Infrared Physics & Technology 68 (2015) 98–109

Contents lists available at ScienceDirect

Infrared Physics & Technology

journal homepage: www.elsevier.com/locate/infrared

Small infrared target detection based on low-rank and sparse representation

YuJie He *, Min Li, JinLi Zhang, Qi An

Xian Research Inst. of Hi-Tech, Xi'an 710025, China

HIGHLIGHTS

- A low-rank and sparse representation model named as LRSR is proposed.
- The proposed model combines the low-rank representation and sparse representation.
- Based on the LRSR, a infrared small target detection method is proposed.
- The presented method yields high detection probability and robustness to noise.

ARTICLE INFO

Article history:
Received 14 July 2014
Available online 24 November 2014

Keywords:
Low rank and sparse representation
Low rank representation
Sparse representation
Infrared small target detection

ABSTRACT

The method by which to obtain the correct detection result for infrared small targets is an important and challenging issue in infrared applications. In this paper, a low-rank and sparse representation (LRSR) model is proposed. This model can describe the specific structure of noise data effectively by utilizing sparse representation theory on the basis of low-rank matrix representation. In addition, LRSR based infrared small target detection algorithm is presented. First, a two-dimensional Gaussian model is used to produce the atoms that construct over-complete target dictionary. Then, the reset image data matrix is decomposed by the LRSR model to obtain the background, noise and target components of the image. Finally, the target position can be determined by threshold processing for the target component data. The experimental results in single objective frame, multi-objective image sequences, and strong noise background conditions demonstrate that the proposed method not only has high detection performance in effectively reducing the false alarm rate but also has strong robustness against noise interference.

1. Introduction

Infrared imaging guide is one of the most accurate guide manners in the precision guide weapons. Thus, infrared small target detection is one of the key techniques for infrared guidance systems and is thus a popular topic of research for military applications. On one hand, small targets are usually submerged in background clutter and heavy noise with low Signal Noise Ratio (SNR) because of the long observation distance during the transmission and scattering in the atmosphere. On the other hand, the targets in the images appear as dim points which make the targets have no obvious feature and texture information useful. Therefore, infrared small target detection becomes difficult because of these two factors [1].

Thus far, methods of infrared target detection can be classified into two categories [2]: detection based on single frame and detection based on sequential frames. The sequential detection methods are processed on the basis of the prior information of the target and background. Thus, these methods cannot achieve satisfactory performance because information can hardly be obtained in military applications. Moreover, the capability of sequential methods commonly depends on the results of single frame detection. Given these characteristics, single frame detection algorithms have attracted considerable attention from researchers, who have proposed various single frame detection methods. These detection methods can be divided into two classes.

One class is the detection of targets using image filtering [3–6]. This type of algorithms enhances targets by estimating the background through filtering, such as the max-mean filter, max-median filter, top-hat filter, and TDLMS filter. These filters can preserve the edges of structural backgrounds and remove clutter by predicting backgrounds and subtracting the filtered image from the original image. However, these methods suffer from a high false-alarm rate caused by the low SNR of infrared images. Another type of method is detecting targets by using

* Corresponding author. Tel.: +86 13572962905.
E-mail address: ksy5201314@163.com (Y. He).

http://dx.doi.org/10.1016/j.infrared.2014.10.022
1350-4495/

machine learning theory. These methods transform the detection task into a pattern recognition problem to facilitate the application of some powerful mathematical tools, such as PCA [7], PPCA [8], NLPCA [9] and Fisher linear discriminator [10]. These methods also generate the model or extract the features of targets by using training samples. Such details are used to determine the target location.

There are many detection methods based on single-frame proposed in the open literature. Zhao et al. [11] presented a detection algorithm based on the sparse representation. This method improve and optimize the target representation by modeling the small targets as a combination of certain target samples which are generated using Gaussian Intensity Model and employ the thresholding to sparsity concentration index to locate the target position. But the SR method focuses target modeling and is lack of description ability for background. Shao et al. [12] presented a method using the contrast mechanism of human visual system which improved the algorithm proposed by Kim in [13]. This method uses a morphological processing to make the targets distinguished easily from background and locate the true target through threshold segmentation. Gao et al. [14] proposed a detection method based on the IPI model. This algorithm transforms the detection task into an optimization problem which recovers low-rank and sparse matrices and can be solved effectively by low-rank theory. Although this method only employs one low-rank subspace assumption of background which usually comes from a mixture of multi-low-rank subspaces, but it works better if there are heavy noise and clutter in the images with highly heterogeneous background. Chen et al. [15] proposed a detection method based on peer group filter (PGF), bi-dimension empirical mode decomposition (BEMD) and local inverse entropy. In this method, the PGF is used to remove the noise and enhance the target, then, the proposed BEMD can estimate the background and locate the target by removing the background and segmenting the intrinsic mode functions using the local inverse entropy. Based on the BEMD, Chen et al. [16] also proposed a spatial–temporal detection method. Such method combines the advantages of time-domain difference and BEMD, which is based on sequential images.

On the basis of the original data drawn from several low-rank subspaces, Low-Rank Representation (LRR) [17] was recently proposed for subspace segmentation or recovery. LRR can decompose the data matrix into the clean matrix described by the self-expressive dictionary with low-rank coefficients and the sparse noise. Considering the underlying structure revealing and background modeling ability of LRR with large errors or outliers, we propose a low-rank and sparse representation (LRSR) model in this paper. This model adds the sparse representation of the special structure into the LRR. Infrared small target detection method based on LRSR is also presented. This method employs sparse representation for the small target on the basis of the low-rank decomposition of an infrared image to separate the target from noise. The background, noise, and target can be modeled by using the proposed method. Thus, the proposed method has better detection performance than baseline algorithms, as demonstrated by the experiment results.

The remainder of this paper is organized as follows. Section 2 describes the robust principal component analysis and low-rank representation model. Section 3 presents the proposed LRSR with its solution, complexity, and convergence. Section 4 provides a detailed description of infrared dim and small target detection method, which is proposed on the basis of LRSR. Section 5 presents the experimental results and the comparisons between the proposed method and conventional algorithms. Section 6 provides the conclusions and future works.

2. Low-rank representation

Considering the given observation matrix $X \in R^{m\times n}$ was generated from a low-rank matrix $X_0 \in R^{m\times n}$ with some of its entries corrupted by an additive error $E \in R^{m\times n}$, the original data X_0 can be recovered by the following regularized rank minimization problem which is adopted by the established robust principal component analysis (RPCA) method [18]

$$\min_{Z,E} \operatorname{rank}(X_0) + \lambda\|E\|_l, \quad \text{s.t. } X = X_0 + E \tag{1}$$

where λ is a parameter and $\|\cdot\|_l$ indicates a certain regularization strategy, such as the squared Frobenius norm $\|\cdot\|_F$ used for modeling specify the Gaussian disturbance, the l_0 norm $\|\cdot\|_0$ adopted for characterizing the random corruptions, and the $l_{2,0}$ norm $\|\cdot\|_{2,0}$ used to deal with sample specific corruptions and outliers. Based on the RPCA which assumes that the underlying data structure is a single low-rank subspace, the IPI model was proposed to small infrared target detection in [14]. However, considering the data is usually drawn from a union of multiple subspaces in most cases, the Low Rank Representation (LRR) is presented recently in [18] as depicted in Fig. 1.

