

# Adaptive kernel regression and energy concentration criterion for infrared dim small target detection

Mingyang Ma,<sup>a,b,c</sup> Dejiang Wang,<sup>a,c,\*</sup> He Sun,<sup>a,c</sup> and Tao Zhang<sup>a,c</sup>

<sup>a</sup>Chinese Academy of Sciences, Changchun Institute of Optics, Fine Mechanics and Physics, Changchun, China

<sup>b</sup>University of Chinese Academy of Sciences, Beijing, China

<sup>c</sup>Chinese Academy of Sciences, Key Laboratory of Airborne Optical Imaging and Measurement, Changchun, China

**Abstract.** It is always a challenging task to detect a small target with low signal-to-noise ratio under complex background in infrared images. To address this problem, an effective algorithm based on background subtraction is proposed. First, we add the gradient feature into the kernel regression model to acquire an edge-preserving background estimation. The smoothing matrix of the kernel function is reestablished by a rotation angle and an elongation scale. Further, a multi-scale first-order directional derivative filter is presented to calculate these factors adaptively. Second, to segment the real target from the subtracted image, we model the imaging process of the small target using the point spread function of the optical system. According to the analysis of the imaging size and the energy distribution of target, an energy concentration criterion is constructed and used for target extraction. Finally, comparison of experimental results demonstrates that the proposed algorithm achieves robust performances on background suppression and extracts the target accurately with a high detection probability and low false alarm rate. © 2021 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.OE.60.12.123101](https://doi.org/10.1117/1.OE.60.12.123101)]

**Keywords:** infrared dim small target detection; kernel regression; first-order directional derivative; energy concentration criterion.

Paper 20210549 received May 25, 2021; accepted for publication Oct. 11, 2021; published online Dec. 1, 2021.

## 1 Introduction

Infrared (IR) small-target detection with a far distance plays a significant role in many applications, such as precise guidance,<sup>1</sup> infrared search and tracking systems (IRST),<sup>2</sup> and remote sensing.<sup>3</sup> The measurements required from IR images are obtained from the difference of natural IR radiant between target and background. If the detection system is equipped with an angular sensor and a range finder, the measurement will also be augmented by the angle and range information. In general, the images always contain clutter due to several factors, such as the sources of target and background radiation, transmission properties of atmosphere, and limitations of IR detector.<sup>4,5</sup> Especially, when these factors get worse, the density of clutter will increase dramatically. This clutter will bring many false alarms to the detection results. Thus, detecting dim small targets in complex background is a necessary capability for a robust IR detection system. Although many detection algorithms have been developed in recent decades, the requirement of high detection probability and low false alarm rate is still not reached.<sup>3,6</sup> The fundamental problems mainly include three aspects. (1) The background edge. In general, there are always many luminous edges in the complex background. Due to the obvious gradient variation, this clutter is difficult to suppress and is retained in the detection results.<sup>7</sup> (2) The dimness characteristic of small target. Due to the long imaging distance, the small target captured by the infrared focal plane array is a weak light spot. Consequently, it has a low signal-to-noise ratio (SNR) and lacks available features to be utilized for detection.<sup>8,9</sup> (3) Disturbance from the detector noise. This internal noise generated by the IR detector exists widely in IR images.

---

\*Address all correspondence to Dejiang Wang, [wangdj04@ciomp.ac.cn](mailto:wangdj04@ciomp.ac.cn)

Because it shares the similar sparse characteristic to the small target, it will confuse the detection process and bring many false alarms.<sup>10</sup>

Over the past few decades, many algorithms on infrared small-target detection have been reported, which can be broadly grouped into three categories: human vision system (HVS) method,<sup>11,12</sup> background estimation method,<sup>13,14</sup> and data structure method.<sup>15,16</sup> The HVS method makes use of the local difference principle to distinguish the actual target from complex background.<sup>17</sup> Some researchers define the local contrast feature to detect small target, such as local contrast measurement (LCM)<sup>18</sup> and its improved versions.<sup>19–22</sup> For example, Chen et al.<sup>23</sup> added the accommodative scales and weighted map into the LCM to restrain various background. Han et al.<sup>24</sup> proposed a ratio-difference joint LCM to enhance the small target salience in the complex background. Wang et al.<sup>25</sup> combined the LCM and high-boost filter to achieve high-speed local contrast. Besides, some algorithms put the difference filter into the infrared small-target detection.<sup>26</sup> Wang et al.<sup>27</sup> combined the temporal filtering and difference of Gaussian operator to improve the detection accuracy. Kim proposed<sup>28</sup> min-local-LoG filter to overcome the sensitivity to edge clutter. However, the detection results of these algorithms are easily affected by the heavy clutter and noise. In addition, the morphological filtering method such as top-hat transform is another widely used tool to detect small target.<sup>29,30</sup> Comparing with the aforementioned filters, it has a better anti-interference ability and fits the multiscale feature of targets.<sup>31,32</sup> Unfortunately, they rely on the priori target information, such as size and shape. And the detection results are sensitive to the background edge. In recent years, some researchers pay attention to the gradient feature of small target and propose the facet model for target detection.<sup>33</sup> Based on facet model, Cao et al.<sup>34</sup> designed a detection algorithm using the discontinuity of small target and derivative dissimilarity measure. Bi et al.<sup>35</sup> utilized the multiorder directional derivatives to suppress background and enhance the target in each derivative subband. Although these algorithms achieve good performance, false alarms exist more or less. The advantage of HSV-based method is simple operation. However, when the clutter and noise are heavy, they are more dependent on the postprocessing.

The second category is designed based on the background estimation. The small target is detected by subtracting the estimated background from the original image. The detection result of this category depends on the accuracy of background estimation. Existing algorithms mainly include two approaches. One is the least mean squares (LMS) filter proposed by Hadhoud and Thomas.<sup>36</sup> It predicts the background image by minimizing the difference between the original image and the desired output image. Subsequently, some improvements based on LMS are carried out. Cao et al.<sup>37</sup> integrated the consistency principle of neighboring pixels into the LMS filter to improve the estimation precision. Bae et al.<sup>38</sup> proposed the edge directional two-dimensional (2D) LMS filter to suppress the background edge. However, these filters rely on the step size and the correctness of the reference image. The other approach is the statistical regression. In this kind of approach, the most successful of the IR dim small-target detection methods is kernel regression (KR).<sup>39,40</sup> It establishes the background estimation model using a polynomial constructed by the gray values of neighboring pixels. The background is estimated as the zero-order term of the polynomial.<sup>41</sup> Unlike the LMS filter, this approach is a nonparametric regression algorithm. It only relies on the data itself and does not need a reference image. However, KR filter is sensitive to the background edge, which will decrease the background estimation performance and increase the false alarm rate.

The third category concentrates on the data structure. In this category, the image is regarded as a linear superposition of the target, background, and noise, and the small target is separated from backgrounds by optimizing objective functions.<sup>42</sup> The robust principal component analysis (RPCA) is one of the popular ideas.<sup>43</sup> Based on this method, Gao et al.<sup>44</sup> proposed infrared patch-image model. Subsequently, Dai et al.<sup>45</sup> weighted each column in the patch image to suppress the complex background in the IR image. These algorithms do not obtain a better suppression performance for the background edge. To overcome this issue, Wang et al.<sup>46</sup> proposed the total variation regularization and principal component pursuit method. Meanwhile, to achieve more robust detection, the low-rank recovery and sparse representation methods are introduced to infrared small target detection.<sup>47,48</sup> However, these algorithms rely on the construction of objective models and the robust optimization methods. In addition, some methods convert the detection problem into a binary classification problem, such as dictionary representation,<sup>49</sup> neural

works,<sup>50</sup> and support vector machine.<sup>51</sup> Although the detection performance is good, due to excessive dependence on training samples and label selection, these methods are unsuitable to practical applications.

Putting the emphasis on both background suppression and target extraction, a detection algorithm of IR dim small target based on background subtraction category is proposed in this paper. First, to improve the estimation accuracy of background edges, we modify the classic KR method and propose an adaptive kernel function. The smoothing matrix of the kernel function is rewritten as a product form by a rotation matrix and an elongation matrix. Then, a filter model of multiscale first-order directional derivative (MFODD) is constructed to estimate two matrices by minimizing the directional derivative of samples. Second, to distinguish the target and noise, we model the target image using the PSF of optical system. After analyzing the energy distribution and imaging size of small target on IR image, an energy concentration criterion is derived and performed to extract the actual target from the subtracted image. Comparing with some existing state-of-the-art algorithms, the proposed algorithm achieves good detection performances with a high detection probability and a low false alarm rate.

## 2 Adaptive KR Model for Background Removal

In this section, we formulate the classic kernel estimator and provide its weaknesses on the choice of smoothing matrix and drawback on the background edge estimation. Further, to address these problems, a new KR function that can adapt to the shape and orientation of backgrounds is proposed. The smoothing matrix is rewritten as a product form by a rotation matrix and an elongation matrix. Then, an MFODD filtering model is established to estimate two matrices.

### 2.1 Classic KR for Background Estimation

The KR model of image data is written as follows:

$$I_i = z(\chi_i) + \varepsilon_i, \quad i = 1, 2, \dots, N, \tag{1}$$

where  $\chi_i = [x_i, y_i]^T$  is a  $2 \times 1$  vector of the spatial coordinates of a sample pixel,  $I_i$  is the image gray of a sample pixel,  $i$  is the index of the sample pixel,  $z(\cdot)$  is the regression function,  $\varepsilon_i$  is the random noise, which is supposed to be a Gaussian variable with zero mean, and  $N$  is the total number of sample pixels in the detection window.

