applied optics

Point target detection based on multiscale morphological filtering and an energy concentration criterion

Rang Liu,[1,2,3] **Dejiang Wang,**[2,3,*] **Dabiao Zhou,**[1,2,3] **and Ping Jia**[2,3]

[1]*University of Chinese Academy of Sciences, Beijing 100049, China*
[2]*Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China*
[3]*Key Laboratory of Airborne Optical Imaging and Measurement, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China*
**Corresponding author: wangdj04@ciomp.ac.cn*

Received 31 March 2017; revised 1 July 2017; accepted 19 July 2017; posted 20 July 2017 (Doc. ID 291641); published 14 August 2017

The research on optical imaging characteristics of infrared dim point targets in the presence of nonstationary cloud clutter and random noise is necessary for target detection. We analyze the energy concentration of point targets that are less than 3 × 3 pixels in size and deduce a simulation model of the point target imaging process. Then we adopt omnidirectional multiscale structural elements to detect all the possible targets distributing in every direction. The adaptive threshold and the energy concentration criterion are employed to eliminate false alarms. Finally, the trajectory of point targets is obtained after the low-order recursive correlation. The results show that the detection probability of the proposed method reaches 99.8% with 0.2% false alarm probability. It demonstrates that the proposed method has a good performance to suppress complex background and random noise. Also, it has the advantage of low complexity and easy implementation in a real-time system.

OCIS codes: (110.3080) Infrared imaging; (100.2550) Focal-plane-array image processors; (040.2480) FLIR, forward-looking infrared; (110.4280) Noise in imaging systems.

https://doi.org/10.1364/AO.56.006796

1. INTRODUCTION

The technology of infrared dim point target detection has extensive applications in numerous fields, including infrared (IR) search and track (IRST) systems, terminal guidance, external intrusion warning, and medical monitoring [1,2]. Detecting the point target as early as possible can provide more response time for users, so it is of great significance to improve the robustness and efficiency of detection algorithms. However, the point target's imaging size and minutiae are very small due to the long distance from the targets to the infrared focal plane arrays (IRFPAs). Moreover, the imaging process is always affected by bad weather, atmospheric radiation, nonstationary clouds, and heavy noises, which are likely to lead to the radiation intensities of some scenes in the infrared image over the point target [3–5]. Otherwise, the isolated noise is more easily confused with the point target, which can increase false alarms. However, the upside is that the point target will diffuse into an Airy spot because of the diffraction-limited optical system, and the actual size of the point target is larger than the ideal imaging size of geometrical optics. This phenomenon causes the different energy concentrations in subpixel locations between the point targets and random noise.

Over the past few decades, many researchers have paid much attention to dim point target detection, which focused mainly around background clutter suppression and target enhancement. The spatial methods such as max-mean filtering [6], max-median filtering [7], and the two-dimensional projection algorithm [8] are widely used to reduce background clutter. Although these methods are simple and fast, they result in different degrees of background edges and high-frequency noise, which lead to many false alarms. Chen *et al.* [9] presented a local contrast method based on the human vision system (HVS) that achieved a low false alarm rate. Genin *et al.* [10] employed block-matching 3D filtering (BM3D) and the Gaussian mixture model (GMM) to suppress the complex background. It achieves a perfect effect, yet it is time-consuming. Also, some researchers have adopted frequency domain methods such as the wavelet transform [11,12] and Butterworth high-pass filtering [3]. In a word, many algorithms can provide an excellent performance in background suppression but may reduce target details or fail to meet real-time

1559-128X/17/246796-10 Journal © 2017 Optical Society of America

requirements. Some algorithms highlight the noise when enhancing target details.

In recent years, morphological filtering algorithms have been generally applied to target detection and recognition, especially the top-hat transform, which is based on the geometry and set theory [13,14]. The top-hat transform could extract bright and dim image regions corresponding to the structural element. The role of the structural element in morphological operations is similar to the filtering window [15,16]. The effect of background suppression greatly depends on the size and shape of the selected structural element. The conventional top-hat method to detect a point target only uses a single structural element, which cannott cover various targets with changing sizes in real cases and is more likely to be submerged in heavy clutter. Bai *et al.* [14,15] presented a multiscale center-surround top-hat transform through constructing two structural elements and successfully extracted regions of interest that were richer in image details than those made using a single structural element.

In this paper, we propose a low-order recursive method based on omnidirectional multiscale morphological filtering and the characteristic criterion of the energy concentration to detect dim point targets with low signal-to-noise ratio (SNR). We analyze in detail the optical imaging characteristics of dim point targets, including the morphological information, energy concentration characteristics, background clutter, and noise. The eight-direction morphological transform and the adaptive threshold are employed to suppress the clutter and obtain the region of interest (ROI) of target. The optical diffraction leads to the diffusion of the point target, and its energy concentration is lower than random noise. The remaining random noise is eliminated according to the proposed energy concentration criterion. Finally, the low-order recursive method is adopted to detect the moving target from a long distance. The performance of our algorithm is demonstrated by three typical infrared sequences, as well as a comparison with other algorithms. The results prove that our algorithm is effective for point target detection between the performance and computational time.

This paper is organized as follows. Section 2 analyzes the morphological information and the energy concentration of point targets, and a simulation model of point target imaging process is obtained. In Section 3, we describe the proposed low-order recursive method based on omnidirectional multiscale morphological filtering and the energy concentration criterion. Section 4 presents the experiments on real infrared images and the results of the proposed method. Finally, Section 5 offers the conclusion of the paper.

2. IMAGING CHARACTERISTICS

A. Characteristics of Point Target, Background, and Noise

The research on the imaging characteristics of a point target has an important effect on the performance of the detection algorithm. The imaging size of a point target in motion changes from 1 to 3 × 3 pixels, and the morphological information changes with the imaging distance, attitude angle, and imaging position on IRFPA. Figure 1 presents eight typical point targets (the size of the window is 5 × 5 pixels) captured by a long-wave infrared detector. It can be seen clearly that the energy at the central pixel has diffused to neighboring pixels, resulting in the actual situation that the size of the point target is rarely one pixel. In general, the brightest pixel in the target area is considered as the point target's central pixel. Moreover, the point target is more flexible and may appear in the various complex scenes at different scales. Therefore, the gray distribution of the point target is not necessarily isotropic.

Most of the background in infrared images contains low-frequency components such as clouds, mountains, sea, and buildings, etc. The internal change is relatively smooth, such as the continuous clouds that appear in Fig. 2(a). The low-frequency background has a great correlation in spatial distribution. But the background may also contain high-frequency components, mainly in edges and textures, as in Fig. 2(b).