The formula of LRR is as follows:

$$\min_{Z,E} \operatorname{rank}(Z) + \lambda\|E\|_l, \quad \text{s.t. } X = AZ + E \tag{2}$$

where A is a dictionary that linearly spans the union of subspaces. The minimize indicates the lowest-rank representation of data X with respect to a dictionary A. Apparently, Eq. (2) is a highly non-convex optimization problem, but we can relax it by solving the convex problem:

$$\min_{Z,E} \|Z\|_* + \lambda\|E\|_l, \quad \text{s.t. } X = AZ + E \tag{3}$$

where $\|\cdot\|_*$ denotes the nuclear norm of a matrix (i.e., the sum of its singular values). When $A = I$, LRR degenerates to RPCA which can be seen as a special case of LRR model, an appropriate A can ensure that the LRR can reveal the true underlying data structure. Usually,

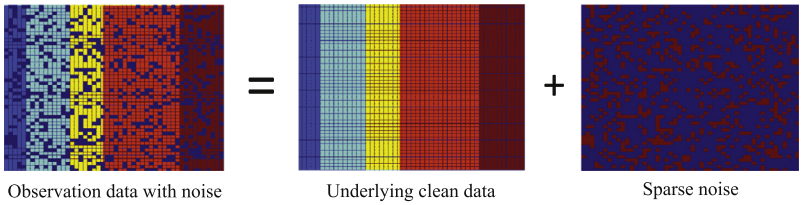

Fig. 1. Diagram of the low-rank decomposition of data.

the observation data X is chosen to be the dictionary. So, Eq. (3) becomes:

$$\min_{Z,E} \|Z\|_* + \lambda\|E\|_l, \quad \text{s.t. } X = XZ + E \tag{4}$$

Due to the ability of LRR to captures the global structure of data even with the grossly contaminated by noise or outliers, the LRR has widespread applications, such as background modeling, removing shadows, and specularities from face images.

3. The low rank and sparse representation (LRSR)

3.1. Model of LRSR

\An infrared image with small target is typically composed of three components, namely, background, target and noise. Generally, the background of the infrared image is transitions slowly and also has the property of non-local self-correlation, the target and noise are small with respect to the whole image. Thus, the infrared image can be regarded as the superposition of low-rank background component and sparse target-noise component. The infrared image has been decomposed by RPCA into background component (denoted as X_0) and target-noise component (denoted as E) [19]. Note that the background data is drawn from a union of multiple subspaces, we decompose the infrared image by using LRR model for obtaining the more accurate results. However, basic LRR cannot distinguish the target and noise data denoted by sparse matrix E. This failure can result in a high false-alarm rate.

Sparse representation based on over-complete dictionary [11,20] can capture the main characteristics of an image by a small amount of nonzero coefficients. This representation can also realize the efficient description of the specific image structure. Therefore, sparse representation is used for infrared small target modeling based on the LRR model to separate the target and noise effectively in this paper.

The noise component is $E = X - AZ$ in Eq. (3) when we use LRR for image decomposition, where $X \in R^{m\times n}$ is the observation data, whereas AZ is the original data represented by the low-rank coefficient. As previously mentioned, the noise component E is composed of the true noise $\tilde{E}$ with small target T as $E = \tilde{E} + T$. In this study, we use the sparse representation to model the small target T, and $T = DS$, where D denotes the over-complete dictionary, whereas S indicates the sparse coefficient. Then we obtain:

$$E = X - AZ = \tilde{E} + DS \tag{5}$$

Eq. (3) could then be rewritten as follows:

$$\min_{Z,\tilde{E}} \|Z\|_* + \lambda\|\tilde{E}\|_l, \quad \text{s.t. } X = AZ + \tilde{E} + DS \tag{6}$$

Considering the known over-complete dictionary $D \in R^{p\times q}$, $q \gg p$, and the sparse representation of target as $T = DS$, thus the optimal sparse representation is the solution of:

$$\min \|S\|_0, \quad \text{s.t. } T = DS \tag{7}$$

where $\|S\|_0$ is the norm l_0 which counts the nonzero entries of a vector. Given that l_0 is highly non-convex, Eq. (7) is an NP problem. Thus, we use l_1 norm of coefficients instead of l_0 norm as:

$$\min \|S\|_1, \quad \text{s.t. } T = DS \tag{8}$$

According to Eqs. (6) and (8), S should be the solution of the optimization problem:

$$\min \|S\|_1, \quad \text{s.t. } DS = X - AZ - \tilde{E}, \quad \text{and} \quad \|Z\|_* < c, \quad \|\tilde{E}\|_l < \varepsilon \tag{9}$$

where c and ε has small positive values. Little transformation in Eq. (9) is applied, we have LRSR model:

$$\min_{\tilde{E},Z,S} \|Z\|_* + \lambda\|\tilde{E}\|_l + \beta\|S\|_1, \quad \text{s.t. } X = AZ + \tilde{E} + DS \tag{10}$$

where λ and β are regular coefficients that balance the three components. In this study, X denotes infrared image, AZ indicates background, $\tilde{E}$ is noise signal and DS denotes target signal. However, the meaning of each component can be different with dictionary D and A selection in other applications.

3.2. Solution of the LRSR model

We use the original data X as the dictionary A and introduce two auxiliary variables and convert Eq. (10) to the following equivalent formula to solve the problem by adopting the augmented Lagrange multiplier method:

$$\min_{J,\tilde{E},Z,S,L} \|J\|_* + \lambda\|\tilde{E}\|_l + \beta\|L\|_1 \tag{11}$$
$$\text{s.t. } X = XZ + \tilde{E} + DS, \quad Z = J, \quad S = L$$

which equals to solving the following augmented Lagrange function

$$\begin{aligned}\mathcal{L} = {} & \|J\|_* + \lambda\|\tilde{E}\|_l + \beta\|L\|_1 + tr\left[Y_1^T(X - XZ - \tilde{E} - DS)\right] \\ & + tr\left[Y_2^T(Z - J)\right] + tr\left[Y_3^T(S - L)\right] \\ & + \frac{\mu}{2}(\|X - XZ - \tilde{E} - DS\|_F^2 + \|Z - J\|_F^2 + \|S - L\|_F^2)\end{aligned} \tag{12}$$

where Y_1, Y_2 and Y_3 are Lagrange multipliers, whereas $\mu > 0$ is a penalty parameter. Eq. (12) can be solved with respect to J, Z, L, and E by fixing other variables, and then the multipliers Y_1, Y_2 and Y_3 are updated. We can employ different regular strategies for noise in different applications. The solution steps are outlined in Algorithm 1 in the case of l_1 norm. The step 2, 3 and 4 of Algorithm 1, which are convex problems, have closed-form solutions. Step 2 can be solved by the following Lemma 1:

Lemma 1 [21]. *For matrix $Y \in R^{n\times d}$ and $\mu > 0$, the problem as follows has the only one analysis solution.*

$$\arg\min_{M\in R^{n\times d}} \mu\|M\|_* + \frac{1}{2}\|M - Y\|_F^2$$

This solution can be described by singular value thresholding operator.

$$SVT_\mu(Y) = U\text{diag}[(\sigma - \mu)_+]V^T \tag{13}$$

$$(\sigma - \mu)_+ = \begin{cases} \sigma - \mu & \sigma > \mu \\ 0 & \text{otherwise} \end{cases}$$

$U \in R^{n\times r}$, $V \in R^{d\times r}$ and $\sigma = (\sigma_1, \sigma_2, \sigma_3 \cdots \sigma_r) \in R^{r\times 1}$ can be achieved by singular value decomposition of matrix Y, $Y = U\Sigma V^T$ and $\Sigma = \text{diag}(\sigma)$.

Algorithm 1. Solving Eq. (12) by inexact ALM algorithms.

Input: data matrix X, parameters λ and β
Initialize: $Z = J = 0, E = 0, S = L = eye(k,d), Y_1 = Y_2 = Y_3 = 0, \mu = 10^{-6}, \mu_{\max} = 10^{10}, \rho = 1.2, \varepsilon = 10^{-6}$
Output: Z, E, S

1: **while** not converged **do**
2: Fix the others and update J by

$$J = \arg\min \frac{1}{\mu}\|J\|_* + \frac{1}{2}\|J - (Z + Y_2/\mu)\|_F^2$$

3: Fix the others and update E by

$$E = \arg\min \frac{\lambda}{\mu}\|E\|_1 + \frac{1}{2}\|E - (X - XZ - DS + Y_1/\mu)\|_F^2$$

4: Fix the others and update L by

$$L = \arg\min \frac{1}{\mu}\|L\|_1 + \frac{1}{2}\|L - (S + Y_3/\mu)\|_F^2$$

5: Fix the others and update Z by

$$Z = (I + X^TX)^{-1}[X^TX - X^TE - X^TDS + J + (X^TY_1 - Y_2)/\mu]$$

6: Fix the others and update S by

$$S = (I + D^TD)^{-1}[D^TX - D^TE - D^TXZ + L + (D^TY_1 - Y_3)/\mu]$$

7: Update three Lagrange multipliers

$$Y_1 = Y_1 + \mu(X - XZ - E - DS), \quad Y_2 = Y_2 + \mu(Z - J), \quad Y_3 = Y_3 + \mu(S - L)$$

8: Update the parameter μ, $\mu = \min(\rho\mu, \mu_{\max})$
9: Check the convergence conditions

$$\|X - XZ - DS\|_\infty < \varepsilon, \quad \|Z - J\|_\infty < \varepsilon, \quad \|S - L\|_\infty < \varepsilon$$

10: **end while**;

The step 3 can be solved by Lemma 2:

Lemma 2 [22]. *The following type formula* $\min\ \varepsilon\|X\|_1 + \frac{1}{2}\|X - W\|_F^2$ *has only one analysis solution*

$$S_\varepsilon[W] = \begin{cases} w - \varepsilon & w > \varepsilon \\ w + \varepsilon & w < -\varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

When the noise is represented by $l_{2,1}$ norm, the step 3 can be solved by Lemma 3.

