Z. Peng, Z. Wang, X. Wang, M. Li, Infrared small target detection by density peaks searching and maximum-gray region growing, *IEEE Geoscience and Remote Sensing Letters* 16 (12) (2019) 1919–1923.
- [2] S. Huang, Y. Liu, Y. He, T. Zhang, Z. Peng, Structure-adaptive clutter suppression for infrared small target detection: Chain-growth filtering, *Remote Sensing* 12 (1) (2020) 47.
- [3] Y. Lyu, L. Peng, T. Pu, C. Yang, J. Wang, Z. Peng, Cirrus detection based on rpca and fractal dictionary learning in infrared imagery, *Remote Sensing* 12 (1) (2020) 142.
- [4] X. Fan, Z. Xu, J. Zhang, Y. Huang, Z. Peng, Dim small targets detection based on self-adaptive caliber temporal-spatial filtering, *Infrared Physics & Technology* 85 (2017) 465–477.
- [5] X. Guan, Z. Peng, S. Huang, Y. Chen, Gaussian scale-space enhanced local contrast measure for small infrared target detection, *IEEE Geoscience and Remote Sensing Letters* 17 (2) (2019) 327–331.
- [6] L. Peng, T. Zhang, Y. Liu, M. Li, Z. Peng, Infrared dim target detection using shearlet's kurtosis maximization under non-uniform background, *Symmetry* 11 (5) (2019) 723.
- [7] I.S. Reed, R.M. Gagliardi, L.B. Stotts, Optical moving target detection with 3-d matched filtering, *IEEE Transactions on Aerospace and Electronic Systems* 24 (4) (1988) 327–336.
- [8] K.A. Melendez, 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.
- [9] U.M. Braga-Neto, M. Choudhury, J.I. Goutsias, Automatic target detection and tracking in forward-looking infrared image sequences using morphological connected operators, *Journal of Electronic Imaging* 13 (4) (2004) 802–814.
- [10] C. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, A.G. Hauptmann, Infrared patch-image model for small target detection in a single image, *IEEE Transactions on Image Processing* 22 (12) (2013) 4996–5009.
- [11] E.J. Candès, X. Li, Y. Ma, J. Wright, Robust principal component analysis?, *Journal of the ACM (JACM)* 58 (3) (2011) 1–37.
- [12] V.T. Tom, T. Peli, M. Leung, J.E. Bondaryk, Morphology-based algorithm for point target detection in infrared backgrounds, in: *Signal and Data Processing of Small Targets 1993*, vol. 1954, International Society for Optics and Photonics, 1993, pp. 2–11.
- [13] M.M. Hadhoud, D.W. Thomas, The two-dimensional adaptive lms (tdlms) algorithm, *IEEE Transactions on Circuits and Systems* 35 (5) (1988) 485–494.
- [14] S.D. Deshpande, M.H. Er, R. Venkateswarlu, 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] G.-D. Wang, C.-Y. Chen, X.-B. Shen, Facet-based infrared small target detection method, *Electronics Letters* 41 (22) (2005) 1244–1246.
- [16] C. Guo, Q. Ma, L. Zhang, Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform, in: *2008 IEEE Conference on Computer Vision and Pattern Recognition IEEE*, 2008, pp. 1–8.
- [17] J. Han, Y. Ma, J. Huang, X. Mei, J. Ma, An infrared small target detecting algorithm based on human visual system, *IEEE Geoscience and Remote Sensing Letters* 13 (3) (2016) 452–456.
- [18] S. Qi, G. Xu, Z. Mou, D. Huang, X. Zheng, A fast-saliency method for real-time infrared small target detection, *Infrared Physics & Technology* 77 (2016) 440–450.
- [19] H. Takeda, S. Farsiu, P. Milanfar, Kernel regression for image processing and reconstruction, *IEEE Transactions on Image Processing* 16 (2) (2007) 349–366.
- [20] Y. Li, Y. Zhang, Robust infrared small target detection using local steering kernel reconstruction, *Pattern Recognition* 77 (2018) 113–125.
- [21] C.P. Chen, H. Li, Y. Wei, T. Xia, Y.Y. Tang, A local contrast method for small infrared target detection, *IEEE Transactions on Geoscience and Remote Sensing* 52 (1) (2013) 574–581.
- [22] X. Wang, G. Lv, L. Xu, Infrared dim target detection based on visual attention, *Infrared Physics & Technology* 55 (6) (2012) 513–521.
- [23] J. Han, Y. Ma, B. Zhou, F. Fan, K. Liang, Y. Fang, A robust infrared small target detection algorithm based on human visual system, *IEEE Geoscience and Remote Sensing Letters* 11 (12) (2014) 2168–2172.
- [24] H. Deng, X. Sun, M. Liu, C. Ye, X. Zhou, Small infrared target detection based on weighted local difference measure, *IEEE Transactions on Geoscience and Remote Sensing* 54 (7) (2016) 4204–4214.
- [25] H. Deng, X. Sun, M. Liu, C. Ye, X. Zhou, Entropy-based window selection for detecting dim and small infrared targets, *Pattern Recognition* 61 (2017) 66–77.
- [26] X. Bai, Y. Bi, Derivative entropy-based contrast measure for infrared small-target detection, *IEEE Transactions on Geoscience and Remote Sensing* 56 (4) (2018) 2452–2466.
- [27] Y. Wei, X. You, H. Li, Multiscale patch-based contrast measure for small infrared target detection, *Pattern Recognition* 58 (2016) 216–226.
- [28] Y. Shi, Y. Wei, H. Yao, D. Pan, G. Xiao, High-boost-based multiscale local contrast measure for infrared small target detection, *IEEE Geoscience and Remote Sensing Letters* 15 (1) (2017) 33–37.
- [29] L. Peng, T. Zhang, S. Huang, T. Pu, Y. Liu, Y. Lv, Y. Zheng, Z. Peng, Infrared small-target detection based on multi-directional multi-scale high-boost response, *Optical Review* 26 (6) (2019) 568–582.
- [30] Y. Dai, Y. Wu, Y. Song, Infrared small target and background separation via column-wise weighted robust principal component analysis, *Infrared Physics & Technology* 77 (2016) 421–430.
- [31] T.-H. Oh, Y.-W. Tai, J.-C. Bazin, H. Kim, I.S. Kweon, Partial sum minimization of singular values in robust pca: Algorithm and applications, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38 (4) (2015) 744–758.
- [32] Y. Dai, Y. Wu, Y. Song, J. Guo, Non-negative infrared patch-image model: Robust target-background separation via partial sum minimization of singular values, *Infrared Physics & Technology* 81 (2017) 182–194.
- [33] X. Wang, Z. Peng, D. Kong, P. Zhang, Y. He, Infrared dim target detection based on total variation regularization and principal component pursuit, *Image and Vision Computing* 63 (2017) 1–9.

