An infrared dim target detection algorithm based on density peak search and region consistency

Baohua Zhang1,3 · Jinhui Zhu1 · Xiaoqi Lu2,3 · Yu Gu1,3 · Jianjun Li1,3 · Xin Liu1,3 · Ming Zhang1,3

Received: 9 March 2021 / Accepted: 8 June 2021 / Published online: 14 July 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
To suppress background clutter and improve detection accuracy, we propose a dim target detection algorithm based on density peak search and region consistency. A density peak search algorithm is first applied to extract candidate targets, and these are then classified and marked according to the local mosaic probability factor, which is important in order to suppress the background clutter and accurately strip the candidate target region from the background. Based on the regional stability of the dim targets, local mosaic gradient factors are used to screen real targets from candidates, and a facet kernel filter is used to extract the irregular contours of dim targets with the aim of enhancing them. Our experimental results show that compared with existing algorithms, the proposed method has better detection accuracy and robustness in various complex scenarios.

Keywords Small target detection · Density peak search · Local mosaic model · Infrared images

1 Introduction

An infrared search and tracking system (IRST) has the advantages of good concealment and robust anti-interference ability, and this approach is widely used in many applications such as military early warning systems, precision guidance and remote sensing (Zeng et al. 2006; Wang et al. 2019). The method of infrared target detection is a key aspect of this technology (Chen et al. 2014). However, there are many challenges associated with dim target detection; for example, the target may only contain dozens of pixels in each frame, or may lack texture and shape information, and the signal-to-clutter ratio in each frame may
be very low (Ye et al. 2017; Huang et al. 2019b, a). Dim targets are often submerged in complex backgrounds (such as cloud edges, ocean waves, and high-brightness noise).

Infrared dim target detection methods can be divided into two types: single-frame (Dai and Wu 2017) and multi-frame methods (Lv et al. 2018; Wan et al. 2016; Zhang et al. 2019). To detect such targets, Lv used a multi-frame iteration process to estimate the background. However, optimisation of the coefficients of the filter required a high number of iterations, so the efficiency of this algorithm was poor (Lv et al. 2018). Wan proposed a detection method based on the inter-frame difference, in which the differences between adjacent frames were used to locate the target, but the contours extracted for the target were incomplete, and the interval between the frames severely affected the detection accuracy (Wan et al. 2016). Zhang used a quaternion discrete cosine transform to extract the features of dim targets from multi-frame images, and this approach made full use of the spatiotemporal features of four channels. However, target misalignment between multiple frames reduced the detection accuracy, and the algorithm had a high level of complexity and poor real-time performance (Zhang et al. 2019). Single-frame methods have many advantages, such as low complexity, high execution efficiency and strong real-time performance, and have therefore attracted more attention.

Most traditional single-frame methods, such as those involving morphological filtering (Zeng et al. 2006) or probability statistical filtering (Deshpande et al. 1999; Soni et al. 1993; Tomasi and Manduchi 1998; Barnett, 1989), detect targets by suppressing the background. Although these methods are easy to implement in real time, their performance is greatly reduced by clutter and interference when a complex background is present. By exploiting the characteristics of the human visual system, Chen proposed a local contrast measure (LCM) that detected a target based on the difference in brightness between the target and the adjacent areas (Chen et al. 2014). This method was not suitable for scenes with a high level of background brightness, and a large number of improved methods were subsequently developed, such as those based on multiscale patch-based contrast measures (Wei et al. 2016; Nie et al. 2018). These methods were unable to extract the features of the target, and yielded high false positive rates caused by high-intensity clutter and high-brightness noise.

Huang proposed a dim target detection method based on maximum grey region growing, which segmented the target and background by selecting seed growing points (Huang et al. 2019b, a). However, this method was only suitable for scenarios in which the number of targets was much smaller than the number of candidates. Xia combined a random walk algorithm with the local contrast features to achieve efficient target detection, but the selected seed points in the algorithm needed to be labelled and classified, which was inefficient (Xia et al. 2018). Qin introduced the concept of facet kernel filtering to suppress clutter and enhance the targets(Qin et al. 2019), the selection of candidate points in this method was simple, but due to its dependence on a single characteristic of the target, the false alarm rate was high for complex backgrounds such as those containing strong clutter and bright edges.

In order to achieve robust detection of dim targets in images with complex backgrounds, we propose an infrared dim target detection algorithm based on a density peak search and region consistency. The main contributions of our algorithm are as follows.