Lemma 3. *Let Q be a given matrix. If the optimal solution to*

$$\min\ \lambda\|W\|_{2,1} + \frac{1}{2}\|W - Q\|_F^2$$

is W^*, *then the i-th column of* W^* *is*

$$W^*(:,i) = \begin{cases} \frac{\|q_i\| - \lambda}{\|q_i\|} q_i & \lambda < \|q_i\| \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

3.3. Complexity and convergence

Assuming the observation data X and dictionary A have the same dimension $m \times n$, because of the involvement of singular value decomposition of $n \times n$ matrix in each iteration, then, the major computation of solution lies in Step 2. The complexity of Step 2 has lower bound $O(r_A^3)$, where r is the rank of A, when the dictionary A has an orthogonal basis. Thus, the complexity of Step 3 and 4 is the same as $O(mnr_A)$, whereas complexity of step 5 and 6 are $O(nr_A^2)$. Therefore, the complexity of the whole algorithm is $O(r_A^3 + nr_A^2 + mnr_A)$. The complexity of the algorithm has upper bound $O(n^3)$ that is the same as the complexity of LRR, while we use the observation data X as dictionary.

Inexact ALM algorithm, which involves two variable matrix iterating alternately, has been proven to converge [18]. However, convergence is difficult to prove because the algorithm in this paper has five iterating variable matrices and the objective of the optimization problem in Eq. (11) is not smooth. Fortunately, theoretical results in the literature [23] state that two conditions can guarantee convergence for Algorithm 1: The first condition is that dictionary matrix A is of full column rank; the second condition is that $\Delta_k = \|(Z_k, J_k) - \arg\min \mathcal{L}\|$ is monotonically decreasing, where $\arg\min \mathcal{L}$ is the optimal solution in the k-th iteration, $\arg\min \mathcal{L}$ denotes the ideal solution. According to the literature [18], the first condition is easy to follow, whereas the second condition can be ensured by convexity of the Lagrange function.

4. LRSR based small target detection

As mentioned earlier, an infrared image can be viewed as combination of three components, such as background, target, and noise. Thus, we decompose the original infrared image into three components images by using LRSR model. Then, the target location is determined by thresholding the target image. The whole method of small target detection based on LRSR is depicted in Fig. 2, and the steps of the detection method are as follows:

1. Image data reconstruction. Given that the small infrared target is represented by the over-complete dictionary D, we need to reconstruct the original image data $X \in R^{m \times n}$ to the matrix $\tilde{X} \in R^{k^2 \times w}$ by using a sliding window $k \times k$, which has the same size as the dictionary atomic from top and left to down and right with sliding step s. Each sub-block extracted by the sliding window was vectorized as a column of matrix $\tilde{X}$, where w is the number of the sub-blocks.
2. Dictionary generation. Two dictionaries in LRSR should be generated. One is used to describe the collection data of multi-subspaces denoted by A. The other is the over-complete dictionary used to represent the small target denoted by D. We usually use the original data itself as dictionary A, which has sufficient capability to determine the underlying structure of data. Moreover, choosing the over-complete dictionary for signal sparse representation is a basic and important problem. This selection also determines the quality of the sparse representation. In this paper, we utilize Gaussian Intensity Model, which is widely used in infrared small targets area to simulate small targets [9,24],

$$T(i,j) = I_{\max} \exp\left(-\frac{1}{2}\left[\frac{(i - x_0)^2}{\sigma_x^2} + \frac{(j - y_0)^2}{\sigma_y^2}\right]\right) \tag{16}$$

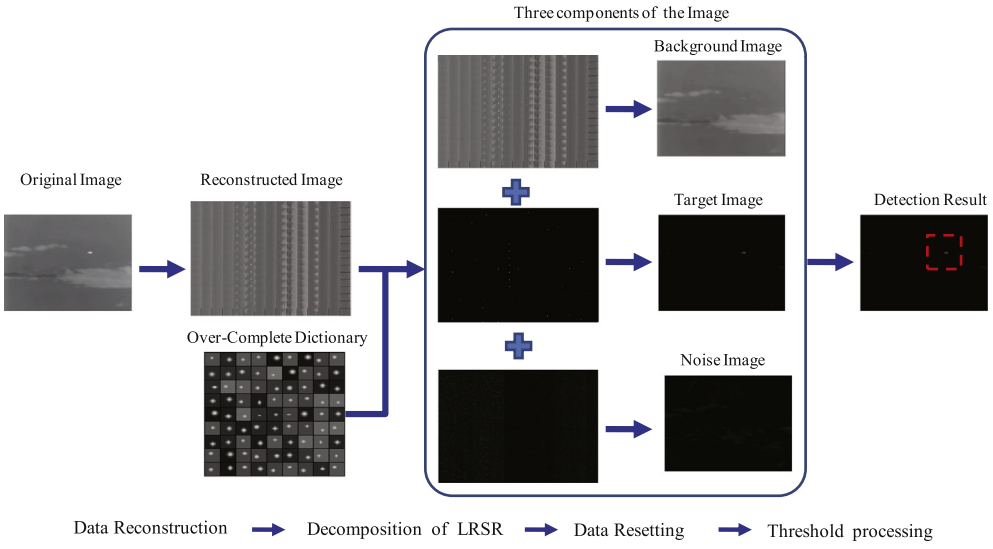

Fig. 2. Diagram of the small detection method based on LRSR.

where $I_{\max}$ is the intensity of the peak, (x_0, y_0) is the position of center, $T(i,j)$ is the intensity in the position (i,j), σ_x and σ_y is the standard deviations in row and column direction, respectively. We can obtain the small infrared target with different size, brightness and center position by adjusting the $I_{\max}, (x_0, y_0)$, as well as σ_x and σ_y. We have to restrict the relationship among the parameters to produce a realistic target. Thus, Eq. (16) can be rewritten by

$$T(i,j) = I_{\max} \exp\left(-\frac{1}{2}\left[\frac{(i-x_0)^2}{\sigma_x^2} + \frac{(j-y_0)^2}{\sigma_y^2}\right]\right), \quad \text{s.t.} \begin{cases} t_1 < \frac{I_{max}}{\sigma_x} < t_2 \\ t_1 < \frac{I_{max}}{\sigma_y} < t_2 \end{cases} \tag{17}$$

where t_1 and t_2 are the constant and $t_1 = 20, t_2 = 60$ in this study. Finally, we vectorize each target $(k \times k)$ generated by Eq. (17) as a column of dictionary $D = [d_1, d_2, d_3 \cdots] \in R^{k^2 \times v}$, v is the number of the atoms of D. The dictionary generated is illustrated in Fig. 3.

3. Image decomposing and locating of the target. Original infrared image X can be decomposed by using LRSR model with the reconstructed image data $\tilde{X}$ and the over-complete dictionary D as input to slove Eq. (12). After obtaining the background component AZ and real noise component $\tilde{E}$, we could locate the target by performing the following steps: First, we calculate the target component as $T = \tilde{X} - AZ - \tilde{E}$. Then, T is inverse transformed to obtain the real target image $\tilde{T}$. The same position pixel in the final target image has different values from the adjacent sub-blocks, because the sliding step is usually less than the size of the window. Thus, we adopt the calculation $v = median(x)$, where x is the different value from different sub-blocks having that have equal location. Finally, the target image $\tilde{T}$ can be located by threshold processing. If $\rho M \leqslant \tilde{T}(x,y)$, then we realize (x,y) as a target point, where M is the maximum value of $\tilde{T}$, and ρ is the threshold.