Although the expression of  $z(\cdot)$  remains unspecified, under the assumption that the regression function is locally smooth to some order  $D$ , we can utilize a generic local expansion at any sample point  $\chi_i$  to approximate  $z(\chi)$ . This relationship can be described by the  $D$ -order Taylor series as follows (we omit the high order expansion for simplicity):

$$z(\chi_i) \approx \xi_0 + \xi_1^T(\chi_i - \chi) + \frac{1}{2}(\chi_i - \chi)^T \xi_2(\chi_i - \chi) + \dots, \tag{2}$$

where  $\xi_0 = z(\chi)$  is the gray value of pixel to be estimated,  $\xi_1$  and  $\xi_2$  are vectors of the first and second order of  $z(\chi)$ , respectively, which can be calculated as

$$\xi_1 = \nabla z(\chi) = \left[ \frac{\partial z(\chi)}{\partial x}, \frac{\partial z(\chi)}{\partial y} \right]^T, \tag{3}$$

$$\xi_2 = \nabla^2 z(\chi) = \frac{1}{2} \left[ \frac{\partial^2 z(\chi)}{\partial x^2}, 2 \frac{\partial^2 z(\chi)}{\partial x \partial y}, \frac{\partial^2 z(\chi)}{\partial y^2} \right]^T, \tag{4}$$

where  $\nabla$  and  $\nabla^2$  denote the gradient and Hessian operators, respectively.

$\{\xi_d\}_{d=0}^D$  can be derived by minimizing the random noise in Eq. (1). According to the least-square theory, the estimation of  $\{\xi_d\}_{d=0}^D$  is established by the following optimization model:

$$\min_{\{\xi_d\}} \sum_{i=1}^N \varepsilon_i = \min_{\{\xi_d\}} \sum_{i=1}^N [I_i - \xi_0 - \xi_1^T(\chi_i - \chi) - \xi_2^T \times \text{vech}\{(\chi_i - \chi)(\chi_i - \chi)^T\} - \dots]^2 K_G(\chi_i - \chi), \tag{5}$$

where  $K_G(\cdot)$  is the kernel function that is used to assign the weights for samples,  $G$  is the  $2 \times 2$  smoothing matrix.  $\text{vech}(\cdot)$  is defined as the half-vectorization operator of the ‘‘up-triangular’’ portion of a symmetric matrix, as follows:

$$\text{vech}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = [a, b, d]^T. \tag{6}$$

This optimization problem can be described as a weighted least-squares problem. In fact, we just focus on the gray value of the pixels. So, it is only necessary to derive  $\xi_0$  from Eq. (5) and the local gradient  $\{\xi_d\}_{d=2}^D$  can be ignored in the estimation. Regardless of the high order term of the estimator, the estimation of background is simplified as

$$\hat{z}(\chi) = \hat{\xi}_0 = e_1^T (\chi_x^T W_x \chi_x)^{-1} \chi_x^T W_x I, \tag{7}$$

where  $e_1$  is a column vector, in which the first element is 1 and other elements are equal to 0.  $I = [I_1, I_2, \dots, I_N]^T$ ,  $W_x = \text{diag}[K_G(\chi_1 - \chi), K_G(\chi_2 - \chi), \dots, K_G(\chi_N - \chi)]$ ,

$$\chi_x = \begin{bmatrix} 1 & (\chi_1 - \chi)^T & \text{vech}^T\{(\chi_1 - \chi)(\chi_1 - \chi)^T\} \\ 1 & (\chi_2 - \chi)^T & \text{vech}^T\{(\chi_2 - \chi)(\chi_2 - \chi)^T\} \\ \vdots & \vdots & \vdots \\ 1 & (\chi_N - \chi)^T & \text{vech}^T\{(\chi_N - \chi)(\chi_N - \chi)^T\} \end{bmatrix}.$$

A common and intuitive choice of  $K_G(\cdot)$  is to use spatial distance between the estimated pixel and samples. Based on the principle of radiometric distance, the weights can be assigned using the following Gaussian kernel function:

$$K_{G_i}(\chi_i - \chi) = \frac{1}{2\pi \det(G_i)} \exp\{-(\chi_i - \chi)^T G_i^{-1} (\chi_i - \chi)\}, \tag{8}$$

where  $G_i$  is the covariance matrix of the kernel function of  $i$ 'th sample pixel, which is also defined as the smoothing matrix. It controls the shape of the kernel function and further decides the quantity of samples. In other words, the choice of the smoothing matrix determines the estimation performance of KR.

In the previous works, the smoothing matrix is usually established by a simplified model, as follows:<sup>28</sup>

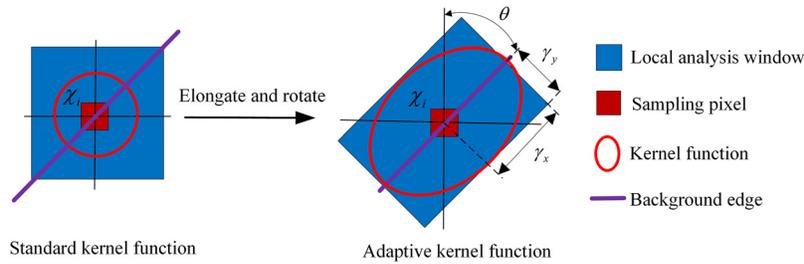
$$G_i = h\mu_i E, \tag{9}$$

where  $h$  is a global smoothing parameter,  $\mu_i$  is a scalar parameter,  $E$  is a  $2 \times 2$  identity matrix.

Research often concentrates on the optimization algorithm to derive the aforementioned parameters. However, under the model of Eq. (9), the smoothing matrix is a diagonal matrix. Because the elements of  $G_i$  on subdiagonal are zero, the row and column direction of kernel function is irrelevant. In other words, this classic model is not suitable for background edges with multidirection characteristics. Furthermore, the background cannot be estimated well.

## 2.2 Smoothing Matrix Selection Based on Multiscale First-Order Directional Derivative

To obtain a better background estimation, the kernel function should adapt to the direction and scale of background edge. With this intuition in mind, we add the gradient feature and the scale of samples into the smoothing matrix. According to the geometric interpretation of the covariance, the smoothing matrix can be rewritten as<sup>52</sup>



**Fig. 1** Schematic representation illustrating the effects of the improved smoothing matrix on the scale and direction of the kernel function.

$$G = R\Upsilon R^T, \tag{10}$$

where  $G$  is the abbreviation of  $G_i$ ,  $R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$  and  $\Upsilon = \begin{bmatrix} \gamma_x & 0 \\ 0 & \gamma_y \end{bmatrix}$   $\theta$  is the rotation factor, which is defined as the clockwise angle from the vertical axis in a normal coordinate system,  $\gamma = [\gamma_x, \gamma_y]$  represents the elongation factors on the horizontal and vertical directions. These two factors of  $\theta$  and  $\gamma$  denote the scale and direction of the kernel function.

Figure 1 schematically shows how factors affect the spreading of kernels on a background edge. First, the kernel rotates with the gradient direction until reaching the dominant orientation of background edge. Then, the kernel elongates the maximum scale according to the length of the background edge. In the final result, the factors of  $\theta$  and  $\gamma$  represent the orientation of background edge and the corresponding energy of the direction, respectively. Due to the addition of  $\theta$  and  $\gamma$ , the kernel function will get a better estimation on the background edge.

To estimate the parameters of  $\theta$  and  $\gamma$ , an MFODD filter is constructed based on the facet model. It is modeled as follows:

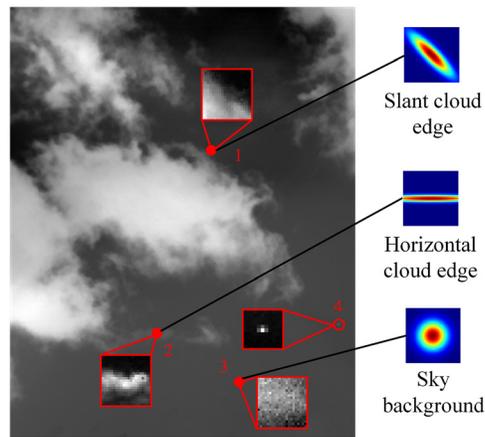
$$\frac{\partial I_i}{\partial \vec{l}_{\theta,\gamma}} = \{(K_{2,\gamma} - 3.4K_{7,\gamma} - 2K_{9,\gamma}) \sin \theta + (K_{3,\gamma} - 3.4K_{10,\gamma} - 2K_{8,\gamma}) \cos \theta\} \otimes f, \tag{11}$$

where  $f$  is the original IR image,  $\otimes$  is the convolution operator,  $\frac{\partial I_i}{\partial \vec{l}_{\theta,\gamma}}$  is the directional derivative of the sample pixel gray  $I_i$  on the direction  $\theta_i$  and scale  $\gamma_i$ ,  $\gamma_i$  is the abbreviation of  $[\gamma_{i,x}, \gamma_{i,y}]$ .  $K_{j,\gamma}$  is the fitting coefficient under scale  $\gamma$  in the facet model, which can be obtained as

$$K_{j,\gamma} = \frac{P_j(r, c)}{\sum_{c' \in \Omega_{\gamma_y}} \sum_{r' \in \Omega_{\gamma_x}} P_j^2(r', c')}, \quad j = 1, 2, \dots, 10, \tag{12}$$

where  $P_j(r, c)$  is the two-dimensionality discrete orthogonal polynomials, which is constructed as  $\{1, r, c, r^2 - 2, rc, r^2 - 2, r^3 - 3.4r, (r^2 - 2)c, (c^2 - 2)r, c^3 - 3.4c\}$ ,  $[r, c]^T \in \Omega_{\gamma_x} \times \Omega_{\gamma_y}$  is the coordinates of pixels in the local analysis window,  $\Omega_{\gamma_x} \times \Omega_{\gamma_y}$  represents the combination of the coordinate set of rows or columns under scale  $[\gamma_x, \gamma_y]$ . Let the location of the current sample pixel is  $(0, 0)$ , the symmetric set  $\Omega_{\gamma_x}$  and  $\Omega_{\gamma_y}$  can be defined as  $\Omega_s = \{-2s, -2s + 1, \dots, 2s - 1, 2s\}$ , where  $s$  is the scale index. Defining  $S$  as the upper scale limit of the analysis window, the set of scale index on a single direction is denoted as  $\Theta = \{1, \dots, S\}$ . Hence,  $\Omega_{\gamma_x} \times \Omega_{\gamma_y}$  has  $S^2$  combinations and the scale  $\gamma$  of the smoothing matrix has  $S^2$  cases.