- [34] L. Zhang, L. Peng, T. Zhang, S. Cao, Z. Peng, Infrared small target detection via non-convex rank approximation minimization joint l₂, l₁ norm, *Remote Sensing* 10 (11) (2018) 1821.
- [35] T. Zhang, H. Wu, Y. Liu, L. Peng, C. Yang, Z. Peng, Infrared small target detection based on non-convex optimization with l_p-norm constraint, *Remote Sensing* 11 (5) (2019) 559.
- [36] J. Weickert, Coherence-enhancing diffusion filtering, *International Journal of Computer Vision* 31 (2–3) (1999) 111–127.
- [37] Y. Dai, 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* 10 (8) (2017) 3752–3767.
- [38] L. Zhang, Z. Peng, Infrared small target detection based on partial sum of the tensor nuclear norm, *Remote Sensing* 11 (4) (2019) 382.
- [39] E. Elhamifar, R. Vidal, Sparse subspace clustering, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2009, pp. 2790–2797.
- [40] G. Liu, Z. Lin, Y. Yu, Robust subspace segmentation by low-rank representation, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010, pp. 663–670.
- [41] G. Liu, S. Yan, Latent low-rank representation for subspace segmentation and feature extraction, in: 2011 International Conference on Computer Vision, IEEE, 2011, pp. 1615–1622.
- [42] L. Zhuang, H. Gao, Z. Lin, Y. Ma, X. Zhang, N. Yu, Non-negative low rank and sparse graph for semi-supervised learning, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2012, pp. 2328–2335.
- [43] X. Lu, Y. Wang, Y. Yuan, Graph-regularized low-rank representation for destriping of hyperspectral images, *IEEE Transactions on Geoscience and Remote Sensing* 51 (7) (2013) 4009–4018.
- [44] K. Tang, R. Liu, Z. Su, J. Zhang, Structure-constrained low-rank representation, *IEEE Transactions on Neural Networks and Learning Systems* 25 (12) (2014) 2167–2179.
- [45] W. Li, Q. Du, B. Zhang, Combined sparse and collaborative representation for hyperspectral target detection, *Pattern Recognition* 48 (12) (2015) 3904–3916.
- [46] W. Li, J. Liu, Q. Du, Sparse and low-rank graph for discriminant analysis of hyperspectral imagery, *IEEE Transactions on Geoscience and Remote Sensing* 54 (7) (2016) 4094–4105.
- [47] R. Liu, Z. Lin, F. De la Torre, Z. Su, Fixed-rank representation for unsupervised visual learning, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2012, pp. 598–605.
- [48] P. Li, J. Bu, B. Xu, Z. He, C. Chen, D. Cai, Sparse fixed-rank representation for robust visual analysis, *Signal Processing* 110 (2015) 222–231.
- [49] L. Wei, X. Wang, J. Yin, A. Wu, Self-regularized fixed-rank representation for subspace segmentation, *Information Sciences* 412 (2017) 194–209.
- [50] C. Wang, S. Qin, Adaptive detection method of infrared small target based on target-background separation via robust principal component analysis, *Infrared Physics & Technology* 69 (2015) 123–135.
- [51] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, Y. Ma, Robust recovery of subspace structures by low-rank representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35 (1) (2012) 171–184.
- [52] Y. He, M. Li, J. Zhang, Q. An, Small infrared target detection based on low-rank and sparse representation, *Infrared Physics & Technology* 68 (2015) 98–109.