5. Experiments and analysis

5.1. Design of experiments

Three groups of experiments are designed to evaluate the target detection performance. The first group is single target detection experiments. Six real images with different backgrounds are decomposed, and the detection results are obtained directly using the proposed method. The second group of experiments is for the infrared sequences with some synthetic targets. We embed small synthetic targets, which are generated from five real targets, into the images chosen from four real sequences. Then, the detection

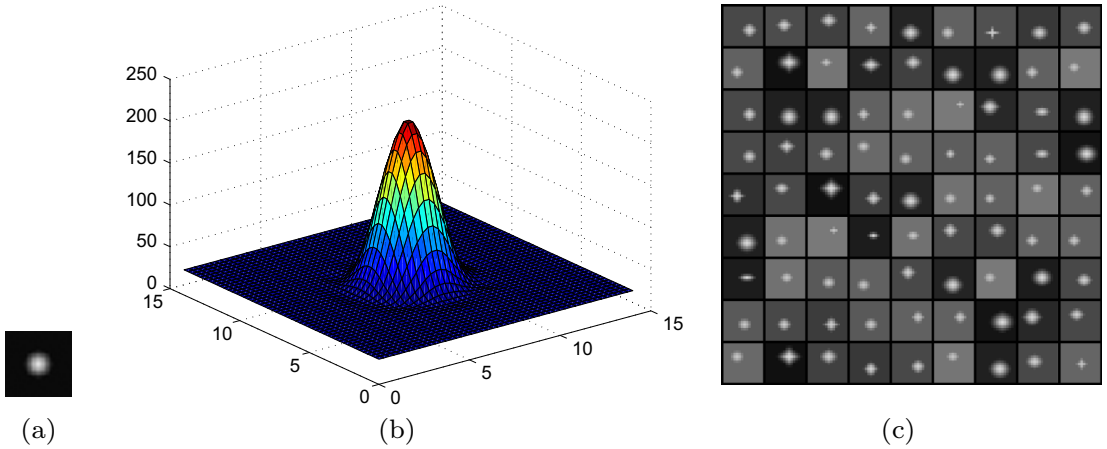

Fig. 3. Diagram of the over-complete target dictionary. (a) Sample of an atom, (b) 3-D mesh of an atom and (c) part of the over-complete dictionary.

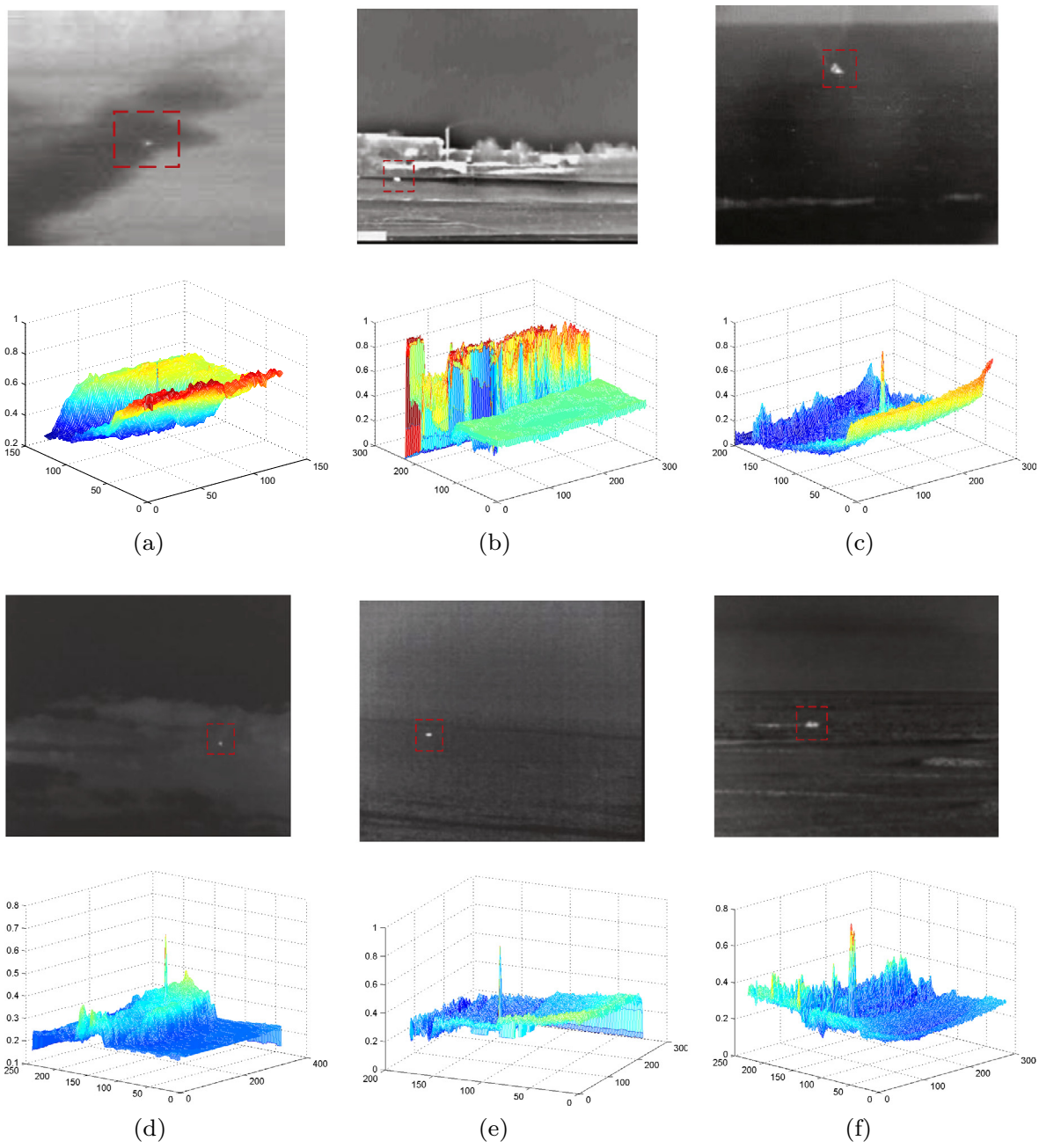

Fig. 4. Real infrared images with small dim targets against different backgrounds. (a) Rive-ground background, (b) riverbank background, (c) sea-sky background, (d) sea background, (e) sky-cloud background and (f) coast background.

performance is quantitatively evaluated by using some objective evaluation criteria, such as local SNR, detection probability, and false-alarm rate. We also test the conventional single-frame detection methods for comparison, such as TDLMS, Top-Hat, Max-Median, SLMD based on RPCA [19] and the detection method based on SR [11]. The third group of experiments evaluates the robustness to noise of the different methods. In these groups of experiments the size of sliding window and sliding step are experimentally determined as 16×16 and 10, respectively, based on computational cost consideration. The detection threshold is set to 0.85. However, the penalty parameters λ and β, controlling the tradeoff among the low-rank representation, noise component and the target component, are different for different conditions.