In the MFODD model,  $\theta_i$  denotes the dominate direction of the sample pixel  $i$  and  $\gamma_i$  reflects the number of neighboring pixels who are similar to the sample pixel on this side. Taking Fig. 1 as an example, the sample is located on the background edge. When the direction is perpendicular to the maximal gray gradient of the sample and the scale is equal to the length of the edge, the kernel function gets a good estimation on background. At this moment, the absolute value of the MFODD filtering gets the minimum response. With this intuition in mind, the estimation of  $\theta_i$  and  $\gamma_i$  is represented as follows:



**Fig. 2** Example of adaptive kernel under different background.

$$[\hat{\theta}_i, \hat{\gamma}_i] = \arg \min \left| \frac{\partial I_i}{\partial \vec{l}_{\theta, \gamma}} \right| \quad \gamma \in \Theta_x \times \Theta_y, \quad \theta \in \Xi, \quad (13)$$

where  $\Xi$  is the angle set of  $\theta$ . In the target detection, we mainly concentrate on the direction of 0 deg, 45 deg, 90 deg,  $-45$  deg. So, only these four angles are considered in the rotation factor estimation of the proposed adaptive kernel function.  $\gamma$  is used to describe the energy of the sample on the orientation. So, the value of  $\gamma$  should be the maximum target scale  $S$ .

Figure 2 is a visual illustration of the adaptive kernel function on a complex cloud background. The size of the real target located in region 4 is  $2 \times 3$  pixels. So, in the MFODD filtering, we set  $S = 3$ . (This setting will be further analyzed in Sec. 3.1.) In region 1, the direction of the cloud edge in the local area is slant. So,  $\gamma = 1$  and  $\theta = -45$  deg can be estimated from Eq. (13). Analogously, due to the horizontal and narrow shape of cloud in region 2, the smoothing matrix will obtain a large scale and a rotation factor with 90 deg. When the background is clean such as region 3, a standard Gaussian formation can be estimated by MFODD filter. Both the local enlargements and the kernel functions of three regions are shown in Fig. 2.

Submitting the estimation results of the MFODD filter into Eq. (7), the background can be estimated as  $\hat{f}_B(x, y) = \hat{z}(\chi)$ . Further, a subtraction operation is adopted to segment the target image and remove the clutter and noise:

$$f'(x, y) = f(x, y) - \hat{f}_B(x, y). \quad (14)$$

In the subtracted image  $f'(x, y)$ , the target becomes salient and the background clutter is extremely suppressed. As for some residual clutter that is not completely eliminated, a global threshold can be used for removal.

### 3 Energy Concentration Criterion for Target Extraction

After a subtraction operation, a “pure” target image is obtained. However, some isolated noises such as the detector noise will also remain in the subtracted image. Due to sharing the similar feature of bright and small size with the target, it is difficult to eliminate in the background removal and will increase the false alarm rate. To extract the actual target accurately, an energy concentration criterion is proposed in this section.

#### 3.1 Imaging Size and Energy Distribution of Infrared Dim Small Target

The linear theory of image formation, related to Gaussian optics, allows expressing the imaging process as a convolution product between the radiance distribution on the object plane and the

point spread function (PSF).<sup>53</sup> To obtain the image, the energy captured by the optical system is integrated and sampled by the detector. This whole imaging process can be described as

$$f_T(x, y) = \zeta[g(x, y) \otimes \text{PSF}], \quad (15)$$

where  $g(x, y)$  is the real target in the object space, PSF is the point spread function of optical system,  $\zeta[\cdot]$  is the sampling function of IR detector, which represents the process of discretization and sampling, and  $f_T(x, y)$  is the image of target.

With the imaging distance of the target farther away, the real target  $g(x, y)$  is approximately a point. Thus, the target image  $f_T(x, y)$  can be described by the PSF of the optical system. The ideal PSF of the optical system is represented as a first-order Bessel function

$$\text{PSF} = \left[ \frac{2J_1(\pi\rho/\lambda F)}{\pi\rho/\lambda F} \right]^2, \quad (16)$$

where  $J_1$  is the first-order Bessel function,  $\lambda$  is the response wavelength of the detector,  $F$  is the optical parameter, which is defined as the ratio between the focal length and the aperture of optical system,  $\rho = \sqrt{(x - x_0)^2 + (y - y_0)^2}$  is the radius of the detector,  $(x_0, y_0)$  is the imaging location of target center on the detector. For example,  $(x_0, y_0) = (0, 0)$  represents that the target locates on the pixel center of IR detector.

Based on Eqs. (15) and (16), we first analyze the imaging size of the target. The distribution of PSF on the 2D image plane is defined as Airy spot. Let  $Z$  denote  $\frac{\pi\rho}{\lambda F}$ . According to the character of Bessel function, we can note that when  $Z = 1.22\pi$ , it gets the minimal value  $J_1(Z) = 0$ . Further, the diameter of the Airy spot can be derived as follows:

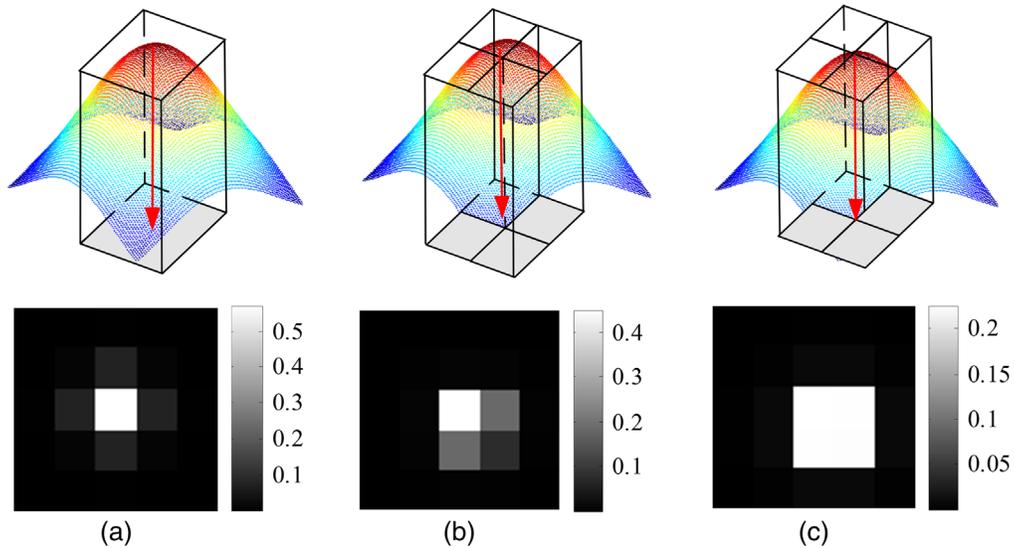
$$d = 2.44\lambda F, \quad (17)$$

where  $d$  is the diameter of Airy spot.

In the real world, the long-wave IR detection system is usually adopted to capture the long distance target.<sup>54</sup> Submitting some common parameters of  $\lambda = 9$  to  $10 \mu\text{m}$  and  $F = 2$  into Eq. (16), the size of Airy spot can be derived as  $43.9$  to  $48.8 \mu\text{m}$ . In addition, Eq. (15) shows that the target image is the discrete sampling result of Airy spot by the IR detector. It is assumed that the shape of the detector element is rectangle. According to the common pixel sizes of long-wave IR detector ( $15, 25, 30 \mu\text{m}$ ), we can note that the target size will change from  $1 \times 1$  to  $3 \times 3$  pixels. In other words, the maximum target scale is no more than  $S = 3$ .

Second, we analyze the energy distribution of target. Equation (16) indicates that the energy of PSF is related to the location of  $(x_0, y_0)$ . Thus, the energy distribution of target is affected by the projected location of target center. Next, we adopt the simulations to analysis the energy distribution of small target under different projected location. Assuming that the pixel size is  $25 \mu\text{m}$  and the deviations of the row and column direction keep coincident. The result is shown in Fig. 3. The first row is the three-dimensional (3D) schematic diagrams describing the projected location of the small target center on the pixel. And the second row is the corresponding energy distributions in the 2D space. In Fig. 3(a), the target is projected on the pixel center. In Figs. 3(b) and 3(c), the projected center target is  $8$  and  $12 \mu\text{m}$  away from the pixel center, respectively.