- [53] X. Wang, Z. Peng, D. Kong, Y. He, Infrared dim and small target detection based on stable multisubspace learning in heterogeneous scene, *IEEE Transactions on Geoscience and Remote Sensing* 55 (10) (2017) 5481–5493.
- [54] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena* 60 (1–4) (1992) 259–268.
- [55] W. Stefan, R.A. Renaut, A. Gelb, Improved total variation-type regularization using higher order edge detectors, *SIAM Journal on Imaging Sciences* 3 (2) (2010) 232–251.
- [56] T. Chan, A. Marquina, P. Mulet, High-order total variation-based image restoration, *SIAM Journal on Scientific Computing* 22 (2) (2000) 503–516.
- [57] K. Bredies, K. Kunisch, T. Pock, Total generalized variation, *SIAM Journal on Imaging Sciences* 3 (3) (2010) 492–526.
- [58] X. Liu, Y. Chen, Z. Peng, J. Wu, Z. Wang, Infrared image super-resolution reconstruction based on quaternion fractional order total variation with l_p quasinorm, *Applied Sciences* 8 (10) (2018) 1864.
- [59] X. Liu, Y. Chen, Z. Peng, J. Wu, Infrared image super-resolution reconstruction based on quaternion and high-order overlapping group sparse total variation, *Sensors* 19 (23) (2019) 5139.
- [60] X. Liu, Y. Chen, Z. Peng, J. Wu, Total variation with overlapping group sparsity and l_p quasinorm for infrared image deblurring under salt-and-pepper noise, *Journal of Electronic Imaging* 28 (4) (2019) 043031.
- [61] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, et al., Distributed optimization and statistical learning via the alternating direction method of multipliers, *Foundations and Trends in Machine Learning* 3 (1) (2011) 1–122.
- [62] W. Deng, W. Yin, Y. Zhang, Group sparse optimization by alternating direction method, in: Wavelets and Sparsity XV, vol. 8858, International Society for Optics and Photonics, 2013, p. 88580R.
- [63] I. Selesnick, Total variation denoising (an mm algorithm), NYU Polytechnic School of Engineering Lecture Notes 32.
- [64] I.W. Selesnick, P.-Y. Chen, Total variation denoising with overlapping group sparsity, in: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, IEEE, 2013, pp. 5696–5700.
- [65] J. Liu, T.-Z. Huang, I.W. Selesnick, X.-G. Lv, P.-Y. Chen, Image restoration using total variation with overlapping group sparsity, *Information Sciences* 295 (2015) 232–246.
- [66] M. Yin, J. Gao, Z. Lin, Laplacian regularized low-rank representation and its applications, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38 (3) (2015) 504–517.
- [67] E.T. Hale, W. Yin, Y. Zhang, Fixed-point continuation for l₁-minimization: Methodology and convergence, *SIAM Journal on Optimization* 19 (3) (2008) 1107–1130.
- [68] V.E. Shestopal, Solution of the matrix equation ax - xb = c, *Mathematical Notes of the Academy of Sciences of the USSR* 19 (3) (1976) 275–276.
- [69] C. Gao, L. Wang, Y. Xiao, Q. Zhao, D. Meng, Infrared small-dim target detection based on markov random field guided noise modeling, *Pattern Recognition* 76 (2018) 463–475.
- [70] Z. Lin, A. Ganesh, J. Wright, L. Wu, M. Chen, Y. Ma, Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix, *Coordinated Science Laboratory Report no. UIUC-ENG-09-2214, DC-246*.