5.2. Experiment on single target detection

In this subsection, Fig. 4 shows the application of the proposed detection method to test six real infrared images with small dim target against different backgrounds. The six original infrared images are rive-ground background image with size 128×128, riverbank background with size 200×256, sea-sky background image with size 200×256, sea background with size 320×240, sky-cloud background with size 320×240 and coast background with size 320×240. The target detection experimental results on these six images by our method are shown in Fig. 5. From this figure, we can see the original infrared image are effectively decomposed into three different components, such as background

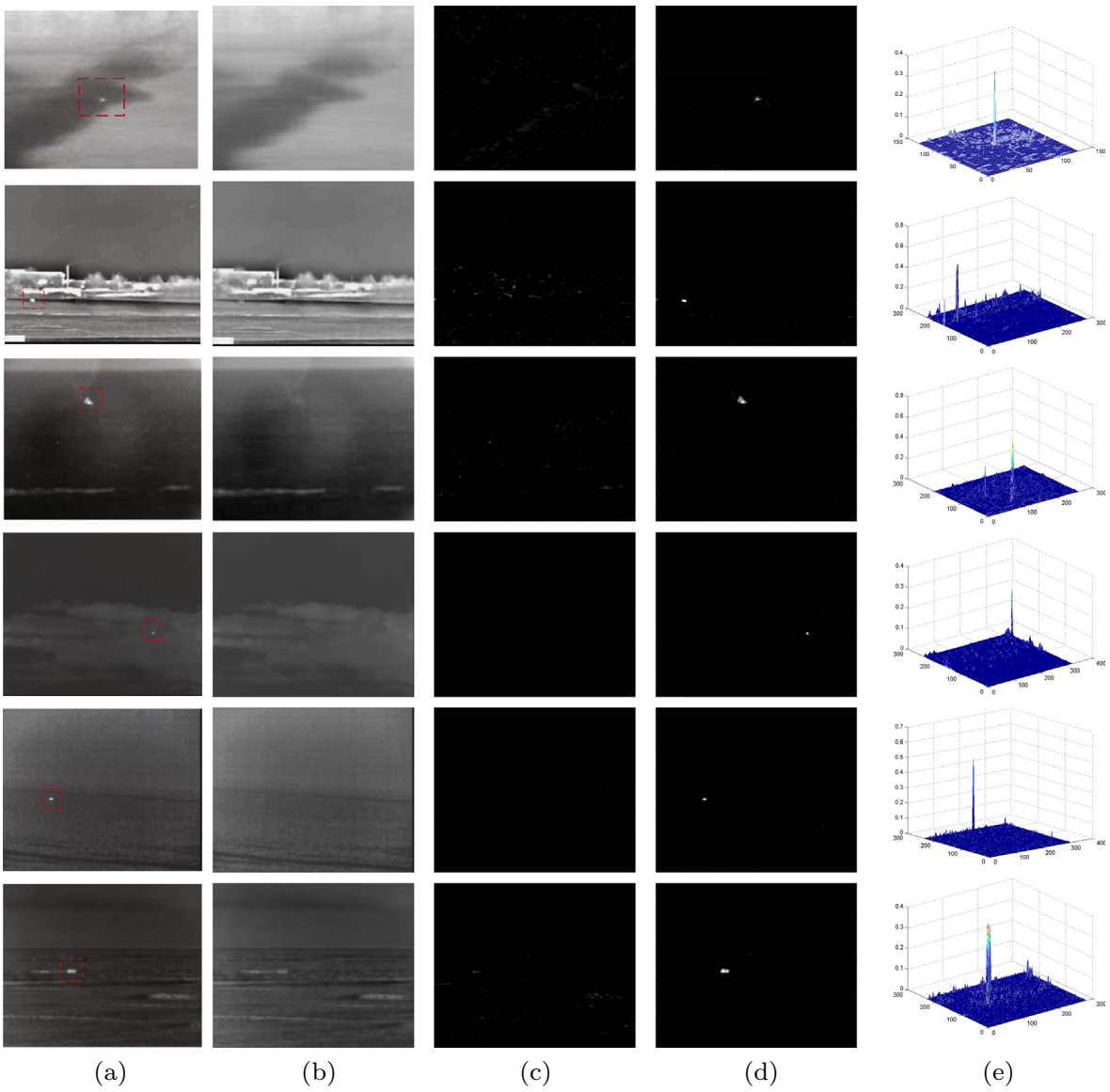

Fig. 5. Image decomposition results obtained by proposed algorithm. (a) Original image, (b) background data, (c) noise component, (d) target component and (e) 3-D mesh of target component.

component image shown in Fig. 5(b), noise component image shown in Fig. 5(c), and target component image show in Fig. 5(d). Fig. 5(e) shows the 3-D mesh surfaces of the target component image. The results indicate that our method can decompose the real infrared images into three components and effectively accomplish the task of target detection with different clutter backgrounds not only for the dim point targets but also for the porphyritic target such as the target in Fig. 4(c) and (f), through separating the target component from other elements.

5.3. Experiment on infrared sequences with multiple targets

5.3.1. Generation of experimental data

Multi-target detection experiments on four sequences are performed in this subsection to verify the performance of the proposed method more objectively. However, infrared small target detection has no test database available for public and widely used. Thus we synthesized the test database by embedding the multi-targets into several background images selected from four real

Fig. 6. The five real infrared target images.

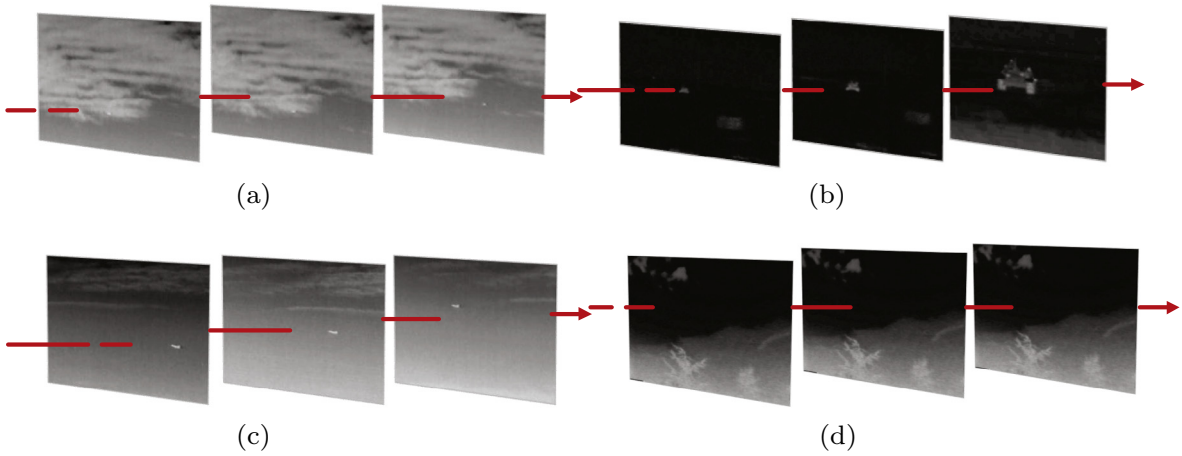

Fig. 7. The four real infrared image sequences. (a) Sequence 1, (b) sequence 2, (c) sequence 3 and (d) sequence 4.

sequences with different clutters. The synthesized targets are obtained by resizing five real infrared targets extracted from Fig. 6. The representative images for four sequences are shown in Fig. 7. The background of each sequence from left to right is thick cloud sky, ground road, thin cloud sky and trees sway with the wind.

The synthesizing process for the test images is as follows: First, the five real small targets extracted from Fig. 6 were formed target group r, where the original targets sizes are $3\times4, 3\times5, 8\times6, 8\times10, 5\times5$, respectively. Second, the number of target k was determined randomly, $k\in[2,10]$ for every image chosen from each sequence. Then, we chose k targets random and repeatable from r denoted by $t_1, t_2 \ldots t_k$ and resized each target t_i by using bicubic interpolation method to obtain the new synthesized target T_i with size of $am\times an, am\times an\in[4,90]$, where a is Random decimal, $a\in(0,3]$, m and n are the original size of small targets. Third, the synthesized images can be achieved by embedding the new targets $T_1, T_2 \ldots T_k$ into the real images chosen from background sequences. The rule of embedding is demonstrated in Eq. (18):

$$I(x,y)=\begin{cases}\max(rT(x-x_0,y-y_0),B(x,y)) & x\in(1+x_0,an+x_0),y\in(1+y_0,am+y_0)\\ B(x,y) & \text{otherwise}\end{cases} \tag{18}$$

where $B(x,y)$ is the background image, (x_0,y_0) is random location that indicates the left upper corner of the simulated target in $B(x,y)$. r is produced randomly, $r\in[h,255]$, h is the maximum pixel value of $B(x,y)$. Finally, the synthesized infrared image was blurred by Gaussian filter to produce realistic image [14].