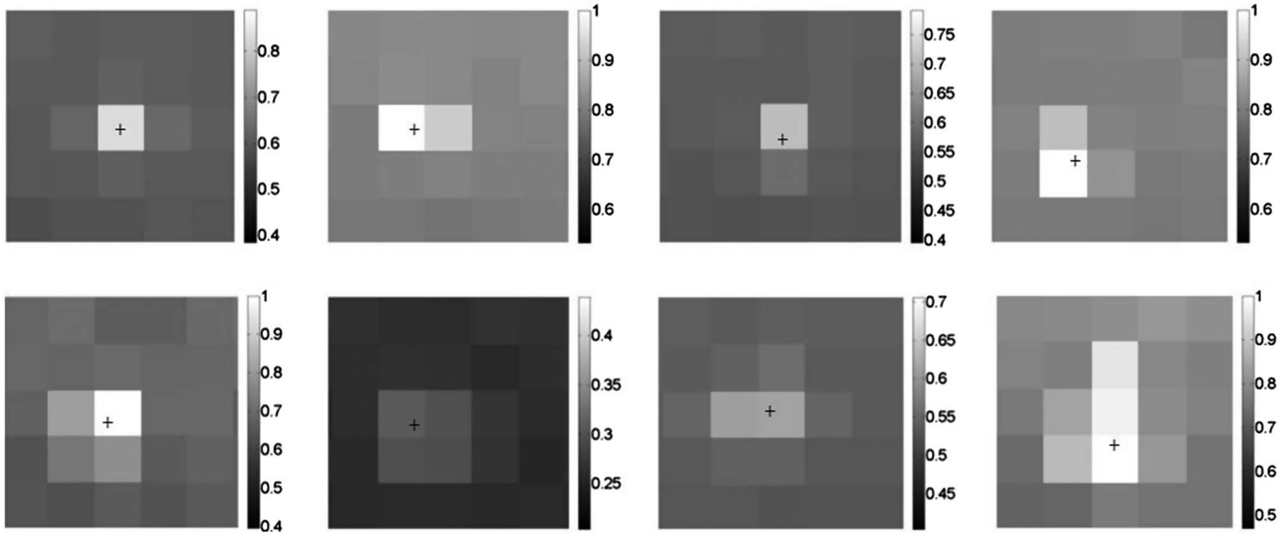
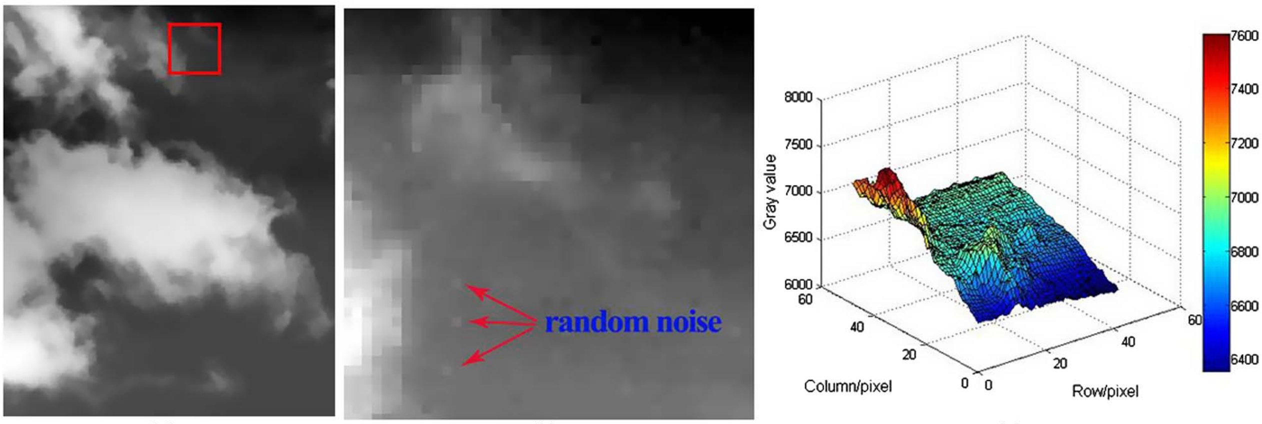


Fig. 1. Typical morphological information of a point target captured by a long-wave infrared detector. The gray value has been normalized to [0,1], and the point target's central pixel has been marked with a cross.



(a) (b) (c)

Fig. 2. Cloud edges and isolated noise, which interfere in point target detection: (a) complex background, (b) magnification of the red box located at (a), and (c) corresponding three-dimensional grayscale map.

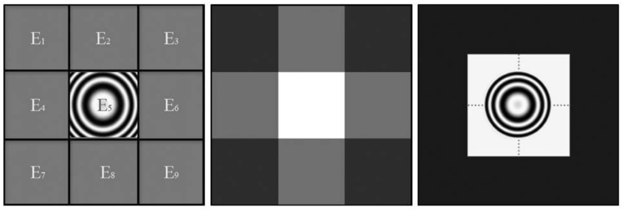
A small amount of isolated noise can be clearly seen in Fig. 2(b), and it is marked with a red arrow. Generally, noise can be divided into two categories: one is the external noise generated by the external environment, and the other is the internal noise generated by the infrared thermal imaging system. In most cases, the isolated noise is position invariant or has random fluctuation that is likely to be mistaken for the point target. The fixed noise can be eliminated by the inter-frame relation. But it is difficult to separate the random noise from the point target using traditional methods. Better yet, the energy concentration of the point target's central pixel is different from the random noise when imaging on IRFPA. We can establish the energy concentration criterion to eliminate the random noise. For simplicity, the energy concentration of the point target's central pixel is called the energy concentration of the point target in the following.

B. Energy Concentration of Point Targets

Here we do not consider point targets whose imaging size is more than 3 × 3 pixels. It is assumed that the energy distribution of the point target is $E_1 \sim E_9$, as demonstrated in Fig. 3(a), and the maximum energy is $E_{\max}$. The energy concentration (EC) of the point target is defined as

$$\mathrm{EC} = \frac{E_{\max}}{\sum_{i=1}^{9} E_i}. \qquad (1)$$

The imaging position of a point target may be in the central pixel, or span two or even four pixels. The EC of the point target will inevitably fluctuate with the changing of imaging position. Diffraction occurs in the long-distance imaging process, so the point target may be regarded roughly as an Airy spot. When the target is imaging in the central pixel, as shown in Fig. 3(b), the Airy spot is located in E_5. The obtained EC is the imaging system's energy concentration, and it is also the maximum EC of the point targets. When the point target is imaging at the intersection angle of the four pixels, such as in Fig. 3(c), the EC of the point target is the lowest. Then we derive the theoretical value of the point target's EC from the perspective of the imaging system.



(a) (b) (c)

Fig. 3. Simulation results of the imaging position of the point target: (a) an Airy spot located in the center, (b) maximum energy concentration, and (c) minimum energy concentration of point target.

In the actual imaging process, the point target is filtered by the optical system, and the high-frequency part of the target edges is suppressed. Many imaging simulation methods do not take the point spread function (PSF) into account. In fact, the imaging process of point targets can be described as follows. First, the target's energy is converged to the focal plane arrays through the optical system. The PSF of the optical system can describe the shape and energy distribution of point targets in this step. Next, the detector will integrate and sample the target, which is situated in the focal plane array, and we finally obtain the digital image containing the point target. Various noises will unavoidably appear in the imaging process. The actual imaging process of the point target can be described by the following formula:

$$g(x,y) = S[\mathrm{PSF} * f(x,y)] + n, \qquad (2)$$

where $f(x,y)$ is the energy of the ideal point target imaging in the position (x,y). PSF is the point spread function of optical systems, and S represents the detector integration and sampling process. n is the noise, which can be approximated as white Gaussian noise whose mean value is zero, and $g(x,y)$ is the energy of the actual point target after degradation. The block diagram in Fig. 4 illustrates the specific imaging process.