Figure 3(a) shows that although the target is projected on the pixel center, the energy does not completely concentrate on a single pixel but diffuses to neighboring pixels. The reason is that the diameter of Airy spot exceeds the pixel size of detector, which has been proved by the above analysis about target size. From Figs. 3(b) and 3(c), it can be noted that with the increasing of deviations between target center and pixel center, the energy of the central pixel decreases gradually. On the contrary, the neighboring pixels get higher energy. When the target is projected at the intersection of four pixels, the energy distribution of the discrete target is extremely uniform. Based on the simulation, we can draw the conclusion that the target energy captured by IR detector will disperse on the image. And this property can be used as an important judgment for distinguishing the actual target and false alarms in the subtracted image.



**Fig. 3** Examples of target image for several deviations between the target center and the pixel center: (a)  $0 \mu\text{m}$ , (b)  $8 \mu\text{m}$ , and (c)  $12 \mu\text{m}$ .

### 3.2 Energy Concentration Criterion Construction

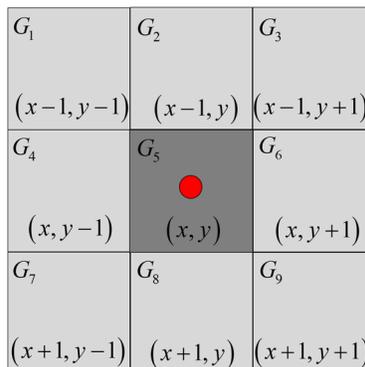
The aforementioned analysis denotes that the imaging size of target will be no more than  $3 \times 3$  pixels. With this intuition in mind, we construct the filtering window as shown in Fig. 4. The detected pixel  $(x, y)$  is the center pixel of this window.  $G_1 \sim G_9$  represent the energies of pixels in the filtering window.

From the last section, we have noted that the target energy distribution is affected by the projected location of a small target center on the detector. When the target projects at the pixel center, such as Fig. 3(a),  $G_5$  of pixel  $(x, y)$  is strong. When the projected center deviates from the pixel center, such as the condition of Figs. 3(b) and 3(c),  $G_1 \sim G_9$  of neighboring pixels will rise gradually with the increasing of the deviation distance. According to this diffusion property, an energy concentration criterion is constructed as follows:

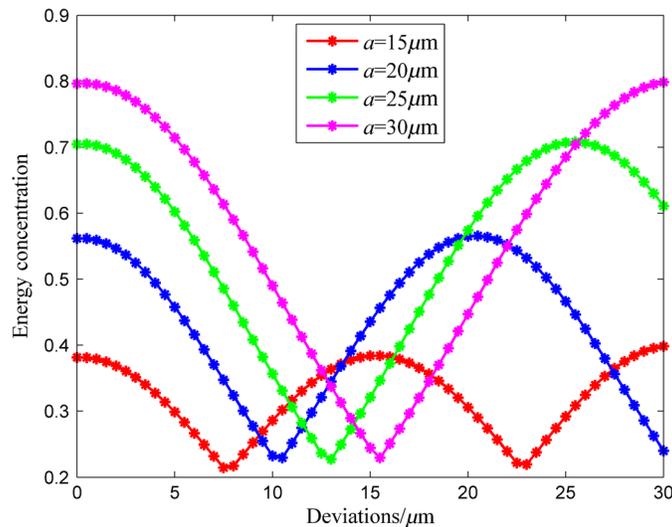
$$G_e = \frac{G_{\max}}{\sum_{i=1}^9 G_i}, \tag{18}$$

where  $G_e$  represents the energy concentration of detected pixel,  $G_{\max}$  is the maximum energy in nine pixels, and the denominator term is the sum of energies of these nine pixels.

According to the definition in Eq. (18), the energy concentration map of the whole image can be obtained from subtracted image by the following calculation:



**Fig. 4** Filtering window of the energy concentration criterion.



**Fig. 5** Target energy concentrations under different deviations between target center and pixel center.

$$G_e(x, y) = \frac{\max\{f'(x + i, y + j)\}}{\sum_{i=1}^9 f'(x + i, y + j)}, \quad (i, j) \in L_m, \tag{19}$$

where  $L_m = \{-1, 0, 1\}$  is the direction vectors.

To extract the actual target in the energy concentration map, it is necessary to determine the variation range of target energy concentration. Hence, we simulate the energy concentration with different deviations between target center and pixel center under some common pixel sizes. The results are plotted in Fig. 5.  $a$  represents the value of pixel size.

We can note that the energy concentration of target changes periodically with the increasing of the deviations of target center. The period is approximated to the pixel size. In addition, Fig. 5 also indicates that when the deviation is equal to the half of the pixel size, the energy concentration reaches the lower limit. Besides, the lower limit is independent of the pixel size and is kept at 0.22. Unlike the lower limit, the upper limit of energy concentration depends on the location and pixel size simultaneously. When the target exactly projects at the pixel center, the energy concentration reaches the maximum value. Further, the peaks of curves between different pixel sizes in Fig. 5 indicate that when the pixel size becomes larger, the target will obtain a higher energy. It illustrates that enlarging the pixel size of detector property is helpful to improve the target salience.

In this paper, we denote the variation range of energy concentration as  $[G_{th1}, G_{th2}]$ . Setting it as the segment thresholds, the actual target can be extracted by the following binary operation:

$$\begin{cases} \text{Target pixel,} & \text{if } G_{th1} \leq G_e(x, y) \leq G_{th2} \\ \text{Nontarget pixel,} & \text{otherwise} \end{cases} \tag{20}$$

According to the proposed energy concentration criterion, the detector noise and other isolated noises can be eliminated and the actual target will be extracted from the subtracted image.

## 4 Experimental Results

### 4.1 Data Acquisition Equipment

To obtain the actual and reliable image resources, a prototypeIRST system is set up to capture the airliners leaving from Changchun Longjia International Airport, which is shown in Fig. 6. The IR camera is mounted on a two-axis motion platform. The pan and tilt axes of the platform could be changed continuously through the serial port. The ADS-B global navigation system

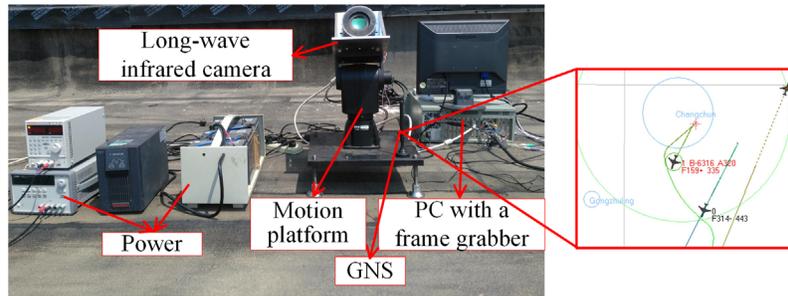


Fig. 6 Data acquisition equipment.

Table 1 Detailed parameters of the IR camera.

Style name	Brief description
Manufacturer	Sofradir
Spectral band pass ( $\mu\text{m}$ )	7.7 to 11.3
Element size ( $\mu\text{m}$ )	$320 \times 256$
Pixel size	30
NETD (mK)	17.32
ADC resolution (bit)	14
Focal length (mm)	38
$F/\#$	2
Integration ( $\mu\text{s}$ )	300

(GNS) is employed to acquire the flight information, including the speed, location, and altitude of the aircraft. The screenshot of GNS is also shown in Fig. 6.

The specific parameters of the long-wave IR camera are presented in Table 1. Submitting the parameters into Eq. (17), we can note that the target size is no more than  $3 \times 3$  pixels. Thus, we set  $S = 3$  in the adaptive KR model. In addition, because the pixel size of our detector is  $30 \mu\text{m}$ ,  $[G_{th1}, G_{th2}] = [0.22, 0.8]$  is set as the segment threshold of the energy concentration criterion in the target extraction.

#### 4.2 Performance Evaluation

In the experimental step, two aspects are adopted to evaluate our proposed method: background suppression performance and detection performance. The signal-to-clutter ratio (SCR) and background suppression factor (BSF) are used as the filter metrics for background suppression, which are defined as<sup>55</sup>

$$\begin{cases} \text{SCR} = \frac{\mu_t - \mu_c}{\sigma_c} \\ \text{BSF} = \frac{\sigma_{in}}{\sigma_{out}} \end{cases}, \quad (21)$$

where  $\mu_t$  and  $\mu_c$  represent the mean values of target region and background region,  $\sigma_c$  represents the stand deviation of the background region,  $\sigma_{in}$  and  $\sigma_{out}$  denote the standard deviation (STD) of the original input image and the STD of the filtered image, respectively.