Four image sequences that contain 1063 small targets are synthesized by using the previous simulation method and all images could be divided into four groups according to the original sequences. The details of four groups of synthetic images are shown in Table 1. As seen in Fig. 8, the simulated targets have different sizes, gray, and local signal to noise ratio (LSNR), which can reflect the real characteristic of infrared small target. Thus, the synthetic images are credible to be used as test data set.

Table 1
Details of four groups of synthetic images.

23	Group 1	Group 2	Group 3	Group 4
Number of images	50	30	30	100
Number of targets	235	153	141	534
Average pixels of target	21.4	32.7	45.8	43.5

5.3.2. Experimental comparison

In this subsection, we compare our method with several typical algorithms, such as Max-Median, Top-Hat, TDLMS, SLMD based on RPCA in the literature [19] and the algorithm based on SR in the literature [11] by using local signal to noise ratio (LSNR), local signal to noise ratio gain (LSNRG) and the receiver operating characteristic (ROC) to be as the objective metrics, where the ROC curve is the curve of detection probability p_d and false-alarm probability F_a. The larger area under the curve, the better the detection performance is.

$$p_d=\frac{N_r}{N_T} \tag{19}$$

$$F_a=\frac{N_w}{N_i} \tag{20}$$

where N_r is the number of true detections, N_T is the number of actual targets, N_w is the number of false detections, and N_i is the number of images. LSNR is as follows:

$$\mathrm{LSNR}=\frac{P_T}{P_B} \tag{21}$$

where P_T is the max target intensity and P_B is the maximum pixel value in neighboring area around the target. The size of the neighboring area around target is twice the size of the target in this paper. Generally speaking, the higher the LSNR of a target is, the greater the contrast between the target and the background in local area. These characteristics can make the detection easy and effective.

$$\mathrm{LSNRG}=\frac{\mathrm{LSNR}_{out}}{\mathrm{LSNR}_{in}} \tag{22}$$

where LSNR_{in} and LSNR_{out} stand for the LSNR values before and after process by our method in local area. The bigger the LSNRG is, the better performance for LSNR gain the algorithm has.

Fig. 9 demonstrates the comparisons between the proposed method and the classical algorithms. In this figure, the first column illustrates the representative images for four groups of synthetic image sequences. Moreover, Fig. 9(a2) and (a4) have heavy noise and low SNR, particularly in Fig. 9(a4), wherein a distant target cannot be observed clearly. Fig. 9(a2) and (a3) presents tank and aircraft targets with big size, whereas Fig. 9(a2) also contains the road and building targets. Fig. 9(a1) and (a3) have sky background. However, the former has strong cloud clutter and several target location are in the clouds and sky junction. As depicted in the second row of Fig. 9, the proposed method can detect correctly not only four flight targets in the sky but also the tank target on the

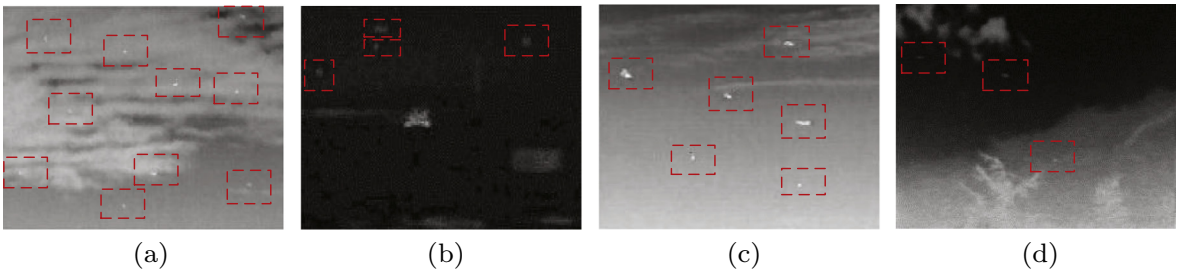

Fig. 8. The representative images for four groups of synthetic image. (a) Group 1, (b) group 2, (c) group 3 and (d) group 4.

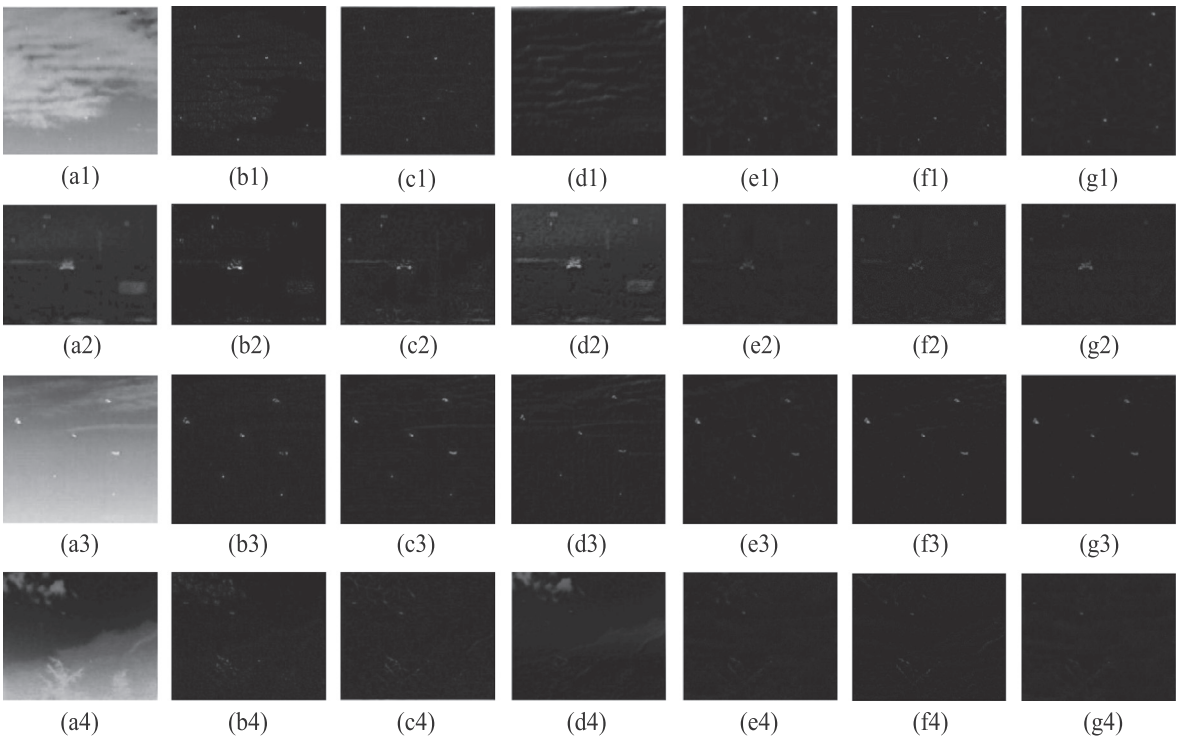

Fig. 9. Results of different methods. (a1)–(a4) The representative images for four groups of synthetic images, (b1)–(b4) max-median with filter size of 9×9, (c1)–(c4) top-hat with filter size of 12×12, (d1)–(d4) TDLMS with parameter $\mu = 10^{-6}$, (e1)–(e4) SR, (f1)–(f4) SLMD with $\lambda = 0.07$ and (g1)–(g4) LRSR with $\lambda = 0.10$, $\beta = 0.05$, $\lambda = 0.07$, $\beta = 0.01$, $\lambda = 0.08$, $\beta = 0.03$ and $\lambda = 0.07$, $\beta = 0.05$ for the four groups, respectively.

road which illustrates that our method also has detection capability to spot target with large pixels. Meanwhile, the road and building, which are obviously not targets, are eliminated effectively. The third row of Fig. 9 presents that SR and LRSR algorithms can effectively maintain the contour information of the targets. Thus, these two algorithms can be further used for target recognition. However, SR method has more clutter because it has weak capability to background modeling. In addition, SLMD method based on RPCA has the similar subjective detection performance with SR method with less noise residual. However, because it does not possess the capability to separate noise and target, SLMD method has inferior detection capability to our LRSR method. In general, the detection method based on LRSR has less clutter and less noise residual for different backgrounds compared with that of the other baseline methods with the target enhanced and more using information preserved simultaneously.