We approach the problem of dim target detection as an abnormal point detection problem, and candidate target points are initially screened out using an unsupervised method based on an improved random walker (RW) algorithm. Local mosaic template regions are then constructed with the candidate target point as the centre pixel, and candidate targets are classified and labelled locally to suppress clutter. Based on the gradient difference and the directional
consistency of dim infrared targets, a local mosaic probability factor (LMPF) is used to separate the target from noise and clutter, and targets are enhanced using a local mosaic gradient factor (LMGF). Our experimental results show that compared with existing algorithms, the proposed detection algorithm yields better performance for images containing complex scenes.

2 Proposed method

The details of the proposed method are as follows. Candidate targets are first identified based on their abnormal attributes, using a density peak search algorithm. They are then selected as the central point, and an improved RW algorithm is then applied to construct the local mosaic template region, thus reducing the feature search range. To suppress clutter, the LMPF is used to classify and mark candidate targets, which is a critical step in separating targets from clutter. Based on the direction consistency of the dim targets, the LMGF is then used to enhance the edges of the targets, which allows them to be extracted from the background. Figure 1 shows a flow chart for the proposed method.

2.1 Density peak search algorithm

The density peak search algorithm (DPeak) is a granular computing model for which theoretical basis rests on two assumptions: (1) the central point is surrounded by neighbouring data points with a lower local density; and (2) two central points are not adjacent to each other (Wang et al. 2017).

For any pixel \( i \), the DPeak model needs to calculate the local density \( \rho_i \) and the distance to the nearest correlation distance \( \delta_i \)

\[
\rho_i = f_i, \tag{1}
\]

\[
\delta_i = \min_{i: \rho_i > \rho_j} d_{ij}, \tag{2}
\]

where \( f_i \) represents the gray value of the pixel, the highest density point \( \delta_i = \max_{j} (d_{ij}) \),

\[
d_{ij} = \left| x_i - x_j \right| + \left| y_i - y_j \right|, \tag{3}
\]

\[
Y = \rho \times \delta, \tag{4}
\]

![Fig. 1 The entire flow chart of the proposed method](https://example.com/fig1.png)
where \( d_{ij} \) is the Manhattan distance, and then calculate the joint feature factor \( \Upsilon \) pixel by pixel, and put them in the queue \( Q \) in descending order, and extract the first \( m \) pixels as candidate target points.

The grey value of the target is normally higher than that of the local neighbouring pixels, meaning that the two assumptions of the DPeak model are applicable to the characteristics of dim targets. The DPeak model can therefore reduce the search range of the target and reduce the calculation time.

### 2.2 Local mosaic probability factor (LMPF) and local mosaic gradient factor (LMGF)

The textural structure of the dim target is essentially different from the clutter. Figure 2 shows a comparison of the gradient distributions: Fig. 2a shows an infrared image, and Fig. 2b shows an enlarged view of the target area in Fig. 2a. The upper image shows the gradient distribution of the background clutter, and the lower image shows the gradient distribution of the target and its neighbourhood. It can be seen that the gradient direction of the target neighbourhood is consistent and the amplitude is uniform; in comparison, the gradient direction for the background clutter is messy, and the amplitude difference is large. Figures 2c and d show three-dimensional gradient distribution maps for the background clutter and the target, respectively. It can be seen that the dim target has different characteristics, such as higher grey values, a higher gradient and a higher density than the adjacent background area. The aim of candidate point screening is to judge whether each point represents a target, based on these local attributes.

The roaming window of the RW algorithm is a rectangular frame, and the outermost pixels are usually classified as background. In a multi-target scene, if the targets are adjacent or touching, these neighbouring targets will be classified as background (Qin et al. 2019), meaning that parts of them will be missed. In order to accurately extract the target from the background, we improve the RW algorithm by introducing a more efficient marking model. Unlike the global traversal method, our improved model traverses only the candidate target area. The roaming window is also changed to an adaptive geometric window, which more closely reflects the contours of the target. On this basis, the outer area of the geometric window (background) is marked as 0-type, and the central area (target) is marked as 1-type, which is useful in terms of accurately segmenting the target. In Fig. 3a, the red frame indicates the central region, and the area between the yellow and red frames represents the transition region. The outer area lies between the blue and yellow frames. The candidate target is used as the central pixel of the local mosaic model, and is marked as a 1-class pixel, while pixels in the outer area of the mosaic model are marked as 0-class.
An infrared dim target detection algorithm based on density…

Since most dim targets have dimensions of less than 9 × 9 pixels, the size of the template is set to 11 × 11 pixels. The shape of the marked pixels in the outer region is close to the target contour, and can accurately approximate the true distribution of the target. The transition region and the outer region are defined as follows:

\[ MR = (CR \oplus \text{shape}) - CR, \]

\[ OR = (MR \oplus \text{shape}) - MR, \]

CR, MR and OR refer to the central region, the transition region and the outer region respectively. \( \oplus \) refers to the expansion operation, and shape refers to the diamond-shaped structural element.