In addition, receiver operating characteristic (ROC) curve is adopted to evaluate the detection performances of different algorithms. It is used to describe the dynamic relationship between the

detection probability and false alarm rate. In this paper, we totally adopt two types of ROC, which are defined, respectively, as follows:<sup>35,56</sup>

$$\begin{cases} P_{d,1} = \frac{N_{\text{true}}}{N_{\text{total}}} \times 100\% \\ P_{fa,1} = \frac{N_{\text{false}}}{N_{\text{false}} + N_{\text{total}}} \times 100\% \end{cases}, \quad (22)$$

$$\begin{cases} P_{d,2} = \frac{M_{\text{true}}}{M_{\text{total}}} \times 100\% \\ P_{fa,2} = \frac{M_{\text{false}}}{M_{\text{sum}}} \times 100\% \end{cases}, \quad (23)$$

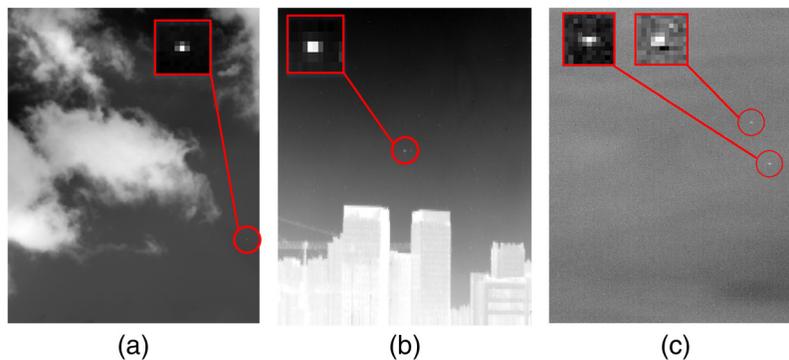
where  $N_{\text{true}}$  and  $N_{\text{false}}$  are the number of the detected true targets and false alarms, respectively,  $N_{\text{total}}$  is the total number of targets.  $M_{\text{true}}$  and  $M_{\text{false}}$  are true and false pixels detected,  $M_{\text{total}}$  is total true pixels, and  $M_{\text{sum}}$  is total pixels in image.

### 4.3 Detection Performance and Comparison

We totally obtain three IR sequences of small targets using the constructed equipment. Three representative images from these sequences are shown in Fig. 7. The type of targets is the civil aircraft. Seqs. 1 and 2 contain cloud and building clutter on sky backgrounds. Seq. 3 contains various detector noise and nonuniformity noise. The small target regions are located and magnified in original images for clearer observation. The target sizes are listed in Table 2. For comparison purpose, other five methods are performed, including Top-hat, LCM, RPCA, LMS, and classic KR.

First, we verify the ability of proposed algorithm on background suppression using the three tested sequences. The results of different algorithms are shown in the form of 3D gray maps in Figs. 8–10. The gray value has been normalized. The target locations are marked by red circles.

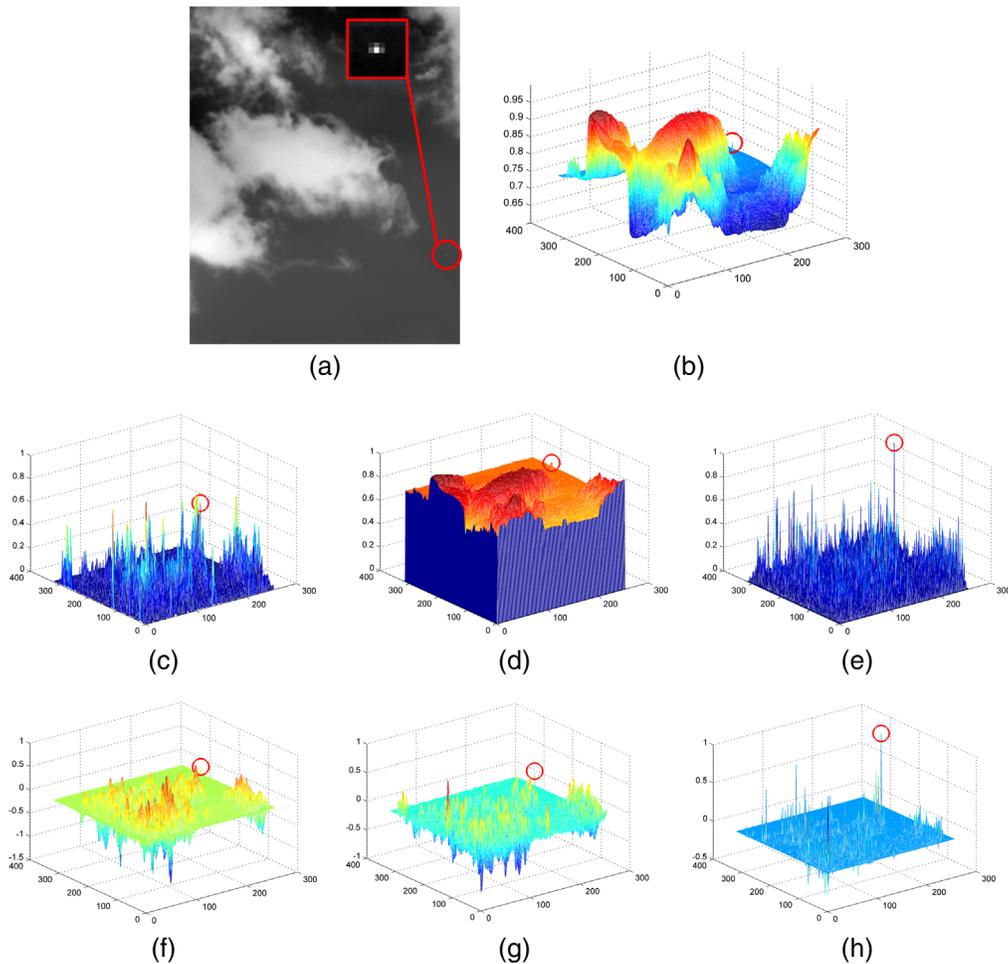
Figures 8(a) and 8(b) show that Seq. 1 contains numerous complex cloud backgrounds. Under this condition, the HVS-based methods including top-hat and LCM have a good performance on target enhancement. However, the cloud backgrounds are not suppressed effectively.



**Fig. 7** Samples on three IR sequences. (a) The raw image of Seq. 1, (b) the raw image of Seq. 2, and (c) the raw image of Seq. 3.

**Table 2** Sizes of targets in Fig. 7.

	Fig. 7(c1)			
	Fig. 7(a1)	Fig. 7(b1)	Target 1	Target2
Target size (pixels)	2 × 3	2 × 2	2 × 3	3 × 3

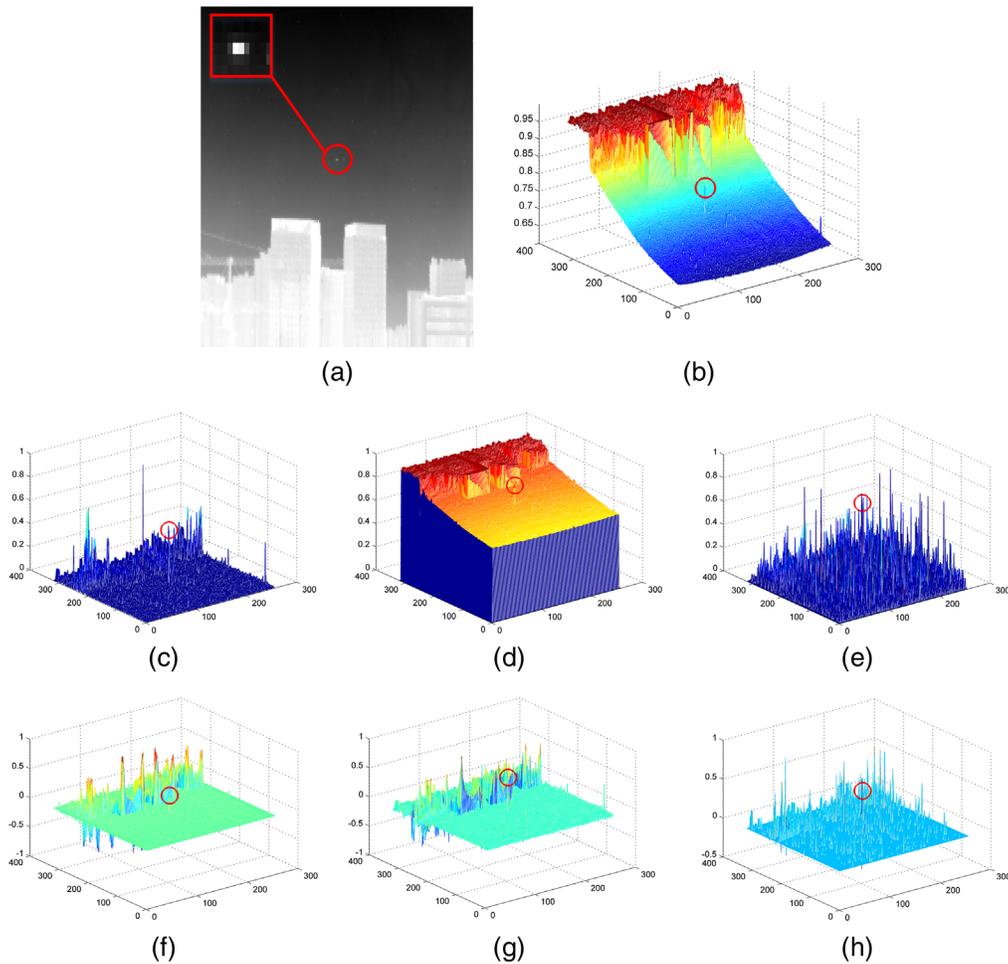


**Fig. 8** Results of background suppression processed by different algorithms for Seq. 1. (a) The raw image sample of Seq. 1, (b) the 3D gray map of sample image, (c) top-hat, (d) LCM, (e) RPCA, (f) LMS, (g) classic KR, and (h) proposed adaptive KR.

Similarly, as shown in Fig. 8(e), the results of RPCA also contain many residual clutter. Figures 8(f)–8(h) show that although the LMS and classic KR achieve better performances on cloud background than top-hat and LCM, they are sensitive to the background edge. By contrast, after being processed by the proposed algorithm, the cloud clutter is suppressed completely, and the target becomes the most salient in the image.