Tianfang Zhang received his bachelor's degree from the School of Optoelectronic Information, University of Electronic Science and Technology of China (UESTC) in 2017. Currently, he is pursuing the PhD degree in School of Information and Communication Engineering, UESTC. His main research interests are infrared small target detection, compressed sensing and optimization.



Zhenming Peng (M'06) received his Ph.D. degree in geodetection and information technology from the Chengdu University of Technology, Chengdu, China, in 2001. From 2001 to 2003, he was a Post-Doctoral Researcher with the Institute of Optics and Electronics, Chinese Academy of Sciences, Beijing, China. He is currently a Professor with the University of Electronic Science and Technology of China, Chengdu. His research interests include image processing, signal processing, and target recognition and tracking. Prof. Peng is a member of Institute of Electrical and Electronics Engineers (IEEE), Optical Society of America (OSA), China Optical Engineering Society (COES), and Chinese Society of Astronautics (CSA).



Hao Wu received his B.Sc degree in mathematics from Sichuan university, Chengdu, China in 2015. He is currently pursuing his Ph.D degree with signal and information processing in University of Electronic Science and Technology of China. His research interests include signal processing, inversion problem and optimal methods.



Yanmin He received her Ph.D. degree from the School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu, China, in 2011. She is currently an Assistant Professor with the School of Information and Communication Engineering, University of Electronic Science and Technology of China. She has authored more than 20 peer-reviewed papers, and is the Inventor or Co-Inventor of ten patents. Her research interests include signal processing and pattern recognition.



Chaohai Li received his Bachelor's degree and master's degree from University of Electronic Science and Technology (UESTC) in 1996 and 2001 respectively. He was appointed as a senior engineer in 2009. Currently engaged in the research of radar and signal processing, participated in a number of national defense pre-research projects and horizontal cooperation projects of scientific research institutes as the project leader or principal investigator.



Chunping Yang received his Ph.D. degree in optical engineering from University of Electronic Science and Technology in 2007. Since 1996, he has been engaged in scientific research and teaching in the Lidar Laboratory of University of Electronic Science and Technology. His main research fields are satellite-to-ground remote sensing system simulation and information processing, infrared image processing and application.