In the local mosaic model, only the central pixel is marked as 1-class, while pixels close to the outer region are marked as 0-class. The pixels in the transition region are classified based on their similarity to the grey level of the central pixel.

To create a new local mosaic region, the transition region and the outer region are updated using Eqs. (5) and (6) when the candidate target points have been obtained. In our approach, the outer area is selected as the background, and the targets are screened using the LMPF, an indicator that compares the probability value for the candidate target region with the probability value for the background.

\[ \text{LMPF} = \begin{cases} \frac{\text{MVP(central area)}}{\text{MVP(OR)}}, & \text{if pixels} \notin Q \text{ and label} = 1 \\ 0, & \text{otherwise} \end{cases}, \]

In Eq. (7), MVP (mean value of probability) is the average of pixel probabilities, and pixels are pixels expanded by candidate target pixels in the central area. LMPF is the gray value of the central area.

It is difficult to use the LMPF to achieve robust detection of dim targets in highlight scenes, especially in low-contrast images. We therefore introduce the LMGF to expand the difference between the target and the background and improve the accuracy of our method. As shown in Fig. 3b, the template is divided into eight parts based on a three-level nested model: these are labelled as S1, S2, S3, S4 (where S represents the central area) and T1, T2, T3, T4 (where T represents the outer area). Here, the LMGF is defined as follows:

\[ \text{LMGF} = \begin{cases} \text{MVP(pixels)}_1, & \text{if} \text{ pixels} \notin Q \text{ and label} = 1 \\ 0, & \text{otherwise} \end{cases}, \]
in Eq. (8), $\text{grad}_{si}$ is the gradient element in the area of $S_i$ ($i = 1, 2, 3, 4$), $N_{si}$ is the number of gradients in $S_i$, $g_{si}$ is the average magnitude of the gradient in $S_i$, $G_{s\text{max}}$ and $G_{s\text{min}}$ are the maximum and minimum values of the average gradient of $S_i$ ($i = 1, 2, 3, 4$).

$$g_{si} = \frac{1}{N_{si}} \sum_{k=1}^{N_{si}} \|\text{grad}_{si}\|^2,$$

$$G_{s\text{max}} = \max_{1 \leq i \leq 4} g_{si},$$

$$G_{s\text{min}} = \min_{1 \leq i \leq 4} g_{si},$$

where, $\text{grad}_{ti}$ is the gradient element in the area of $T_i$ ($i = 1, 2, 3, 4$), $N_{ti}$ is the number of gradients in $T_i$, $g_{ti}$ is the average magnitude of the gradient in $T_i$, $G_{t\text{max}}$ and $G_{t\text{min}}$ are the maximum and minimum values of the average gradient of $T_i$ ($i = 1, 2, 3, 4$).

$$g_{ti} = \frac{1}{N_{ti}} \sum_{j=1}^{N_{ti}} \|\text{grad}_{ti}\|^2,$$

$$G_{t\text{max}} = \max_{1 \leq i \leq 4} g_{ti},$$

$$G_{t\text{min}} = \min_{1 \leq i \leq 4} g_{ti},$$

where, the LMGF is the gray value of the central area. The direction of the gradient in the neighbourhood of the dim target is consistent and the amplitudes are similar, meaning that the gradient direction and amplitude of the clutter are significantly different. The target can therefore be identified based on this information.

### 2.3 Target extraction and enhancement

The input image is filtered using $5 \times 5$ kernel matrix $K$ (Qi et al. 2016) to get the enhanced image, $P_i$ is original image, as shown in Eq. (15):

$$K = \begin{pmatrix}
-4 & -1 & 0 & -1 & -4 \\
-1 & 2 & 3 & 2 & -1 \\
0 & 3 & 4 & 3 & 0 \\
-1 & 2 & 3 & 2 & -1 \\
-4 & -1 & 0 & -1 & -4
\end{pmatrix}, \quad P_e = P_i \ast K,$$

where, the target map $P_m$ is obtained according to Eq. (16), $P_m$ and $P_e$ are calculated by the Hadamard product to get a weighted image $P_w$, $P_w$ is called “Weighted”:

$$P_m = \text{LMPF} \ast \text{LMGF},$$
\[ P_w = P_m \circ P_e, \]  

\[ T_h = \lambda \cdot \max(P_w), \]  

in the enhanced image \( P_e \), both the targets and the clutter are enhanced, while \( P_m \) only maps the target area in the proposed method. Therefore, the weighted image \( P_w \) is more approximate the target contour, indicating that the proposed algorithm can effectively suppress the clutter and enhance the targets. \( T_h \) is a simple threshold the value range of \( \lambda \) is 0.6~0.9.