As shown in Figs. 9(a) and 9(b), the bright buildings have strong radiation intensity in Seq. 2. Its edge regions bring crucial challenges to the background suppression. Under such complex backgrounds, the suppression results of top-hat, LCM, and RPCA contain obvious residuals on the regions of building edges and corner points. Furthermore, the target is almost submerged by residual clutter and even lost. Although the classic KR achieves a better performance than other algorithms, it is sensitive to the building edges. By contrast, the proposed algorithm maintains its robustness to the building background.

Figures 10(a) and 10(b) show that the targets in Seq. 3 are almost swamped by massive detector noises. The sparse property of this noise is similar to the small target, which is easily to result in false alarms. As shown in Figs. 10(c) and 10(d), top-hat and LCM not only enhance the small target but also increase the noise intensity. Due to the existence of detector noise, the low-rank property of the background in Seq. 3 is not obvious, and a precise desired image can not be obtained. Thus, both RPCA and LMS do not get good suppression performances. Although the classic KR can suppress the noise clutter, the enhancement performance of small target is still inferior to the proposed algorithm.



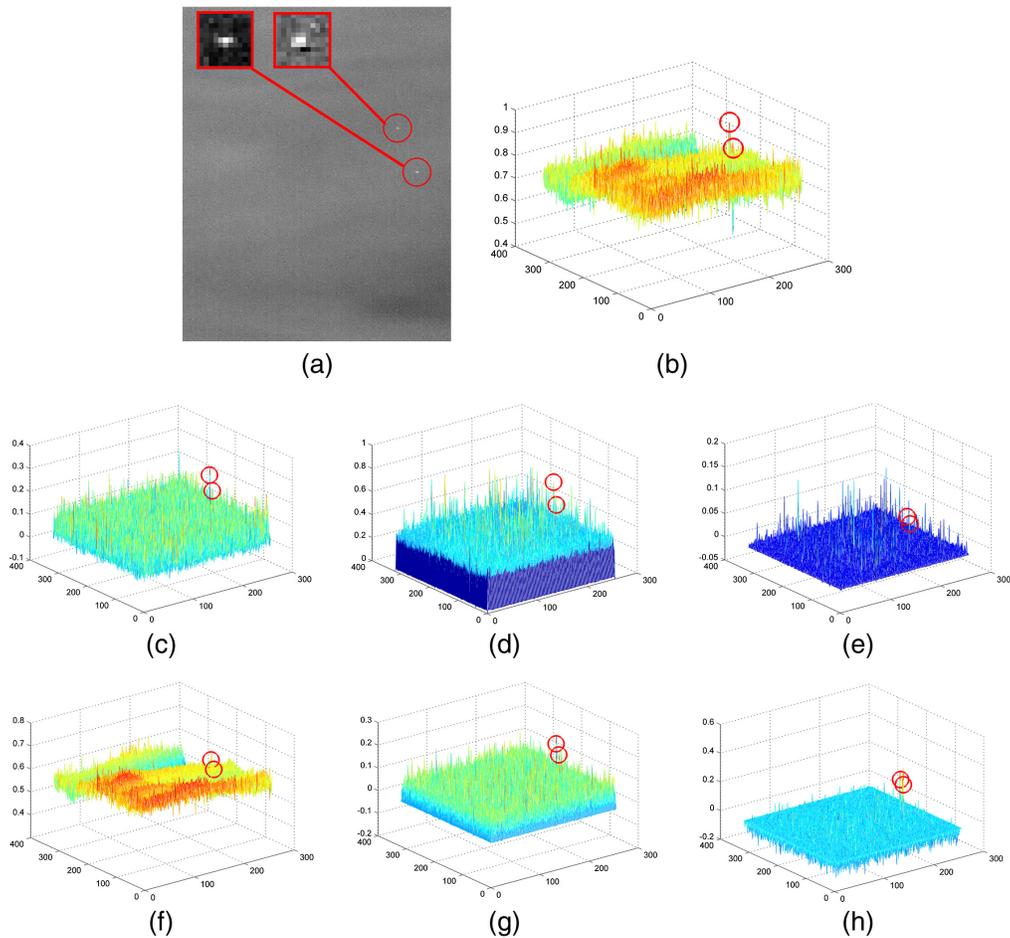
**Fig. 9** Results of background suppression processed by different algorithms for Seq. 2. (a) The raw image sample of Seq. 1, (b) the 3D gray map of sample image, (c) top-hat, (d) LCM, (e) RPCA, (f) LMS, (g) classic KR, and (h) proposed adaptive KR.

Table 3 lists the values of SCR and BSF of all algorithms on tested sequences. The best results are highlighted in bold. As can be seen from Table 3, the proposed method achieves the highest BSF on all sequences and the largest target enhancement on Seqs. 1 and 3. Although top-hat and LCM get highest SCRs in Seq. 2, they do not suppress clutter effectively. The quantitative results reflect that the proposed algorithm can enhance the target salience and maintain the robustness to various IR backgrounds simultaneously.

Second, the subtracted images obtained from the aforementioned processes are utilized to extract the actual target and verify the performance of the proposed energy concentration criterion. Comparison algorithms adopt the constant false alarm rate (CFAR) method as the segment algorithm. Detailed steps of CFAR are referred in Ref. 9. The extraction results of all algorithms on three sequences are shown in Figs. 11–13. The candidate points are marked by red rectangles.

Figures 11 and 12 show that the extraction results of top-hat and LCM contain many false alarms on the background edge. In contrast, RPCA, LMS, and classic KR achieve better performances than the HVS-based method. However, the background edges are not suppressed effectively and result in the increasing of false alarms.

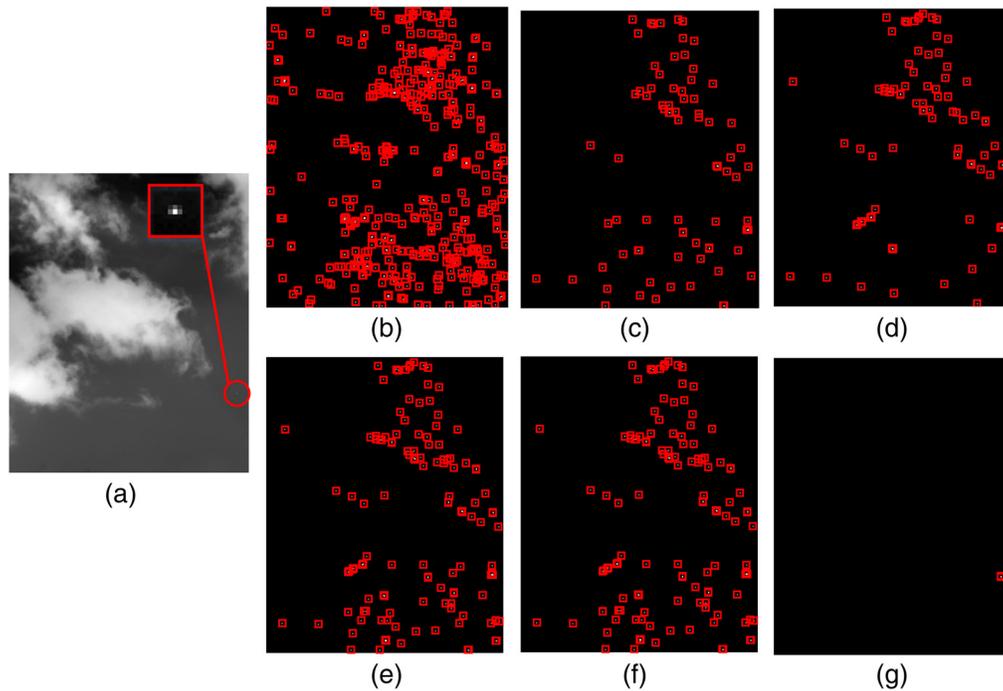
Figure 13 shows that the noise background brings a great challenge for small target detection. All the comparison algorithms do not distinguish the actual target and the detector noise well. There are many false alarms in the final detection results. On the contrary, the proposed method has the least false alarms and extracts the target accurately.



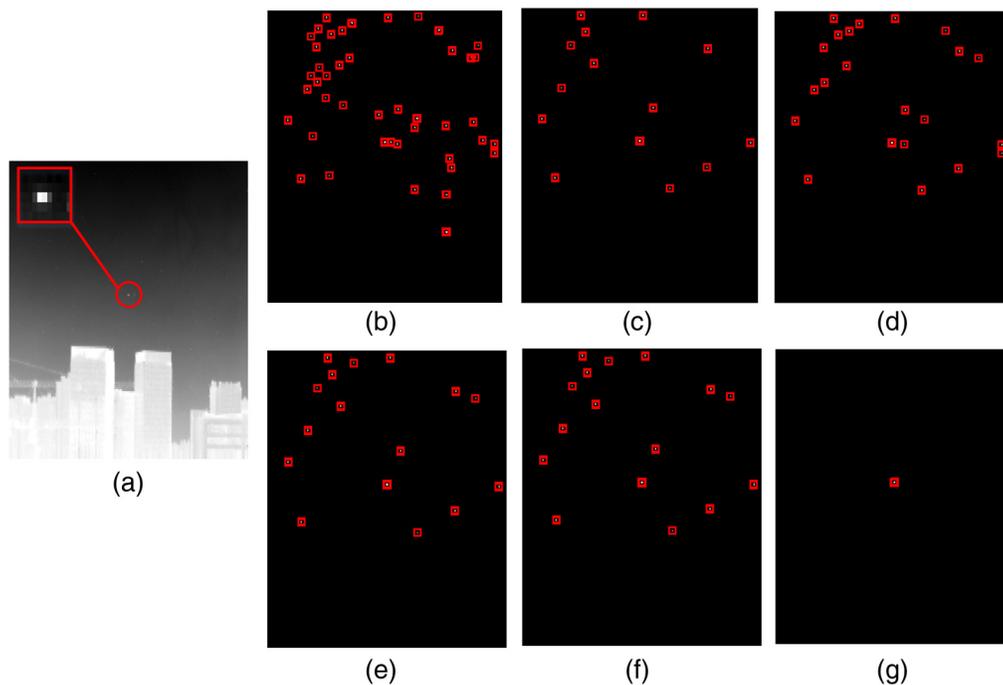
**Fig. 10** Results of background suppression processed by different algorithms for Seq. 3. (a) The raw image sample of Seq. 1, (b) the 3D gray map of sample image, (c) top-hat, (d) LCM, (e) RPCA, (f) LMS, (g) classic KR, and (h) proposed adaptive KR.