With the imaging distance of the target farther away, it is gradually reducing to a point on IRFPA whose PSF is limited by the optical system. The ideal PSF of the optical system can be represented as a first-order Bessel function,

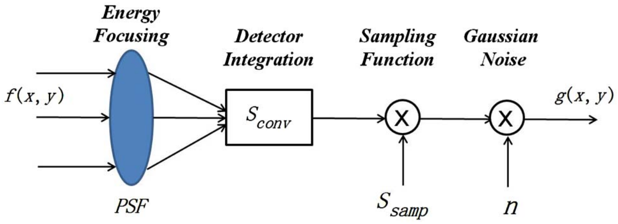


Fig. 4. Basic block diagram of the specific imaging process.

$$\mathrm{PSF} = \left[\frac{2J_1(\pi r/\lambda F)}{\pi r/\lambda F}\right]^2, \tag{3}$$

where J_1 is the first-order Bessel function, and r is the radius of the target which is calculated from $r = \sqrt{x^2 + y^2}$. F is the parameter of the optical system, and $F = \frac{f}{D}$. f and D represent focal length and aperture of the optical system, respectively. λ is the wavelength of incident light. F and λ are assumed to be 2 μm and 9 μm, respectively, in the following simulation.

The first-order Bessel function is appropriate for the ideal diffraction-limited system. But in the real optical system, there may be a Gauss function because of the aberration and distortion. A Gauss function can reflect the dispersion characteristics of horizontal and vertical directions. The dispersion characteristics of the optical system could be considered isotropic in the deduction process of energy concentration.

The imaging process is the course of converting a continuous mode to a discrete signal, which can be regarded as the process of aperture convolution and discrete sampling in the integration time. The convolution function of a detector element can be considered as a smooth window, which is determined by the shape of array. The detector elements are usually rectangular, so the convolution function of a single detector element is represented by a rectangular function as

$$S_{\mathrm{conv}} = \mathrm{rect}\left(\frac{x}{a}, \frac{y}{b}\right), \tag{4}$$

where a, b are the dimensions of single detector element. Limited by the manufacturing process, the detector element cannot be fully engaged, and the fill factor of the detector is less than 100%. Assuming that the fill factor of the detector is μ, the convolution function of single detector element is expressed as

$$S_{\mathrm{conv}} = \mathrm{rect}\left(\frac{x}{\mu a}, \frac{y}{\mu b}\right). \tag{5}$$

The next step is the sampling process of the detector. We will get the digital image after the discrete sampling, and this process adopts the dressing function, which is expressed as

$$S_{\mathrm{samp}} = \frac{1}{ab}\mathrm{comb}\left(\frac{x}{a}, \frac{y}{b}\right) = \frac{1}{ab}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta\left(\frac{x}{a} - m, \frac{y}{b} - n\right). \tag{6}$$

Equations (5) and (6) are combined to obtain the aperture convolution and discrete sampling model of the infrared detector, shown as

$$S = S_{\mathrm{conv}} * S_{\mathrm{samp}} = \frac{1}{ab}\mathrm{rect}\left(\frac{x}{\mu a}, \frac{y}{\mu b}\right) * \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta\left(\frac{x}{a} - m, \frac{y}{b} - n\right). \tag{7}$$

We plug Eqs. (3) and (7) into Eq. (2), and then the output of the whole system is obtained:

$$\begin{aligned} g(x, y) &= S[\mathrm{PSF} * f(x, y)] + n \\ &= \frac{1}{ab}\left[\frac{2J_1(\pi r/\lambda F)}{\pi r/\lambda F}\right]^2 * f(x, y) * \mathrm{rect}\left(\frac{x}{\mu a}, \frac{y}{\mu b}\right) \\ &\quad * \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta\left(\frac{x}{a} - m, \frac{y}{b} - n\right) + n. \end{aligned} \tag{8}$$

Next, we use the model proposed above to calculate the EC of the point target. It is assumed that the fill factor of the long-wave infrared detector is 0.9, and the element size is 30 μm. All operations are performed in a small pixel interval; thus the influence of noise n will not be considered in the simulation. The relationships between the EC and the across-pixel characteristics of the point target are obtained after the energy normalization shown in Fig. 5.

In the simulation result, the independent variables i and j represent the across-pixel number that deviates from the center of the detector element in the horizontal and vertical directions ($i, j \in [-0.5, 0.5]$), and then the coordinate (i, j) is regarded as the imaging position of the point target's central pixel on IRFPA. It is obviously that the coordinate (0, 0) represents the point target that is located in the central pixel. The dependent variable in the z axis represents the energy of the point target imaging in the position (i, j), which has been normalized to [0,1] so that it is also the EC of the point target in this position.

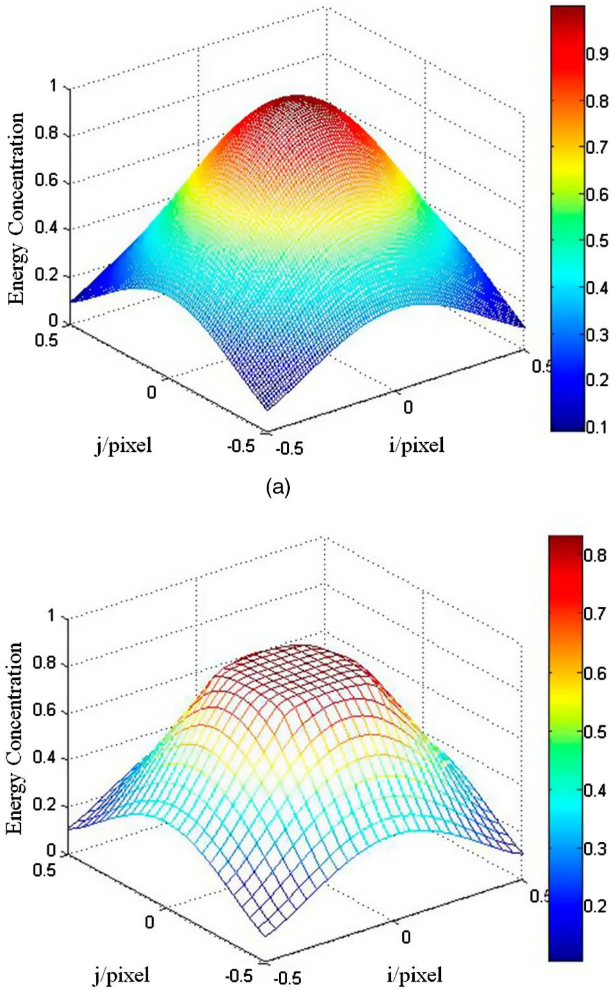
From the simulation result illustrated in Fig. 5(b), we can see that the maximum EC of the point target that the infrared detector can detect is 0.842 when the point target is imaging in the central pixel. The farther away from the central pixel, the smaller the EC will be. The minimum EC is 0.115, and the point target is imaging at the intersection of four pixels in this case.