### 2.4 Proposed detection method

As shown in Algorithm 1, the proposed method selects candidate targets by constructing a feature space for the local density-closest correlation distance, and then constructs the local mosaic template area with the candidate target point as the central pixel. In order to suppress clutter, the candidate target regions are classified and labelled based on the LMPF, and the real targets are finally selected from the candidate targets based on the LMGF.
Algorithm 1 The Proposed Detection Method

Input: infrared image $P_i$, parameters $m$ and $\mu$.

Output: Detection result.

1: for $i = 1$ to $N$ do (the number of pixels is $N$)
2: For $i$, obtain $\rho_i$ according to (1).
3: For $i$, calculate $\delta_i$ according to (2).
4: Obtain 2D Decision Map and Compute $Y$ according to (4).
5: end for

6: Store and Sort all $Y$ in the queue $Q$ in descending order
7: Find the first $m$ pixels as candidate targets and record their positions.
8: for $j = 1$ to $m$ do
9: Select the candidate target pixel of $Q[j]$ and labeled as 1
10: Construct the region of $11 \times 11$ pixels.
11: Compute the LMPF according to (7).
12: Compute the LMGF according to (14).
13: Mark pixels segmented as label 1.
14:end for

15: Compute the enhanced map $P_e$ according to (15).
16: Compute the weighted map $P_m$ according to (16).
17: Compute the weighted map $P_w$ according to (17).
18: Segment targets from weighted map using adaptive thresholding.

3 Experiments and analysis

In this section, the experiments are performed to verify the effectiveness of the proposed method. The experimental computing environment in this paper is 3.40 GHz Intel i7-3770 CPU processor, 8 GB memory, and the test software is MATLAB 2018b.

3.1 Experimental setup

(1) Datasets: To verify the robustness and accuracy of the proposed method, six sets of infrared image sequences with different complex backgrounds were selected. Table 1 gives detailed information on these six sequences. Each sequence has a different background: Frame 1 contains the edge of the sky and irregular clouds; Frame 2 has a glare
background (Huang et al. 2019b, a); Frame 3 contains a large amount of ring-shaped point clutter; in Frame 4, the target is surrounded by clouds, which form a great deal of clutter (Qin et al. 2019); Frame 5 shows a dual-target image sequence with high-intensity sky and sea wave clutter as the background; and Frame 6 is a four-target image sequence with low intensity (Xia et al. 2018) (Table 2).

(2) Baseline Methods: This paper selects the existing 9 methods as comparison methods, including Top-Hat (Zeng et al. 2006), Fast-Saliency (Qi et al. 2016), VARD (Nasiri and Chehresa, 2017), LCM (Chen et al. 2014), MPCM (Wei et al. 2016), AAGD (Deng et al. 2016), NRAM (Zhang et al. 2018a, b), Max-Median (Deshpande et al. 1999) and LIG (Zhang et al. 2018a, b).

(3) Evaluation metrics:

\[
BSF = \frac{\sigma_{in}}{\sigma_{out}}, \quad (19)
\]

\[
CG = \frac{C_{out}}{C_{in}}, \quad (20)
\]

\(\sigma_{in}\) and \(\sigma_{out}\) represent the intensity standard deviation of the source image and the processed image respectively; \(C_{in}\) and \(C_{out}\) are the contrast between the source image and the processed image. The contrast is defined as Eq. (21):

Table 1: Details of six infrared target datasets

| Frame | Frame number | Target size | Scene | Clutter |
|-------|--------------|-------------|-------|---------|
| 1     | 100          | 3×3         | Cloudy sky | Strong edges of cloudy shape |
| 2     | 66           | 2×5         | Sky     | bright and heavy noise |
| 3     | 80           | 3×4         | Sky     | a circular clutter |
| 4     | 60           | 4×5         | Cloudy sky | Complicated cloud clutters |
| 5     | 67           | 4×9 and 4×12 | Sea | Bright noise and lined shape clutters |
| 6     | 85           | 5×7, 5×6, 5×5 and 4×4 | Sea and sky | Heavy noise |