The ROC curves of the six methods for four image sequences are depicted in Fig. 10. The curves are worse in Group 2 and Group 4 compared with that of Group 1 and Group 3 because of the low contrast and serious noise. We also can see the detection performance of the SLMD method is in-between the SR method and LRSR method. From Group 1, Group 2 and Group 4, SR method has slightly better performance when the $F_a < 0.5$ compared with that of the other methods. However, the proposed method displays the highest detection performance than that of the others with the same false-alarm rate. This finding indicates that the performance of our method is more stable for different clutter backgrounds compared with that of the other methods.

Table 2 demonstrates the comparison results of detection probability with the same false-alarm rate ($Fa = 2/image$), the average LSNR and the average LSNRG of six algorithms. The average LSNR of four original image sequences is 1.84, 1.42, 1.97 and 1.28. From Table 2, we can clearly see that the proposed method has better performance than baseline methods not only in terms of detection probability but also in terms of average LSNR and LSNRG. Table 2 also demonstrates SLMD method, which presents the similar objective detection performances with SR method in three evaluation indexes, has superior detection capability over other conventional methods, and the TDLMS method has the worst detection result among the baseline methods.

Detection experiments on two infrared images without target are performed to further illustrate the effectiveness of our

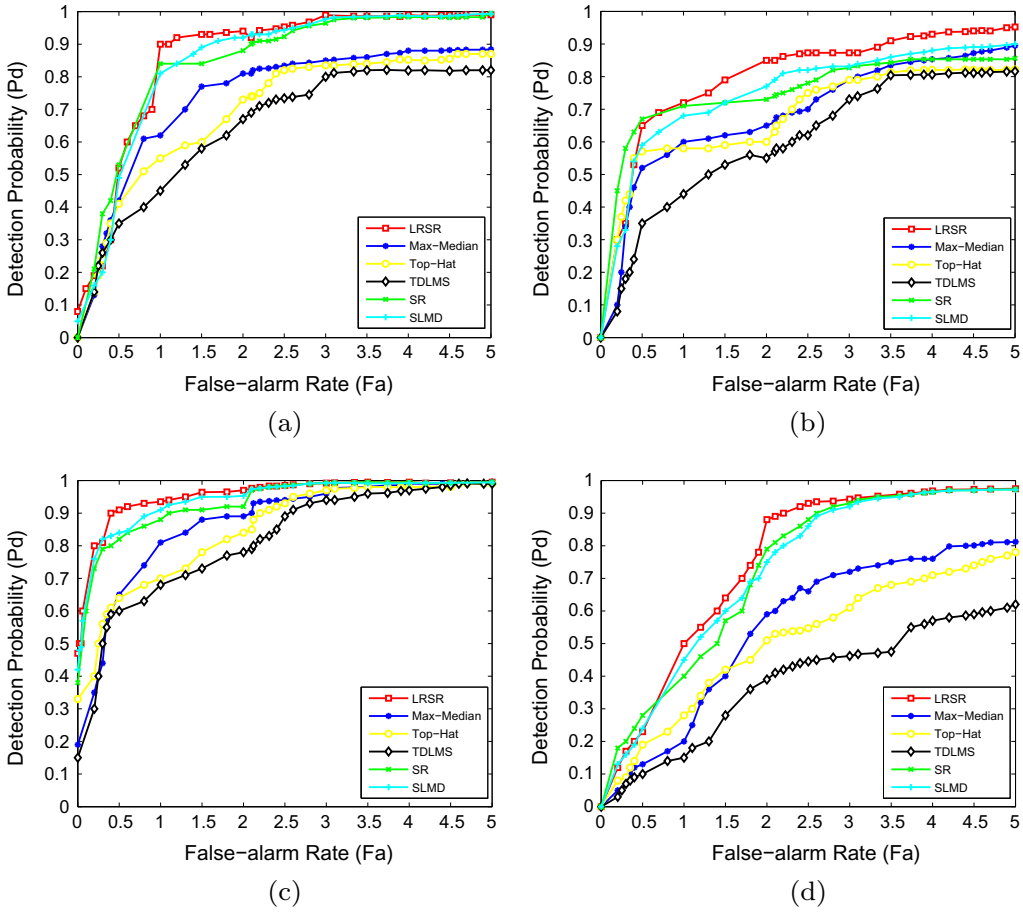

Fig. 10. ROC curves of detection results of four groups image. (a) Group 1, (b) group 2, (c) group 3 and (d) group 4.

Table 2
The details of four groups of synthetic images.

	Group 1	Group 2	Group 3	Group 4
P_d				
Max-Median	0.81	0.65	0.89	0.59
Top-Hat	0.73	0.60	0.84	0.51
TDLMS	0.67	0.55	0.77	0.39
SR	0.88	0.73	0.92	0.79
SLMD	0.90	0.79	0.95	0.76
LRSR	**0.94**	**0.85**	**0.97**	**0.88**
Average of LSNR				
Max-Median	3.37	1.86	3.49	1.73
Top-Hat	3.07	1.74	3.17	1.59
TDLMS	2.63	1.63	3.01	1.50
SR	5.99	3.09	6.40	3.30
SLMD	6.21	3.06	6.92	3.35
LRSR	**8.63**	**3.66**	**8.45**	**4.42**
Average of LSNRG				
Max-Median	1.83	1.41	1.77	1.35
Top-Hat	1.67	1.32	1.61	1.24
TDLMS	1.43	1.23	1.53	1.17
SR	3.26	2.34	3.25	2.58
SLMD	3.85	2.17	3.69	2.64
LRSR	**4.69**	**2.77**	**4.29**	**3.45**

Bold fonts indicate the best performance.

proposed method. In this case, the less noise the result has, the lower the false alarm rate is, and the detection performance is better. The comparison results are shown as Fig. 11.

When there is no small infrared target in the image, the sparse coefficients S and target component DS are zero ideally for minimizing $\|Z\|_* + \lambda\|\tilde{E}\|_l + \beta\|S\|_1$. So our proposed model falls back to Eq. (4), and noise component is concentrated in $\tilde{E}$. But in fact, our model could move some noise which can be represented by the target dictionary into the target component. Even so, the detection result of our method has fewer clusters than the baseline methods. In Fig. 11, the first column shows two original infrared images, in which there are swimming pool and campus road respectively, without any target. From the first row, it can be seen that the other four methods could detect the edge of pool except the SR and LRSR methods. This is because these two methods have the ability to describe small target, they also can eliminate some pixels such as the pool edge which are not target obviously. However due to the weakness of background modeling, the result of SR method has more noise than the result of LRSR method has. From the second row, we can see the last three methods have superior detection capability to the previous three methods. In general, as the experimental results shown, our proposed model could work well and still maintains good performance when there is no target in the image.

5.4. Experiments on the robustness of noise

We added salt and pepper noise with density of 0.02 to the first two groups, whereas Gaussian noise with standard deviation $\sigma = 0.005$ and mean $\mu = 0$ was added to the last two groups to test the robustness of the proposed method to the noise that may appear in the images. Moreover, we compared the false-alarm rate with $P_d = 0.7$, average LSNR and average LSNRG of six methods. The average LSNR of four original image sequences with noise

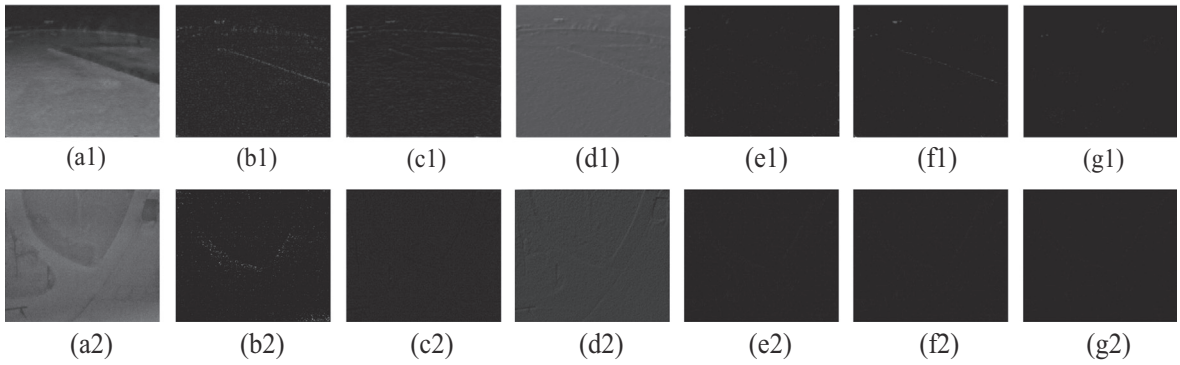

Fig. 11. Results of different detection methods for images without target. (a1)–(a2) The two images without targets, (b1)–(b2) Max-Median with filter size of 9×9, (c1)–(c2) Top-Hat with filter size of 12×12, (d1)–(d2) TDLMS with parameter $\mu = 10^{-6}$, (e1)–(e2) SR, (f1)–(f2) SLMD with $\lambda = 0.07$ and (g1)–(g2) LRSR with $\lambda = 0.07$, $\beta = 0.01$, $\lambda = 0.08$, $\beta = 0.03$ for the two images, respectively.