However, the random noise almost occupies only one pixel, and there is no diffusion effect, just like a pulse point. The energy of noise is highly concentrated and its EC is about 1. Therefore, we can set up a certain threshold criterion of EC to separate the point target and the random noise. The high threshold of EC can be set as 0.9, and the low threshold as 0.1. The random noise can be eliminated easily after the EC criterion is adopted.

When calculating the SNR and EC of the point target in the specific infrared image, it is carried out in the local neighborhood around the point target, as demonstrated in Fig. 6. The calculation of SNR is as follows:

$$\mathrm{SNR} = \frac{|\overline{\mu_t} - \overline{\mu_b}|}{\sigma_b}, \tag{9}$$

where $\overline{\mu_t}$ is gray mean of the target area, and $\overline{\mu_b}$ is the gray mean of the local background area, so the molecular term can be seen as the absolute energy of the point target. σ_b is the



(a)

(b)

Fig. 5. Simulation results of imaging system: (a) result of point target passing through the PSF of the optical system, and (b) result of energy concentration of point target after the convolution integral and discrete sampling by the detector.

standard deviation of the local background area, which is calculated from

$$\sigma_b^2 = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} [\mu_b(i,j) - \overline{\mu_b}]^2, \tag{10}$$

where M, N are the sizes of the local background areas.

According to the definition of SNR, we calculate the EC of the point target in the specific image as

$$\mathrm{EC} = \frac{\mu_{t\ \max} - \overline{\mu_b}}{\sum_t |\mu_t - \overline{\mu_b}|}, \tag{11}$$

where $\mu_{t\ \max}$ is the maximum gray value of point target area μ_t, and the denominator term is the sum of the gray values of all pixels in the target area (3 × 3 pixels) minus the local background area, which can be seen as the absolute total energy of the point target.

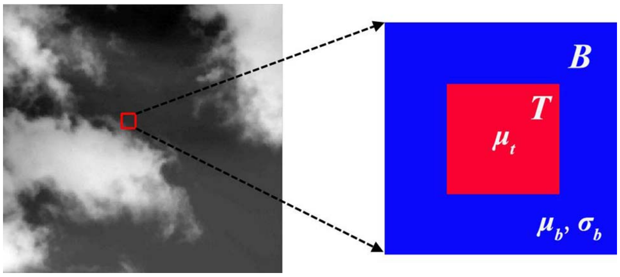


Fig. 6. The right is the enlarged local area of the point target and its surrounding background, which is in the red box, left, while T is the point target, and B is the background area around the target.

3. POINT TARGET DETECTION

A. Omnidirectional Multiscale Morphological Filtering

To carry out the effective detection of an infrared point target, the most important thing is to suppress background clutter. Morphological filtering is used to extract the corresponding form of the image with a certain structural element to realize the target recognition. The most basic operations in morphological filtering are dilation and erosion.

We let $I(x,y)$ represent the original grayscale image, and $b(s,t)$ is the structural element. $\oplus$ represents the dilation operation, and Θ represents the erosion operation. The dilation and erosion of $I(x,y)$ and $b(s,t)$ are, respectively, defined as

$$I \oplus b = \max_{s,t}(I(x-s, y-t) + b(s,t)), \tag{12}$$

$$I \Theta b = \min_{s,t}(I(x+s, y+t) - b(s,t)). \tag{13}$$

The open operation in morphological filtering can get rid of the region that is less than the structural element, and the closed operation can connect two regions that are very close. The open operation and closed operation of $I(x,y)$ by $b(s,t)$, denoted by $I \circ b$ and $I \bullet b$, are defined based on dilation and erosion, shown as follows:

$$I \circ b = [I(x,y) \Theta b(s,t)] \oplus b(s,t), \tag{14}$$

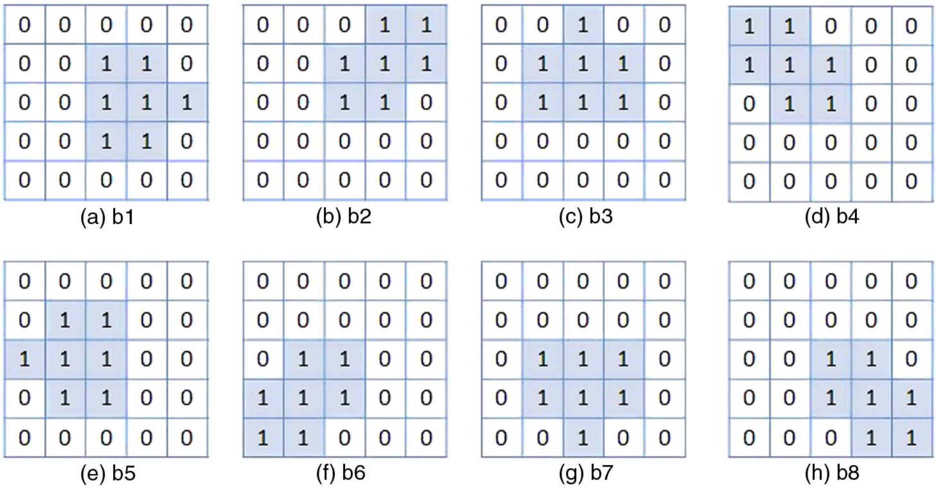
$$I \bullet b = \left[I(x,y) \oplus b(s,t)\right] \Theta b(s,t). \tag{15}$$

We use the structural element larger than the target to make an open operation, so as to obtain the background image. Then we subtract the background image from the original image, which is the top-hat transform. The top-hat transform can extract bright details from the image. The bottom-hat transform is the difference between the closed operation and the original image, which can extract dark details of the image. Top-hat (TH) and bottom-hat (BH) are defined as

$$\mathrm{TH} = I(x,y) - (I \circ b)(x,y), \tag{16}$$

$$\mathrm{BH} = (I \bullet b)(x,y) - I(x,y). \tag{17}$$

On the basis of the original image, we add the top-hat transform results and then subtract the bottom-hat transform results. The final result is that the background can be suppressed and the bright ROI, which may contain point targets,