Table 2: Parameter settings of different methods

| No. | Methods | Parameter settings |
|-----|---------|--------------------|
| 1   | Top-Hat | Local window size: 5×5 |
| 2   | Fast-Saliency | Gaussian filter size: 5×5, \(\sigma = 1.5\) |
| 3   | VARD    | External, Guard and Internal Patch size: 15×15, 11×11, 7×7 |
| 4   | LCM     | Local window size: 8×8, sliding steps of row and column: 4 |
| 5   | MPCM    | \(L = 3\), Local window size: 3×3, 5×5, 7×7 |
| 6   | AAGD    | Internal and external window size:3×3,19×19 |
| 7   | NRAM    | Patch size:50×50, sliding step:10, \(\lambda = 1/\sqrt{\min (m,n)}\), \(\mu^0 = \sqrt{\min (m,n)}, \sigma = \sqrt{\min (m,n)/2.5}\) |
| 8   | Max-Median | Local window size: 15×15 |
| 9   | LIG     | \(K=0.2, n=19\) |
| 10  | Proposed | window size: 11×11, \(\mu = 100\) |
where, $f_t$ is the maximum intensity value of the target area and $f_b$ is the average value of the background area. In this paper, we use BSF and CG as evaluation indicators. BSF can be used to evaluate the ability of the algorithm to suppress clutter globally, and its value is proportional to the suppression effect, while CG can be used to evaluate the effect of target enhancement, and its value is proportional to this effect.

$$C = |f_t - f_b|,$$  \hspace{1cm}  \text{(21)}

$$TPR = \frac{\text{number of true target detection}}{\text{total number of actual targets}},$$ \hspace{1cm}  \text{(22)}

$$FPR = \frac{\text{number of detected false targets}}{\text{total number of pixels in images}}.$$ \hspace{1cm}  \text{(23)}

TPR and FPR represent the true positive and false positive rates, respectively, and are used to draw the ROC curve and evaluate the algorithm: the closer the position of this curve to the upper left corner of the graph, the better the ability of the algorithm.

### 3.2 Qualitative comparisons

Figure 4 shows a comparison of the detection results from the source images based on the LMPF, the LMGF and the weighted method, the weighted method represents $P_w$. It can be seen that for Frames 1, 2 and 4, the detection results based on the LMPF still contain point shadow clutter. This indicates that the clutter suppression ability of the LMGF is better than that of the LMPF. However, in terms of target edge retention and target enhancement, the LMPF yields better performance than the LMGF, except for Frame 5; due to the influence of the strong light and sea clutter, the effectiveness of the LMPF is slightly lower than that of the LMGF in this case.

In order to verify the performance of this algorithm, we used six datasets in our detection experiments, including both single-target and multi-target image sequences. To qualitatively analyse the effect of each step of the proposed method, each source image was processed using LMPF, LMGF and the weighted model, respectively, and the BSF and CG were calculated for the outputs. Table 3 shows the results. It can be seen that the results of the weighted model were better than those of the LMPF or the LMGF models, meaning that the weighted model can enhance the target and suppress background clutter more effectively.

In order to evaluate the performance of our algorithm, six datasets were used as experimental data, including both single-target and multi-target image sequences. Nine alternative methods were selected for comparison: the top-hat, max-median, fast-saliency, LCM, MPCM, VARD, AAGD, NRAM and LIG methods. Figures 5 and 6 show the detection results for these six datasets. The four datasets in Fig. 5 all contained single-target sequences, and their backgrounds included high-intensity light, cloud clutter and complex irregular clutter. The two datasets shown in Fig. 6 were two-target and four-target image sequences, respectively, and were used to verify the multi-target detection ability of our algorithm.

The source images in Fig. 5 are four sets of image sequences showing areas of sky. In Frame 1, the background is the sky and irregular cloud edges, forming a lot of clutter; Frame 2 has a light background with strong intensity, and the background brightness
is similar to that of the target; the background of Frame 3 contains a great deal of ring-shaped point clutter; and the target in Frame 4 is surrounded by clouds, which create clutter interference.

The detection results from the top-hat and the max-median methods contain a great deal of clutter, indicating that they have poor effectiveness in terms of background suppression, and the top-hat method loses the target in Frame 2. The detection results from the LCM and the MPCM methods not only contain background clutter, but also have
a high false alarm rate. In contrast, there are few background clusters in the results of the fast-saliency method. The other comparison methods also show some problems: the AAGD method gives a small amount of linear clutter in Frames 1 and 4; the NRAM method loses the target in Frame 1; and the LIG method gives a small number of background clusters in Frame 4. In general, our proposed method effectively suppresses the background clutter and preserves the contours of the target, and gives the best detection performance.