**Table 3** SCR values and BSF values for the suppression results of different algorithms.

Filtering methods	Metrics	Fig. 7(a1)	Fig. 7(b1)	Fig. 7(c1)	
				Target 1	Target2
Top-hat	SCR	2.1738	5.8998	2.9845	2.6566
	BSF	0.1525	2.7335	1.0494	1.0308
LCM	SCR	2.5088	<b>8.8872</b>	4.9181	2.8982
	BSF	0.5156	3.4790	2.1202	2.0951
RPCA	SCR	6.6556	2.5366	1.6259	1.8206
	BSF	1.7678	5.7667	1.2000	0.6819
LMS	SCR	2.0356	2.3886	1.7298	1.6877
	BSF	1.6234	6.4451	1.1048	1.0650
Classic KR	SCR	2.0392	3.7353	1.6402	1.1676
	BSF	1.8282	8.669	0.8419	0.8079
Proposed	SCR	<b>7.5474</b>	4.3097	<b>5.7998</b>	<b>6.3052</b>
	BSF	<b>10.4693</b>	<b>9.0577</b>	<b>3.2489</b>	<b>2.4412</b>

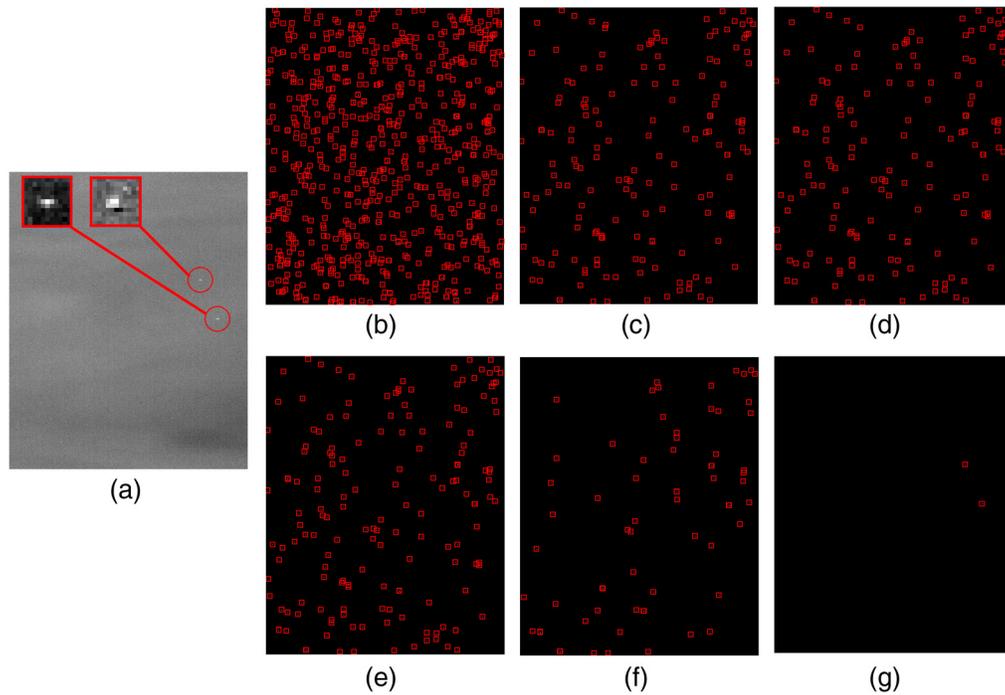


**Fig. 11** Detection results of different algorithms for Seq. 1. (a) The raw image, (b) top-hat, (c) LCM, (d) RPCA, (e) LMS, (f) classic KR, and (g) proposed energy concentration criterion.

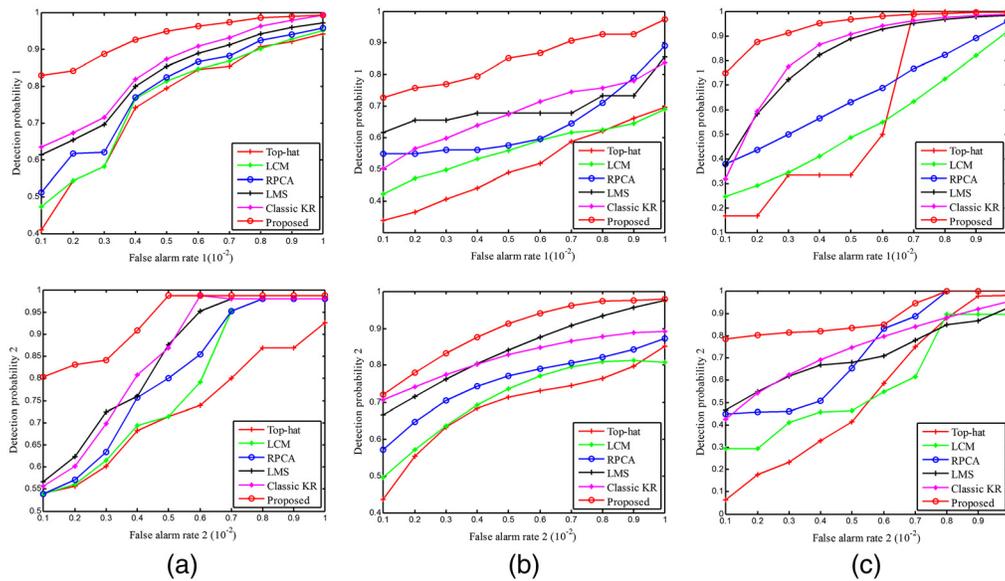


**Fig. 12** Detection results of different algorithms for Seq. 2. (a) The raw image, (b) top-hat, (c) LCM, (d) RPCA, (e) LMS, (f) classic KR, and (g) proposed energy concentration criterion.

Two types of ROC curves on three sequences are shown in Fig. 14. The corresponding calculations are shown in Eqs. (22) and (23), respectively. For Seq. 1, due to the influence of complex clouds, comparison algorithms do not achieve high detection probabilities. While, the proposed algorithm shows a strong ability to detect more than 80% of the small targets under a low false alarm rate. For Seq. 2, the strong edge of building background seriously affects the



**Fig. 13** Detection results of different algorithms for Seq. 3. (a) The raw image, (b) top-hat, (c) LCM, (d) RPCA, (e) LMS, (f) classic KR, and (g) proposed energy concentration criterion.



**Fig. 14** Two types of ROC curves of different algorithms for (a) Seq. 1, (b) Seq. 2, and (c) Seq. 3.

detection results of comparison algorithm. However, as shown in Fig. 14(b), the proposed method is robust and achieves the best detection. For Seq. 3, when the false alarm rate is very low, the ROC curves of the proposed method are not lower than 75%. And the detection probability of proposed method rises the fastest to 100% than comparison methods. While, comparison algorithms cannot get good performance under low false alarm rate. In general, trends of ROC curves in Fig. 14 indicate that under the same false alarm rate, the proposed algorithm achieves the highest detection probability in three sequences. This demonstrates that our proposed algorithm is more robust than other comparison algorithms.

The aforementioned experiment results and quantitative evaluations show that the proposed method has good abilities on target enhancement and background suppression. Also, the performance and robustness of the proposed method are both superior to other comparison methods.

## 5 Conclusions

In this paper, a robust algorithm is proposed for background removal and small-target detection in IR images. First, to suppress the complex background and the strong edge, we add the gradient feature and scale of samples into the kernel function. The smoothing matrix is reestablished by a product of a rotation matrix and an elongation matrix. Then, an MFODD filter is modeled to estimate the rotation angle and the elongation factor. Second, to extract the actual target in the subtracted image, we analysis the imaging size and the energy distribution of the small target on the IR detector. Further, an energy concentration criterion is constructed as a segment method. The experimental results demonstrate that the proposed algorithm achieves better performances than other existing state-of-the-art algorithms. The quantitative evaluations in terms of SCR, BSF, and ROC indicate that it can conspicuously suppress background and detect dim small target.

## Acknowledgments

This work was supported by the Nation Natural Science Foundation of China (Grant Nos. 61675202, 61627819, 61727818, and 61905240).