Table 3
The performance comparison of different methods with different noise.

	Group 1	Group 2	Group 3	Group 4
P_d				
Max-Median	2.1	2.8	1.8	5.8
Top-Hat	2.7	3.2	2.1	6.0
TDLMS	3.6	4.0	2.6	7.8
SR	1.5	1.9	1.2	4.6
SLMD	1.3	2.3	0.98	4.2
LRSR	**1.1**	**1.4**	**0.9**	**3.1**
Average of LSNR				
Max-Median	2.35	1.20	2.01	0.75
Top-Hat	1.95	1.10	1.78	0.81
TDLMS	1.76	0.98	1.57	0.70
SR	3.89	1.54	3.30	0.97
SLMD	4.72	1.51	4.18	1.04
LRSR	**6.01**	**1.95**	**5.08**	**1.55**
Average of LSNRG				
Max-Median	1.69	1.35	1.64	1.14
Top-Hat	1.43	1.24	1.45	1.21
TDLMS	1.29	1.10	1.28	1.07
SR	2.84	1.69	2.69	2.10
SLMD	3.24	1.71	3.26	2.35
LRSR	**4.39**	**2.27**	**4.13**	**3.37**

Bold fonts indicate the best performance.

are 1.37, 0.89, 1.23 and 0.66. The results of comparison are depicted as Table 3.

Table 3 demonstrates that each indicator decreases in number because of the different noise. However, the proposed method still has higher LSNR and LSNRG with low false-alarm rate compared with those of the other algorithms. In addition, the average LSNRG of LRSR method has the smallest decreased proportion among all the algorithms. This finding demonstrates that the capability of our method to improve of LSNR is slightly influenced by noise. Consequently, our method has better robustness to noise than baseline methods because the former employs noise component modeling and can represent the noise correctly through the selection of the appropriate model. In addition, the proposed method also inherits the capability of correcting the corruptions automatically of LRR, which is also an important contributor to noise robustness.

6. Conclusion and future work

In this paper, a novel model called LRSR is presented based on low-rank representation combining the sparse representation theory. On the basis of the LRSR model, we proposed a small target detection method which can transform the detection task into separation process of background, noise and target components by solving LRSR. The results of the three groups of experiments validated that the proposed method has better detection performance and greater robustness to different noises than conventional baseline methods. In future works, the research attempt would be to investigate the two dictionary generations in the LRSR model. Such investigation may contribute to the improvement of the detection performance of the proposed detection method or may extend the application range of the LRSR model with the use of different dictionaries.

Conflict of interest

There is no conflict of interest.

Acknowledgments

This work is supported by National Natural Science Foundation of China (Grant No. 61102170). We thank Liu for providing the source codes of LRR. We are also very grateful to the editor and anonymous reviewers for their constructive comments and suggestions that help improve the quality of this manuscript.

References

[1] C. Zheng, H. Li, Small infrared target detection based on harmonic and sparse matrix decomposition, Opt. Eng. 52 (2013). 066401(1–10).
[2] C. Gao, T. Zhang, Q. Li, Small infrared target detection using sparse ring representation, IEEE Trans. Aerospace Electron. Syst. Mag. 27 (2012) 21–30.
[3] M. Zeng, J. Li, Z. Peng, The design of top-hat morphological filter and application to infrared target detection, J. Infrared Phys. Technol. 48 (2006) 67–76.
[4] P. Wang, J. Tian, C. Gao, Infrared small target detection using directional highpass filters based on ls-svm, Electron. Lett. 45 (2009) 156–158.
[5] Y. Cao, R. Liu, J. Yang, Small target detection using two-dimensional least mean square (tdlms) filter based on neighborhood analysis, Int. J. Infrared Millimeter Waves 29 (2008) 188–200.
[6] S.D. Deshpande, H.E. Meng, R. Venkateswarlu, P. Chan, Max-mean and max-median filters for detection of small targets, in: SPIE's International Symposium on Optical Science, Engineering, and Instrumentation, International Society for Optics and Photonics, 1999, pp. 74–83.
[7] T. Hu, J. Zhao, Y. Cao, F.-L. Wang, J. Yang, Infrared small target detection based on saliency and principle component analysis, J. Infrared Millimeter Waves 29 (2010) 303–306.
[8] Z. Liu, X. Zou, C. Chen, X. Shen, Detection of small objects in image data based on the nonlinear principal component analysis neural network, Opt. Eng. 44 (2005). 093604(1–9).
[9] Y. Cao, R. Liu, J. Yang, Infrared small target detection using ppca, Int. J. Infrared Millimeter Waves 29 (2008) 385–395.
[10] R. Liu, H. Zhi, Infrared point target detection with Fisher linear discriminant and kernel Fisher linear discriminant, J. Infrared, Millimeter, Terahertz Waves 31 (2010) 1491–1502.
[11] J. Zhao, Z. Tang, J. Yang, Infrared small target detection based on image sparse representation, J. Infrared Millimeter Wave 30 (2011) 156–161.

[12] X. Shao, H. Fan, G. Lu, J. Xu, An improved infrared dim and small target detection algorithm based on the contrast mechanism of human visual system, J. Infrared Phys. Technol. 55 (2012) 403–408.
[13] S. Kim, Y. Yang, J. Lee, Y. Park, Small target detection utilizing robust methods of the human visual system for irst, J. Infrared, Millimeter, Terahertz Waves 30 (2009) 994–1011.
[14] C. Gao, M. Deyu, Y. Yang, Y. Wang, X. Zhou, A. Hauptman, Infrared patch-image model for small target detection in a single image, IEEE Trans. Image Process. 22 (2013) 4996–5009.
[15] Z. Chen, S. Luo, T. Xie, et al., A novel infrared small target detection method based on bemd and local inverse entropy, J. Infrared Phys. Technol. 66 (2014) 114–124.
[16] Z. Chen, T. Deng, L. Gao, et al., A novel spatial–temporal detection method of dim infrared moving small target, J. Infrared Phys. Technol. 66 (2014) 84–96.
[17] 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.
[18] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, Y. Ma, Robust recovery of subspace structures by low-rank representation, IEEE Trans. Pattern Anal. Mach. Intell. 35 (2013) 171–184.
[19] H.L.C.Y. Zheng, Small infrared target detection based on low-rank and sparse matrix decomposition, Appl. Mech. Mater. 239 (2012) 214–218.
[20] J. Wright, A. Yang, A. Ganesh, S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Trans. Pattern Anal. Mach. Intell. 31 (2009) 210–227.
[21] J. Cai, E.J. Cands, Z. Shen, A singular value thresholding algorithm for matrix completion, IEEE Trans. Pattern Anal. Mach. Intell. 20 (2008) 1956–1982.
[22] Z. Lin, M. Chen, Y. Ma, The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-rank Matrices, arXiv preprint arXiv:1009.5055.
[23] J. Eckstein, D. Bertsekas, On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math. Program 55 (1992) 293–318.
[24] J. Zhao, Z. Tang, J. Yang, et al., Infrared small target detection using sparse representation, J. Syst. Eng. Electron. 22 (2011) 897–904.