The experimental dataset shown in Fig. 6 contains two-target and four-target image sequences with the sea and sky as background, and these are used to verify the multi-target detection performance of our approach. Frame 5 is a dual-target image sequence with a strongly lit sky and clutter from sea waves. The background suppression effects from the top-hat, fast-saliency, VARD, AAGD, NRAM and LIG methods are good. Since there is a great deal of clutter in the detection results from the LCM, MPCM and max-median methods, this means that the background suppression effects are poor. Frame 6 is a four-target image sequence with low brightness. The fast-saliency and VARD methods lose the target, and the LCM, max-median and MPCM methods have poor background suppression effects. The targets detected by the NRAM and the LIG methods are blurry, while the contours extracted with the top-hat and AAGD methods are clear. Compared with the alternative methods, our approach has obvious advantages in terms of background clutter suppression and target enhancement, and yields better performance for all of the datasets.

### 3.3 Quantitative comparisons

The BSF, CG, and ROC curves and the average running time of the algorithm were used to further evaluate the detection performance of our method. The BSF was used to evaluate the background suppression capability, while the CG was applied to evaluate the target enhancement capability of the algorithm. Table 4 shows the results for the BSF, CG and average running time for our algorithm.

In Table 4, the BSF of the proposed method is best for Frames 1, 2, 4, and 5, while the NRAM method performs best for Frames 3 and 4. This shows that the background suppression ability of our method is stronger than that of the alternative methods.

| Methods | LMPF | LMGF | Weighted |
|---------|------|------|----------|
| Frame 1 | 5.640 | 5.754 | 9.480 |
| Frame 2 | 4.955 | 9.277 | 12.370 |
| Frame 3 | 8.396 | 9.069 | 12.003 |
| Frame 4 | 4.474 | 5.328 | 7.112 |
| Frame 5 | 13.748 | 14.020 | 19.690 |
| Frame 6 | 7.385 | 9.981 | 9.983 |
| Frame 1 | 2.486 | 2.470 | 2.686 |
| Frame 2 | 8.121 | 8.121 | 8.121 |
| Frame 3 | 2.671 | 2.457 | 2.682 |
| Frame 4 | 5.250 | 5.248 | 5.257 |
| Frame 5 | 1.563 | 2.013 | 2.525 |
| Frame 6 | 3.344 | 2.297 | 3.564 |

---

The BSF, CG, and ROC curves and the average running time of the algorithm were used to further evaluate the detection performance of our method. The BSF was used to evaluate the background suppression capability, while the CG was applied to evaluate the target enhancement capability of the algorithm. Table 4 shows the results for the BSF, CG and average running time for our algorithm.

In Table 4, the BSF of the proposed method is best for Frames 1, 2, 4, and 5, while the NRAM method performs best for Frames 3 and 4. This shows that the background suppression ability of our method is stronger than that of the alternative methods.
CG is the target local contrast gain. The larger the value of this parameter, the better the background suppression and the target enhancement performance. The CG values from the proposed method were the highest for all image sequences.

In terms of the running time, the fast-saliency method required the shortest time, and AAGD took the shortest time in the remaining five image sequences. The NRAM method required the longest time for all image sequences. The running time for our method was between these two extremes, taking into account the detection accuracy and computing efficiency.

Figure 7 shows a comparison of the ROC curves for these algorithms. It can be clearly seen that the ROC curve for our method is the highest.
3.4 Parameter Sensitivity Analysis

In our method, the target detection problem is treated as a clustering problem with unsupervised learning, meaning that the selection of the number of candidate points is the main parameter that determines the performance. The weights between adjacent pixels strongly affect the location of the target contour, and we use the parameter $\mu$ to determine these weights. In this section, we report the results of experiments performed to analyse the influence of this parameter on the test results.

Since dim targets were often submerged in clutter, we assumed that most of the seed points were false targets and a few were target points, meaning that the target points formed a set of abnormal points. Figure 8 shows TPR diagrams for varying numbers of candidate points $m$ for six groups of image sequences. According to the experimental results, this proposed method sets $m$ to 12.

Figure 9 shows the distribution of the ROC curve for different values of the parameter $\mu$ for the proposed method. It can be seen that if this parameter is set to 100, the results for the ROC curve are best for all the image sequences. If it is set to 200, the curve is slightly lower, and it is set to zero or 300, the target may be missed or false targets may be counted, causing poor detection results. We therefore set this parameter to 100.