0	0	0	0	0
0	0	1	1	0
0	0	1	1	1
0	0	1	1	0
0	0	0	0	0

(a) b1

0	0	0	1	1
0	0	1	1	1
0	0	1	1	0
0	0	0	0	0
0	0	0	0	0

(b) b2

0	0	1	0	0
0	1	1	1	0
0	1	1	1	0
0	0	0	0	0
0	0	0	0	0

(c) b3

1	1	0	0	0
1	1	1	0	0
0	1	1	0	0
0	0	0	0	0
0	0	0	0	0

(d) b4

0	0	0	0	0
0	1	1	0	0
1	1	1	0	0
0	1	1	0	0
0	0	0	0	0

(e) b5

0	0	0	0	0
0	0	0	0	0
0	1	1	0	0
1	1	1	0	0
1	1	0	0	0

(f) b6

0	0	0	0	0
0	0	0	0	0
0	1	1	1	0
0	1	1	1	0
0	0	1	0	0

(g) b7

0	0	0	0	0
0	0	0	0	0
0	0	1	1	0
0	0	1	1	1
0	0	0	1	1

(h) b8

Fig. 7. Eight-direction structural elements, including b_1 of 0° direction, b_2 of 45° direction, b_3 of 90° direction, b_4 of 135° direction, b_5 of 180° direction, b_6 of 225° direction, b_7 of 270° direction, and b_8 of 315° direction. "1" and "0" are the basic binary morphological operators.

can be enhanced [17,18]. The extracted ROI result is expressed as

$$\mathrm{ROI} = I + \mathrm{TH} - \mathrm{BH}. \tag{18}$$

The size and shape of point targets in motion are changing all the time, and a point target is more likely to occur in the complex background at various scales. The traditional top-hat transform method only adopts a single structural element to estimate the background of the infrared image, ignoring the differences of point targets distributing in every direction, which does not produce the best results. In this paper, we adopt the multiscale top-hat and bottom-hat transforms, which are defined as follows:

$$\mathrm{TH}_n = I(x, y) - (I \circ b_n)(x, y), \tag{19}$$

$$\mathrm{BH}_n = (I \bullet b_n)(x, y) - I(x, y), \tag{20}$$

where b_n represents different structural elements, and n is a positive integer.

Since the point target is very tiny in nonstationary clutter background and noise, the dimensions of the selected structural elements should not be too large. We design eight omnidirectional multiscale structural elements that have dimensions of 5 × 5 to extract the point targets distributing in every direction as far as possible. The eight-direction structural elements b_n $(n = 1, 2, \ldots, 8)$ contain 0° direction, 45° direction, 90° direction, …, and 315° direction, as shown in Fig. 7. The morphological structural elements with specific directions can effectively eliminate the nondirectional and continuous background clutter.

The operator "1" in each structural element represents that the dilation or erosion operation is performed in its position, and the distribution direction of "1" is similar to the morphological direction of the point target. Each structural element can not only extract the point targets that are distributed in the specified direction, but can also highlight the candidate points at smaller scales than the structural element in the image, thus avoiding omission. Taking the structural element b_1 as an example, the major bright ROIs that it is able to extract are illustrated in Fig. 8.

The proposed eight-direction structural elements cover almost all the point targets distributing in every direction at all scales. For each kind of structural element, the operations of the top-hat and bottom-hat transforms can be used to detect the bright and dark regions in the direction corresponding to the structural elements.

The omnidirectional bright image regions extracted by b_1 to b_8 can be expressed as follows:

$$\begin{aligned} \mathrm{TH}_1 &= I(x, y) - (I \circ b_1)(x, y), \\ \mathrm{TH}_2 &= I(x, y) - (I \circ b_2)(x, y), \\ &\ldots \\ \mathrm{TH}_8 &= I(x, y) - (I \circ b_8)(x, y). \end{aligned} \tag{21}$$

Equally, the omnidirectional dark regions extracted by b_1 to b_8 can be expressed as follows:

$$\begin{aligned} \mathrm{BH}_1 &= (I \bullet b_1)(x, y) - I(x, y), \\ \mathrm{BH}_2 &= (I \bullet b_2)(x, y) - I(x, y), \\ &\ldots \\ \mathrm{BH}_8 &= (I \bullet b_8)(x, y) - I(x, y). \end{aligned} \tag{22}$$

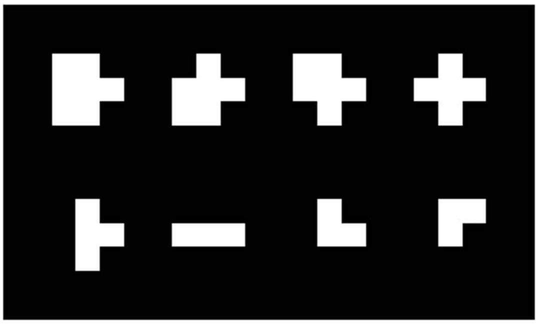
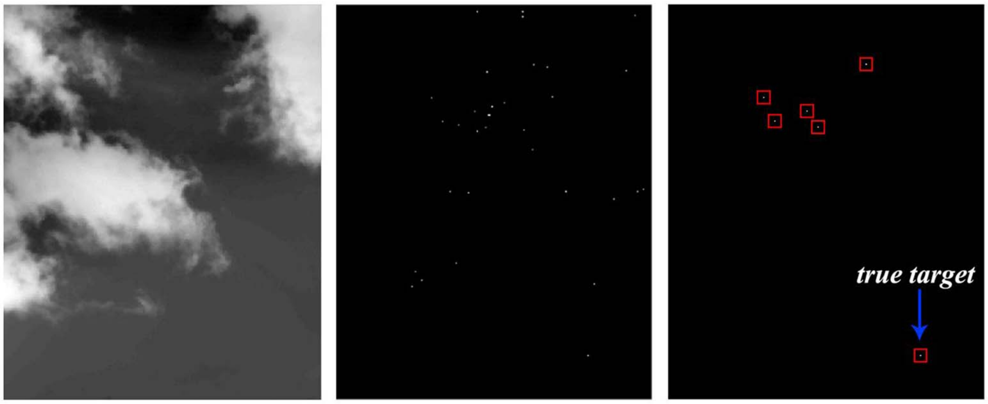


Fig. 8. Major bright ROIs that the structural element b_1 can extract.



(a) (b) (c)

Fig. 9. Results of the algorithm: (a) original image; (b) CFAR result; (c) result of energy concentration threshold.

The ROI_n of the suspected target extracted by structural element b_n is expressed as

$$\mathrm{ROI}_n = I + \mathrm{TH}_n - \mathrm{BH}_n. \tag{23}$$

Each ROI_n is most likely to contain suspected point targets distributing in the detection direction, so we add the eight extracted ROI_n into a big ROI, which is calculated in Eq. (24):

$$\mathrm{ROI} = |\mathrm{ROI}_1 + \mathrm{ROI}_2 + \ldots + \mathrm{ROI}_8|. \tag{24}$$

The constant false alarm rate (CFAR) is used to deal with the above ROI, and then we get the candidate targets. The threshold is calculated in the local small neighborhood that contains the suspected target,

$$\mathrm{Th} = \overline{\mu_b} + k\sigma_b, \tag{25}$$

where Th is the obtained adaptive threshold, and k is a fixed constant that can be regarded as the threshold of the SNR. On theory of binary image, dealing with ROI by threshold Th. To display the results of the proposed algorithm, k is assigned to $\sqrt{2}$ according to the SNR of point targets to be detected. Figure 9(a) presents the original image that contains one point target, which is grabbed by the infrared detector of fixed base, and Fig. 9(b) presents the result after the adaptive threshold. At the same time, we have eliminated the fixed noise through the interframe relation. Results show that the complex clouds disappeared, but there are still many residual candidate points, which are mostly the random noises. We relocate the candidate points to the original image, and then the energy concentration criterion is adopted to eliminate false alarms. The final result is shown in Fig. 9(c) with only a few candidate points. It can determine that the candidate on the lower right is the real target by means of multiframe accumulation. The following analysis is low-order recursive detection in the condition of a moving base.