## References

1. S. Aghaziyarati, S. Moradi, and H. Talebi, "Small infrared target detection using absolute average difference weighted by cumulative directional derivatives," *Infrared Phys. Technol.* **101**, 78–87 (2019).
2. J. Han et al., "Infrared small-target detection under complex background based on subblock-level ratio-difference joint local contrast measure," *Opt. Eng.* **57**(10), 103105 (2018).
3. M. Nasiri and S. Chehresa, "Infrared small target enhancement based on variance difference," *Infrared Phys. Technol.* **82**, 107–119 (2017).
4. S. L. Wolfe, *Introduction to Infrared System Design*, SPIE Optical Engineering Press, Bellingham, Washington (1996).
5. L. Huo et al., "Staircase-scene-based nonuniformity correction in aerial point target detection systems," *Appl. Opt.* **55**, 7149–7156 (2016).
6. K. Bai et al., "Robust contact-point detection from pantograph-catenary infrared images by employing horizontal-vertical enhancement operator," *Infrared Phys. Technol.* **101**, 181–185 (2019).
7. Y. Dai et al., "Non-negative infrared patch-image model: robust target-background separation via partial sum minimization of singular values," *Infrared Phys. Technol.* **81**, 182–194 (2017).
8. Q. Song et al., "Single frame infrared image small target detection via patch similarity propagation based background estimation," *Infrared Phys. Technol.* **106**, 103197 (2020).
9. R. Liu et al., "An omnidirectional morphological method for aerial point target detection based on infrared dual-band model," *Remote Sens.* **10**, 1054 (2018).
10. B. Narayanan, R. C. Hardie, and R. A. Muse, "Scene-based nonuniformity correction technique that exploits knowledge of the focal-plane array readout architecture," *Appl. Opt.* **44**(17), 3482 (2005).
11. S. Zhao et al., "Infrared target detection method based on the receptive field and lateral inhibition of human visual system," *Appl. Opt.* **56**(30), 8555–8563 (2017).
12. Y. Qian et al., "Infrared small target detection based on saliency and gradients difference measure," *Opt. Quant. Electron.* **52**, 151 (2020).

13. Y. Wang et al., "Small target detection using edge-preserving background estimation based on maximum patch similarity," *Int. J. Adv. Rob. Syst.* **14**(6), 1–11 (2017).
14. Y. Li and Y. Zhang, "Robust infrared small target detection using local steering kernel reconstruction," *Pattern Recognit.* **77**, 113–125 (2018).
15. T. Ma et al., "Infrared small target detection based on divergence operator and nonlinear classifier," *Opt. Quantum Electron.* **53**, 369 (2021).
16. M. Ju et al., "ISTDet: an efficient end-to-end neural network for infrared small target detection," *Infrared Phys. Technol.* **114**, 103659 (2021).
17. L. Genin, F. Champagnat, and G. L. Besnerais, "Background first- and second-order modeling for point target detection," *Appl. Opt.* **51**, 7701–7713 (2012).
18. C. Chen et al., "A local contrast method for small infrared target detection," *IEEE Trans. Geosci. Remote Sens.* **52**(1), 574–581 (2014).
19. Y. Chen and Y. Xin, "An effective infrared small target detection method based on visual contrast mechanism," *IEEE Trans. Geosci. Remote Sens. Lett.* **13**(7), 962–966 (2016).
20. J. Liu et al., "Tiny and dim infrared target detection based on weighted local contrast," *IEEE Trans. Geosci. Remote Sens. Lett.* **15**(11), 1780–1784 (2018).
21. C. Xia et al., "Infrared small target detection based on multiscale local contrast measure using local energy factor," *IEEE Geosci. Remote Sens. Lett.* **17**(1), 157–161 (2020).
22. J. Han et al., "A local contrast method for infrared small-target detection using a tri-layer window," *IEEE Geosci. Remote Sens. Lett.* **17**(10), 1822–1826 (2020).
23. Y. Chen et al., "An effective infrared small target detection method based on the human visual attention," *Infrared Phys. Technol.* **95**, 128–135 (2018).
24. J. Han et al., "A local contrast method combined with adaptive background estimation for infrared small target detection," *IEEE Geosci. Remote Sens. Lett.* **16**(9), 1442–1446 (2019).
25. H. Wang et al., "A novel and high-speed local contrast method for infrared small-target detection," *IEEE Geosci. Remote Sens. Lett.* **17**(10), 1812–1816 (2020).
26. E. Lee, E. Gu, and K. Park, "Effective small target enhancement and detection in infrared images using saliency map and image intensity," *Opt. Rev.* **22**(4) 659–668 (2015).
27. X. Wang, C. Ning, and L. Xu, "Spatiotemporal difference-of-Gaussians filters for robust infrared small target tracking in various complex scenes," *Appl. Opt.* **54**(7), 1573–1586 (2015).
28. S. Kim, "Min-local-LoG filter for detecting small targets in cluttered background," *Electron. Lett.* **47**(2) 105–106 (2011).
29. C. Corbane et al., "Fully automated procedure for ship detection using optical satellite imagery," *Proc. SPIE* **7150**, 71500R (2008).
30. X. Bai, "Morphological center operator for enhancing small target obtained by infrared imaging sensor," *Optik* **125**, 3697–3701 (2014).
31. X. Bai, F. Zhou, and B. Xue, "Fusion of infrared and visual images through region extraction by using multi scale center-surround top-hat transform," *Opt. Express* **19**, 8444–8457 (2011).
32. X. Bai, F. Zhou, and B. Xue, "Multiple linear feature detection based on multiple-structuring-element center-surround top-hat transform," *Appl. Opt.* **51**, 5201–5211 (2012).
33. B. Zhu and Y. Xin, "Effective and robust infrared small target detection with fusion of poly-directional first order derivative images under facet model," *Infrared Phys. Technol.* **69**, 136–144 (2015).
34. X. Cao et al., "Infrared small target detection based on derivative dissimilarity measure," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **12**(8), 3101–3116 (2019).
35. Y. Bi et al., "Fast detection of distant, infrared targets in a single image using multiorder directional derivatives," *IEEE Trans. Aerosp. Electron. Syst.* **56**(3), 2422–2436 (2020).
36. M. M. Hadhoud and D. W. Thomas, "The two-dimensional adaptive LMS (TDLMS) algorithm," *IEEE Trans. Circuits Syst.* **35**(5), 485–494 (1998).
37. Y. Cao, R. Liu, and J. Yang, "Small target detection using two-dimensional least mean square (TDLMS) filter based on neighbor analysis," *Infrared Millim. Waves* **29**(2), 188–200 (2008).
38. T. Bae, F. Zhang, and I. Kweon, "Edge directional 2D LMS filter for infrared small target detection," *Infrared Phys. Technol.* **55**, 137–145 (2012).

39. Y. Yang, J. Yang, and K. Yang, "Adaptive detection for infrared small target under sea-sky complex background," *Electron. Lett.* **40**(17), 1083–1085 (2004).
40. Y. Gu et al., "A kernel-based nonparametric regression method for clutter removal in infrared small-target detection applications," *IEEE Trans. Geosci. Remote Sens. Lett.* **7**(3), 469 (2010).
41. H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.* **16**(2), 349–366 (2007).
42. T. Zhang et al., "Infrared small target detection via self-regularized weighted sparse model," *Neurocomputing* **420**, 124–148 (2021).
43. Y. Lyu et al., "Cirrus detection based on RPCA and fractal dictionary learning in infrared image," *Remote Sens.* **12**(1), 142 (2020).
44. C. Gao et al., "Infrared patch-image model for small target detection in a single image," *IEEE Trans. Image Process.* **22**(12), 4996–5009 (2013).
45. Y. Dai, Y. Wu, and Y. Song, "Infrared small target and background separation via column-wise weighted robust principal component analysis," *Infrared Phys. Technol.* **77**, 421–430 (2016).
46. X. Wang et al., "Infrared dim target detection based on total variation regularization and principal component pursuit," *Image and Vision Comp.* **63**, 1–9 (2017).
47. L. Zhang et al., "Infrared small target detection via non-convex rank approximation minimization joint  $l_2$ ,  $l_1$  norm," *Remote Sens.* **10**(11), 1821 (2018).
48. T. Zhang et al., "Infrared small target detection based on non-convex optimization with  $l_p$ -norm constraint," *Remote Sens.* **11**(5), 559 (2019).
49. Y. Sun et al., "Infrared small-faint target detection using non-I.I.D. mixture of Gaussians and flux density," *Remote Sens.* **11**(23), 2831 (2019).
50. Y. Zhang et al., "Small infrared target detection based on low-rank and sparse representation," *Infrared Phys. Technol.* **68**, 98–109 (2015).
51. M. V. Shirvaikar and M. M. Trivedi, "A neural network filter to detect small targets in high clutter backgrounds," *IEEE Trans. Neural Network* **6**, 252–257 (1995).
52. S. Park and J. Lim, "Non-asymptotic rate for high-dimensional covariance estimation with non-independent missing observations," *Stat. Probabil. Lett.* **153**, 113–123 (2019).
53. S. Datcu, L. Ibos, and Y. Candau, "Focal plane array infrared camera transfer function calculation and image restoration," *Opt. Eng.* **43**(3), 648–657 (2004).
54. L. Fortunato et al., "SKYWARD: the next generation airborne infrared search and track," *Proc. SPIE* **9819**, 98190K (2016).
55. P. Lv et al., "A method for weak target detection based on human visual contrast mechanism," *IEEE Trans. Geosci. Remote Sens. Lett.* **16**(2), 261–265 (2019).
56. K. Zhang and X. Li, "Infrared small dim target detection based on region proposal," *Optik* **182**, 961–973 (2019).

**Mingyang Ma** is currently a PhD student at Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His research interests include infrared imaging, target detection, estimation theory, and sensor data fusion.

**Dejiang Wang** is currently a research fellow at Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His research interests include signal processing, infrared image processing, and pattern recognition.

**He Sun** is currently an assistant research fellow at Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. Her research interests include polarization optical technology, image processing, and computer vision.

**Tao Zhang** is currently a research fellow at Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. His research interests include nonlinear control and estimation, signal processing, and imaging techniques.