4 Conclusion

In this paper, a density peak algorithm is used to traverse each pixel in an image to establish a density-distance spatial decision map. Candidate points are then selected based on a proportional relationship between noise points and dim targets. A local mosaic model is built with the candidate pixels as the central points, and the targets can be screened from the candidate targets based on the mapping relationship between the LMPF and the LMGF, which narrows the target search range and reduces the amount...
Table 4  BSF, CG and average running time of different algorithms under each image frame

| Methods | Proposed | Top-hat | Fast-saliency | VARD | LCM | MPCM | AAGD | NRAM | Max–median | LIG |
|---------|----------|---------|---------------|------|-----|------|------|------|-------------|-----|
| BSF     |          |         |               |      |     |      |      |      |             |     |
| Fra1    | 9.480    | 3.754   | 2.726         | 8.718| 1.356| 5.947| 5.111| 7.432| 6.075       | 8.842|
| Fra2    | 12.370   | 5.190   | 1.686         | 6.229| 0.414| 3.218| 4.090| 7.322| 1.045       | 6.107|
| Fra3    | 12.003   | 1.648   | 2.271         | 5.634| 1.339| 2.514| 2.141|       |             |     |
| Fra4    | 10.099   | 0.723   | 2.613         | 7.578| 0.709| 2.080| 1.656| 7.112| 1.227       | 3.499|
| Fra5    | 19.690   | 9.436   | 9.350         | 17.215| 1.730| 7.819| 12.314| 18.115| 4.407       | 17.257|
| Fra6    | 9.983    | 7.683   | 10.346        | 13.714| 1.065| 5.487| 5.638| 17.825| 4.364       | 16.589|
| CG      |          |         |               |      |     |      |      |      |             |     |
| Fra1    | 2.486    | 2.385   | 2.268         | 2.485| 1.914| 1.634| 2.476| 0.565| 1.502       | 2.486|
| Fra2    | 8.121    | 0.948   | 7.555         | 8.105| 3.368| 6.090| 8.050| 8.089| 4.907       | 8.093|
| Fra3    | 2.682    | 2.473   | 2.605         | 2.325| 1.588| 0.479| 0.698| 2.188| 1.228       | 2.682|
| Fra4    | 5.257    | 4.738   | 1.334         | 5.239| 1.714| 1.752| 2.511| 3.999| 2.088       | 5.244|
| Fra5    | 2.252    | 0.962   | 0.608         | 1.199| 1.264| 1.175| 2.206| 1.943| 1.453       | 2.013|
| Fra6    | 3.564    | 3.497   | 1.701         | 3.005| 1.893| 0.862| 2.444| 1.486| 1.100       | 1.571|
| Time    |          |         |               |      |     |      |      |      |             |     |
| Fra1    | 0.580    | 0.570   | 0.026         | 0.067| 0.116| 0.103| 0.049| 0.830| 0.725       | 0.694|
| Fra2    | 1.140    | 0.520   | 0.093         | 0.073| 0.159| 0.137| 0.059| 3.810| 3.234       | 2.499|
| Fra3    | 1.030    | 0.430   | 0.069         | 0.069| 0.150| 0.126| 0.052| 3.110| 2.658       | 2.213|
| Fra4    | 0.820    | 0.510   | 0.078         | 0.080| 0.152| 0.131| 0.066| 2.040| 1.983       | 1.716|
| Fra5    | 1.250    | 0.470   | 0.106         | 0.082| 0.167| 0.153| 0.064| 5.920| 3.472       | 3.175|
| Fra6    | 0.840    | 0.420   | 0.075         | 0.076| 0.146| 0.133| 0.059| 4.090| 2.577       | 2.562|

The bold font indicates the best result.
of calculation required. The proposed method can also accurately extract targets from multi-target image sequences. Experimental results show that our method performs better than alternative approaches, as it not only suppresses clutter and enhances the target, but also has a wide range of applications. In future work, we intend to develop a new positioning method for the target region, reduce the running time further, and build a robust feature space for screening real targets.
Acknowledgements The authors thank the anonymous reviewers and editors for the very constructive comments. This work was supported by the National Natural Science Foundation of China(61962046,61663036,61841204). Inner Mongolia Outstanding Youth Cultivation Fund (2018JQ02). Inner Mongolia Natural Science Foundation (2015MS0604).

Declarations

Conflict of interest The authors declare no conflicts of interest.