B. Low-Order Recursive Method

When the candidate points are to be matched, the only information available is the gray value and motion information. The output frame frequency of the long-wave infrared detector can reach 100 fps, while the point target can bear a constant velocity in a very short time. The moving base detector rotates uniformly in a horizontal direction, so the number of pixels moving in the interframe relation of the target can be maintained at a stable value. The horizontal and vertical coordinates of the target in the kth frame are $x(k)$ and $y(k)$. ϕ_x, ϕ_y are the sensor perturbation errors in the x and y directions, while the perturbation error is usually not more than 1 pixel ($\phi_x, \phi_y \in [-1, 1]$). We can obtain the following formula:

$$\begin{cases} x(k+1) - [x(k) + \phi_x] \approx \text{constant} \\ y(k+1) - [y(k) + \phi_y] \approx \text{constant} \end{cases}. \tag{26}$$

The absolute energy of the point target in a short time is also stable, so the energy feature of the target can be another important feature of the data association in the interframe relation. We take the sum of the difference between 3×3 pixels of the target area and the mean of the local background area as the absolute energy value of the target, as shown in Eq. (27). Note that we also need to relocate the candidate points to the original image because the binarization has been made before. In order to avoid the interference of background clutter, the candidate targets are associated if the energy similarity in the interframe relation is more than 70%,

$$\mathrm{Energy} = \sum_{x=-1}^{1} \sum_{y=-1}^{1} |\mu_t(x, y) - \overline{\mu_b}|. \tag{27}$$

We use the low-order recursive method in the five adjacent frames for target recognition, like a pipeline. Data association is performed when the motion and energy of five candidate targets are all matched, as presented in Fig. 10. The low-order

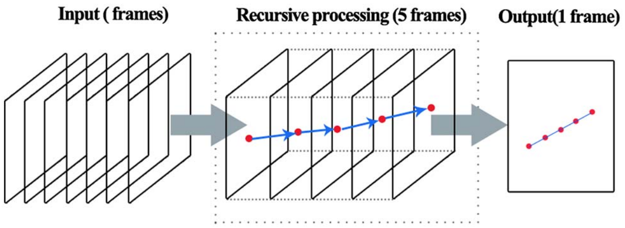


Fig. 10. Diagram of low-order recursive method. First we input the image sequence (100 fps), and then the recursive processing is performed in five adjacent frames to obtain a trajectory output of each of the five frames.

recursive method is simple and reliable in a moving base detector.

4. EXPERIMENTS

A. Data Acquisition Equipment

In order to obtain the true and reliable image resources, the Stirling refrigerator long-wave IRFPA detector is set up in the laboratory to capture aircraft at long distances, as illustrated in Fig. 11. The specific parameters of the detector and optical system are presented in Table 1. The detector adopts the working method of horizontal sweep. The ADS-B global navigation system is employed to monitor civil aviation flight information within 300 km, including the speed, latitude, longitude, and altitude of the aircraft.

B. Processed Results of Proposed Algorithm

In Fig. 12(a) there are three frames extracted from three typical infrared sequences of point targets, including the sky background, the clouds background, and the buildings. Image processing was finished in MATLAB R2014a, and PC configuration is i7- 4790 CPU (3.60 GHz), 8 GB main memory. Figure 12(b) displays the result of the multiscale morphological transform and the adaptive threshold. The fixed noise was eliminated before. The candidate points are mainly high-frequency random noises, which have disappeared after using the EC criterion. There remain about six candidate points in each frame, which are marked in red as shown in Fig. 12(c). The trajectories of each point target are obtained by employing the low-order recursive correlation as presented in Fig. 12(d). Figure 12(e) illustrates the enlarged areas of detected point targets, which distribute at different scales, respectively, including three pixels, two pixels and four pixels. Taking sequence 3 as an example, the ADS-B receiver provides aircraft indicator diagram, which is shown in Fig. 13. After observation and analysis, it can be determined the detected target is B3908 (Airline 1), and 50.3 km away from the detector, calculating through the geodetic coordinate system.

The curves of energy concentration and SNR of the point targets in sequence 3 are summarized in Fig. 14. In the whole moving phase of the point target, the SNR varies between 0.2–2.7, which is hardly visible to naked eyes. The maximum energy concentration of the real point target is 0.75, while the actual value is a little smaller than the theoretical value of 0.842. The reason is that the complex atmospheric composition and the system errors will weaken the energy of point targets in the actual imaging process.

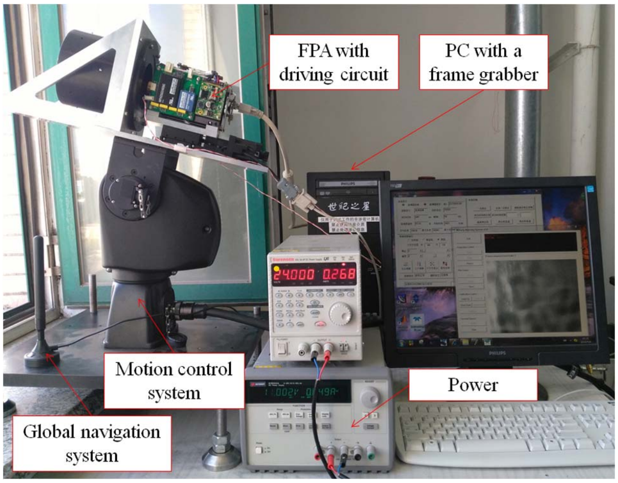


Fig. 11. Data acquisition equipment.

Table 1. Specific Parameters

Parameter	Value
Working wavelength (μm)	7.7–11.3
Resolution	320 × 256
Element size (μm)	30
Output digits	14
Full frame frequency (fps)	100
Focal length (mm)	38
Effective field	14.40 × 11.54

C. Comparison with Other Algorithms

In order to further validate the effectiveness of the proposed algorithm, we compare with common algorithms including the max-median filtering, difference of Gaussian (DoG), BM3D, and GMM. The target detection probability (P_d), the false alarm probability (P_{fa}), and the running time of the algorithm are selected as the evaluation indexes of the results, which are defined as follows:

$$P_d = (N_c/N_t) \times 100\%, P_{\mathrm{fa}} = [N_f/(N_f + N_t)] \times 100\%, \tag{28}$$

where N_c is the number of detected true points, N_f is the number of false alarms, and N_t is the total number of point targets.

There are totally 1507 point targets in the test sequence, including the simple sky background, the complex clouds background, and the buildings. Our method successfully detected 1504 point targets with three false alarms. Therefore, the detection probability of the proposed method reaches 99.8%, and the false alarm probability is 0.2%. The max-median filter detected 1174 point targets with 285 false alarms. The BM3D and GMM methods detected more point targets than the max-median filter and DoG methods; they detected are 1433 and 1418 point targets, respectively. But the number of false alarms of the four comparison algorithms is above 100.