References

Barnett, J.: Statistical analysis of median subtraction filtering with application to point target detection in infrared backgrounds. Proc. SPIE. **1050**, 10–18 (1989)

Chen, C.L.P., Li, H., Wei, Y.T., Xia, T., Tang, Y.Y.: A local contrast method for small infrared target detection. IEEE Trans. Geosci. Remote Sens. **52**, 574–581 (2014)

Dai, Y., Wu, Y.: Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection. IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. **10**, 3752–3767 (2017)

Deng, H., Sun, X., Liu, M., Ye, C., Zhou, X.: Infrared small-target detection using multiscale gray difference weighted image entropy. IEEE Trans. Aerosp. Electron. Syst. **52**, 60–72 (2016)

Deshpande, S.D., Er, M.H., Venkateswarlu, R., Chan, P.: Max-mean and max-median filters for detection of small targets. Proc. SPIE. **3809**, 74–83 (1999)

Huang, Z., Chen, L., Zhang, Y., Yu, Z., Fang, H., Zhang, T.: Robust contact-point detection from pantograph-catenary infrared images by employing horizontal-vertical enhancement operator. Infrared Phys. Technol. **101**, 146–155 (2019a)

Huang, S., Peng, Z., Wang, Z., Wang, X., Li, M.: Infrared small target detection by density peaks searching and maximum-gray region growing. IEEE Geosci. Remote Sens. Lett. **16**, 1919–1923 (2019b)

Lv, P.-Y., Sun, S.-L., Lin, C.-Q., Liu, G.-R.: Space moving target detection and tracking method in complex background. Infrared Phys. Technol. **91**, 107–118 (2018)

Nasiri, M., Chehresa, S.: Infrared small target enhancement based on variance difference. Infrared Phys. Technol. **82**, 107–119 (2017)
Nie, J., Qu, S., Wei, Y., Zhang, L., Deng, L.: An infrared small target detection method based on multiscale local homogeneity measure. Infrared Phys. Technol. 90, 186–194 (2018)
Qi, S., Xu, G., Mou, Z., Huang, D., Zheng, X.: A fast-saliency method for real-time infrared small target detection. Infrared Phys. Technol. 77, 440–450 (2016)
Qin, Y., Bruzzone, L., Gao, C., Li, B.: Infrared small target detection based on facet kernel and random walker. IEEE Trans. Geosci. Remote Sensing. 57, 7104–7118 (2019)
Soni, T., Zeidler, J.R., Ku, W.H.: Performance evaluation of 2-D adaptive prediction filters for detection of small objects in image data. IEEE Trans. Image Process. 2, 327–340 (1993)
Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: Sixth international conference on computer vision (ICCV '98), pp. 839–846. IEEE Computer Society (1998)
Wan, M., Gu, G., Cao, E., Hu, X., Qian, W., Ren, K.: In-frame and inter-frame information based infrared moving small target detection under complex cloud backgrounds. Infrared Phys. Technol. 76, 455–467 (2016)
Wang, X., Yang, L.T., Li, H., Lin, M., Han, J., Apduhan, B.O.: NQA: a nested anti-collision algorithm for RFID systems. ACM Trans. Embed. Comput. Syst. 18, 1–21 (2019)
Wang, G., Yang, J., Xu, J.: Granular computing: from granularity optimization to multi-granularity joint problem solving. Granul. Comput. 2, 105–120 (2017)
Wei, Y., You, X., Li, H.: Multiscale patch-based contrast measure for small infrared target detection. Pattern Recognit. 58, 216–226 (2016)
Xia, C., Li, X., Zhao, L.: Infrared small target detection via modified random walks. Remote Sens. 10, 2004 (2018)
Ye, Y., Shan, J., Bruzzone, L., Shen, L.: Robust registration of multimodal remote sensing images based on structural similarity. IEEE Trans. Geosci. Remote Sens. 55, 2941–2958 (2017)
Zeng, M., Li, J., Peng, Z.: The design of top-hat morphological filter and application to infrared target detection. Infrared Phys. Technol. 48, 67–76 (2006)
Zhang, L., Peng, L., Zhang, T., Cao, S., Peng, Z.: Infrared small target detection via non-convex rank approximation minimization joint l2, 1 norm. Remote Sens. 10, 1821 (2018a)
Zhang, P., Wang, X., Wang, X., Fei, C., Guo, Z.: Infrared small target detection based on spatial-temporal enhancement using quaternion discrete cosine transform. IEEE Access 7, 54712–54723 (2019)
Zhang, H., Zhang, L., Yuan, D., Chen, H.: Infrared small target detection based on local intensity and gradient properties. Infrared Phys. Technol. 89, 88–96 (2018b)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.