For a more intuitive comparison, we introduce an evaluation index named figure of merit (FoM). This is defined from [19],

$$\mathrm{FoM} = P_d\left(1 - \frac{3P_{\mathrm{fa}}}{1 + 3P_{\mathrm{fa}}}\right). \tag{29}$$

FoM is a normalized parameter, such that $0 < \mathrm{FoM} < 1$. A FoM close to 1 indicates near-perfect performance.

The statistical results of different algorithms are shown in Table 2.

As can be seen from the performance table, max-median filtering and DoG are simple, but the false alarm probability is higher, which leads to a lower FoM. BM3D and GMM are very effective at background suppression, but are time-consuming. The proposed algorithm has high detection probability and low

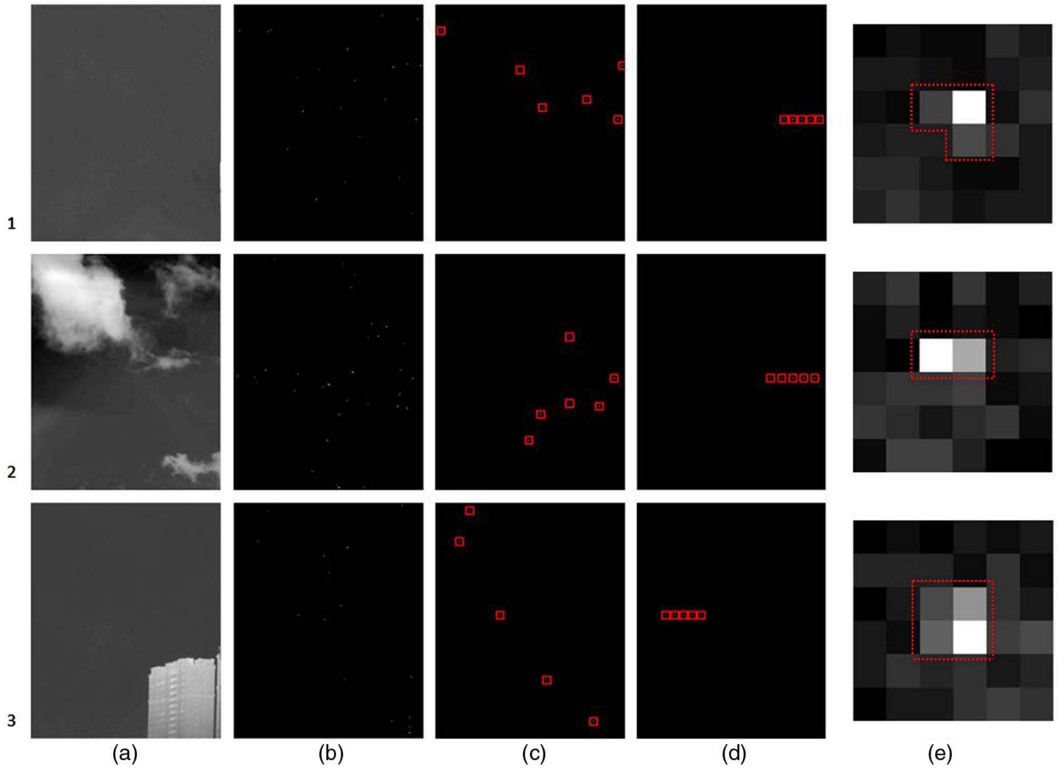
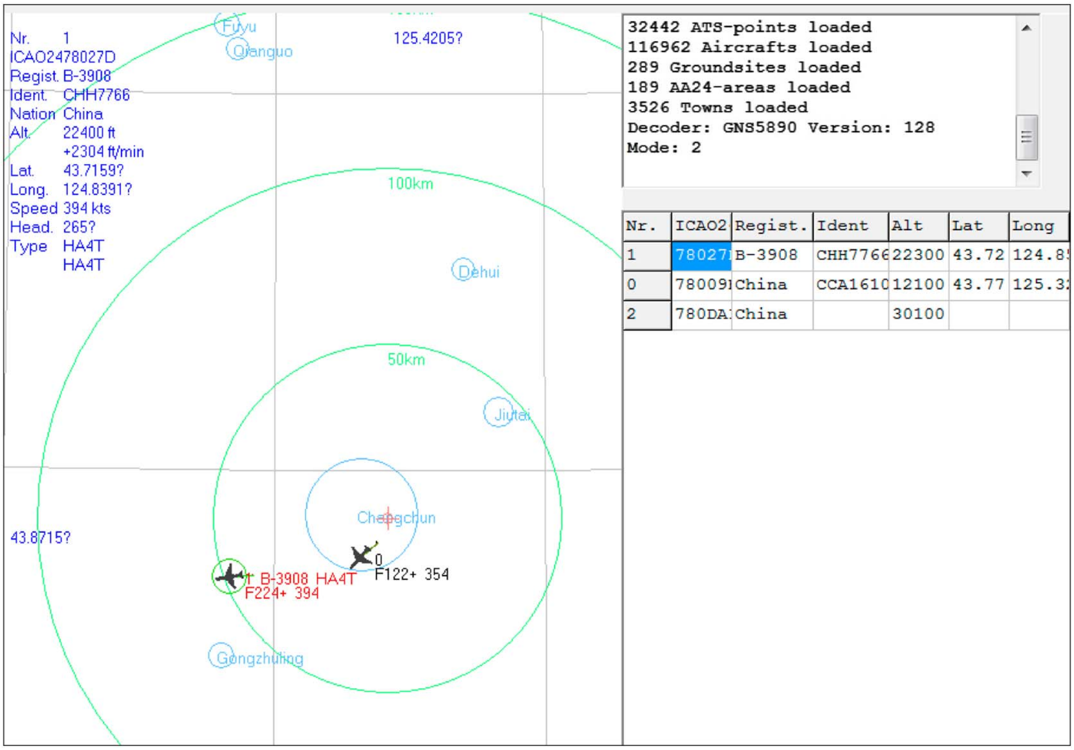


Fig. 12. Three typical frames of the point target, including the sky background, the clouds background, and the buildings. (a) Original image; (b) CFAR result; (c) result of energy concentration threshold; (d) result of the low-order recursive correlation; (e) enlarged area of the detected point target, which has been marked in red.



Nr. 1
ICAO24 78027D
Regist. B-3908
Ident. CHH7766
Nation China
Alt. 22400 ft
+2304 ft/min
Lat. 43.7159?
Long. 124.8391?
Speed 394 kts
Head. 265?
Type HA4T
HA4T

```
32442 ATS-points loaded
116962 Aircrafts loaded
289 Groundsites loaded
189 AA24-areas loaded
3526 Towns loaded
Decoder: GNS5890 Version: 128
Mode: 2
```

Nr.	ICAO2	Regist.	Ident	Alt	Lat	Long
1	78027	B-3908	CHH7766	22300	43.72	124.8
0	78009	China	CCA1610	12100	43.77	125.3
2	780DA	China		30100		

Fig. 13. Indicator diagram provided by ADS-B receiver. The location of the acquisition device is in Changchun, China, which is marked with a red cross. The infrared detector faces the south, and there is an east–west course ahead where it can detect the airliners leaving from Changchun Longjia International Airport.

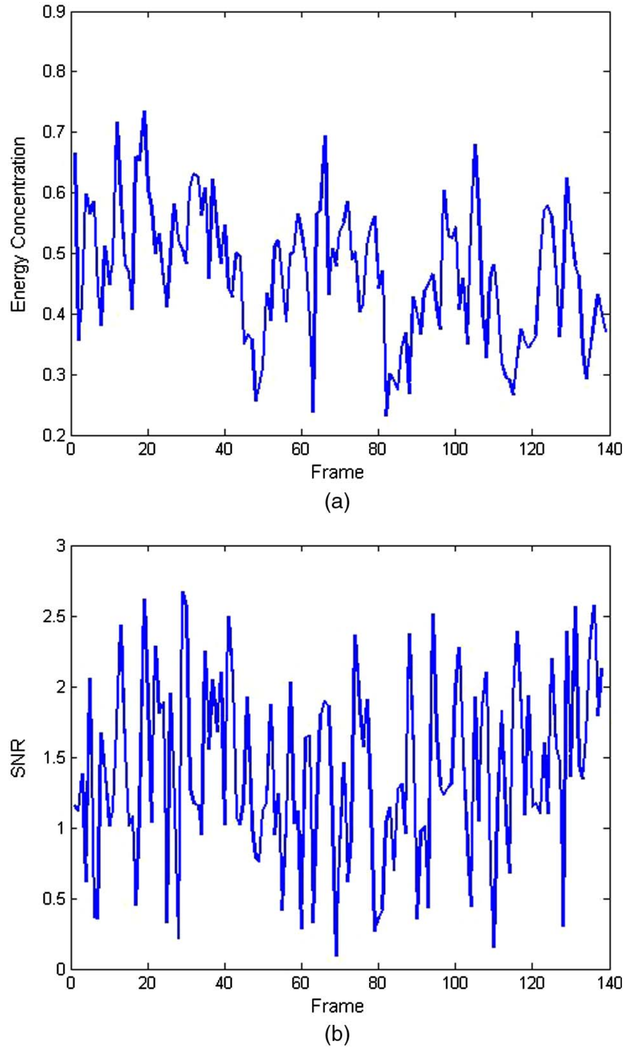


Fig. 14. (a) Curves of energy concentration and (b) SNR of the point target in the sequence 3 detected by the proposed algorithm.

Table 2. Comparison of Algorithms

Method	$P_d/\%$	$P_{\mathrm{fa}}/\%$	FoM	Running Time/s
Max-median filter	77.9	15.9	0.53	0.45
DoG	82.2	10.2	0.63	0.61
BM3D	95.1	6.9	0.79	3.10
GMMO	94.1	7.2	0.77	2.98
Our method	99.8	0.2	0.94	0.32

false alarm rate. At the same time, the proposed algorithm meets the real-time requirements.

5. CONCLUSION

Starting with the analysis of morphological characteristics and energy concentration of dim point targets which are less than 3 × 3 pixels in size, we have proposed a point target detection method based on the omnidirectional multiscale morphological method and EC criterion. First we use the eight-direction 5 × 5 structural elements to detect all the possible targets at different scales. Next, an adaptive threshold is employed to obtain the ROI of the target and improve SNR. Then, we use the EC criterion to eliminate random noise. Finally, the trajectory of the point target is obtained after the low-order recursive correlation. The experimental results show that the proposed algorithm is effective and has a good inhibition effect on the complex background, and the running time of the algorithm meets the real-time requirements for engineering.

Funding. National Natural Science Foundation of China (NSFC) (61675202).

REFERENCES

1. S. H. Revital, H. Ofer, and R. R. Stanley, “Parametric temporal compression of infrared imagery sequences containing a slow-moving point target,” Appl. Opt. **55**, 1151–1163 (2016).
2. C. Yang, J. Ma, and S. Qi, “Directional support value of Gaussian transformation for infrared small target detection,” Appl. Opt. **54**, 2255–2265 (2015).
3. R. Liu, D. Wang, and P. Jia, “Overview on small target detection technology in infrared image,” Laser Optoelectron. Prog. **05**, 44–52 (2016).
4. C. Wang and W. Tian, “Small infrared target fusion detection based on support vector machines in the wavelet domain,” Opt. Eng. **45**, 076401 (2006).
5. X. Li and Y. Zhao, “Approach to dim and small target detection based on fuzzy classification,” Opt. Precis. Eng. **17**, 2312–2320 (2009).
6. S. Deshpande, M. Er, and R. Venkateswarlu, “Max-mean and max-median filters for detection of small-targets,” Proc. SPIE **3809**, 74–83 (1999).
7. Z. A. Mukesh, N. S. Merchant, and D. B. Uduy, “Air-borne approaching target detection and tracking in infrared image sequence,” in *Proceedings of IEEE Conference on Image Processing* (IEEE, 2004), pp. 1025–1028.
8. H. Fan and C. Wen, “Two-dimensional adaptive filtering based on projection algorithm,” IEEE Trans. Acoust. Speech Signal Process. **52**, 832–838 (2004).
9. C. L. P. Chen, H. Li, and T. Y. Wei, “A local contrast method for small infrared target detection,” IEEE T. Geosci. Remote Sens. **52**, 574–581 (2014).
10. L. Genin, F. Champagnat, and G. L. Besnerais, “Background first- and second-order modeling for point target detection,” Appl. Opt. **51**, 7701–7713 (2012).
11. F. Chen and W. Wang, “Target recognition in clutter scene based on wavelet transform,” Opt. Commun. **282**, 523–526 (2009).
12. F. A. Sadjadi, “Infrared target detection with probability density functions of wavelet transform subbands,” Appl. Opt. **43**, 315–323 (2004).
13. M. Zeng, J. Li, and Z. Peng, “The design of top-hat morphological filter and application to infrared target detection,” Infrared Phys. Technol. **48**, 67–76 (2006).
14. 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).
15. 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).
16. C. Corbane, E. Pecoul, L. Demagistri, and M. Petit, “Fully automated procedure for ship detection using optical satellite imagery,” Proc. SPIE **7150**, 71500R (2008).
17. X. Bai, F. Zhou, and B. Xue, “Image enhancement using multi scale image features extracted by top-hat transform,” Opt. Laser Technol. **44**, 328–336 (2012).
18. D. Das, S. Mukhopadhyay, and S. R. Sai-Praveen, “Multi-scale contrast enhancement of oriented features in 2D images using directional morphology,” Opt. Laser Technol. **87**, 51–63 (2017).
19. G. A. Page, B. D. Carroll, A. Prrat, and P. N. Randall, “Long-range target detection algorithms for infrared search and track,” Proc. SPIE **3698**, 48–57 (1999).