


IISTD: Image Inpainting-Based Small Target Detection in a Single Infrared Image

Deyong Lu , Qiang Ling , Yuanyuan Zhang, Zaiping Lin, and Wei An

Abstract—Small target detection is a crucial and challenging task in infrared search and track system. Background estimation-based methods is an effective and important approach for infrared small target detection. Affected by the target pixels, existing background estimation methods may reconstruct an inaccurate background. Based on image inpainting technique, we propose a novel two-stage approach to predict more accurate backgrounds. At the first stage, the inner and outer window-based image inpainting is used to obtain a rough background estimation. Then, a mask of candidate target region is automatically obtained by calculating and evaluating the difference between raw image and rough background. In the second stage, the final accurate background is predicted by mask-based image inpainting. It recovers the removed candidate target area using the information of surrounding background pixels, avoiding target pixels to participate in the calculation of background reconstruction. Finally, the target saliency map is obtained by subtracting the final estimated background from the original image, and a simple adaptive threshold is used to segment the target. Experimental results on real infrared images and sequences demonstrate that the proposed method outperforms other state-of-the-arts. It is simple and effective, with strong robustness and good real-time performance.

Index Terms—Background estimation, image inpainting, infrared image, small target detection.

I. INTRODUCTION

INFRARED search and track (IRST) system has been widely used in military and civilian fields, such as remote sensing, early warning system, precision guidance system, surveillance and reconnaissance. Particularly, with the increasing popularity of unmanned aerial vehicles (UAVs), unauthorized use of UAVs is becoming more frequent. Therefore, the detection of UAVs is a pressing problem today, especially for noncooperative UAVs. Infrared imaging technology provides an important and effective

way for remote detection of small moving targets [1], [2]. Small target detection in IRST system has always been a challenging problem for many reasons. They are as follows.

- 1) Due to the long imaging distance, the target occupies only a few pixels in the infrared image, and usually lacks the features of shape and texture.
- 2) The target radiation is weak and the signal-to-clutter ratio (SCR) is very low.
- 3) In a complex background, the target is corrupted or destroyed by severe clutter and noise.
- 4) In some scenes, the size of the target changes to some extent, which makes it difficult to continuously detect and track the target.

Generally speaking, the mathematical model of an infrared image on the detector array can be expressed as follows:

$$I(x, y) = T(x, y) + B(x, y) + N(x, y) \quad (1)$$

where (x, y) is the pixel location, I , T , B , and N are the original infrared image, the target component, the background component, and the noise component, respectively. The noise component N in the model can be removing, and the model becomes $I = T + B$. The basic idea of infrared small target detection is to use the inherent characteristics of the target and the background to enhance the target and suppress the background, so as to separate the small target and the background from the original infrared image [3]. Most of the small target detection methods are based on some assumptions about target, background, and noise [4], [5], and the applicability of the assumptions will determine its effectiveness and robustness in practical applications [6]. We assume that in most cases, small target is discontinuous with its adjacent region, and its intensity is higher than that of the local neighborhood when dealing with bright targets [7]. This assumption is also based on the characteristics of the infrared sensor system. The radiation intensity of the target is independent of the surrounding background, and generally higher than that of the local background. The infrared radiation intensity of the local background area has a strong spatial correlation, so that the neighborhood information can be used to reconstruct the background covered by the target [8].

In the past few decades, a large number of approaches have been proposed to improve the performance of infrared small target detection methods [4], [5], [9], [10], [11], [12]. Generally speaking, they can be simply categorized into two groups: 1) single frame-based methods; and 2) multiframe-based methods. Many factors lead to the performance degradation of

Manuscript received 27 April 2022; revised 15 June 2022, 6 July 2022, and 1 August 2022; accepted 16 August 2022. Date of publication 22 August 2022; date of current version 2 September 2022. This work was supported in part by the National Key R&D Program of China under Grant 2021YFB3100800, in part by National Natural Science Foundation of China under Grant 62002372 and Grant 62001478, and in part by the Shanghai Aerospace Science and Technology Innovation Fund, and CARD Fundamentals and Frontier Technology Research Fund. (Corresponding author: Qiang Ling.)

Deyong Lu is with the College of Electronic Science, National University of Defense Technology, Changsha 410073, China, and also with the Computational Aerodynamics Institute, China Aerodynamics Research and Development Center, Mianyang 621000, China (e-mail: iamludeyong@qq.com).

Qiang Ling, Zaiping Lin, and Wei An are with the College of Electronic Science, National University of Defense Technology, Changsha 410073, China (e-mail: lq910131@163.com; linzaiping@nudt.edu.cn; anwei@nudt.edu.cn).

Yuanyuan Zhang is with the Tianfu College, Southwestern University of Finance and Economics, Chengdu 610000, China (e-mail: zhangyuanyuan1@tfswufe.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2022.3200380

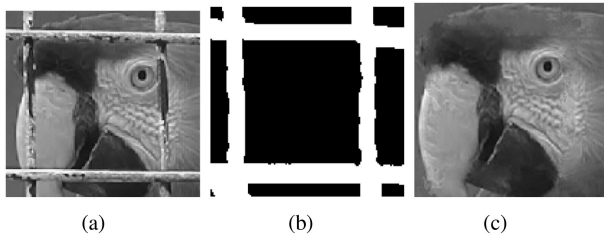


Fig. 1. Example of image inpainting. (a) Original image. (b) Mask to be inpainted. (c) Image inpainting result of [19]. (a) Original image. (b) Mask. (c) Inpainting result.

multiframe-based methods, making single frame-based methods more essential and important [13]. Therefore, this article mainly focuses on single frame-based methods. Zhao et al. [13] comprehensively reviewed existing typical single frame-based methods. Recently, many new methods have been proposed to solve the problem of infrared small target detection [11], [12], [14], [15], [16]. However, there are some disadvantages in the existing methods. First, the robustness and real-time performance of the detection algorithm are not enough. Second, the accuracy of the predicted background is insufficient, and target residues may be present in the reconstructed background. Existing background estimation methods cannot accurately reconstruct the background image. Whether it is a multiwindow or a protection window mechanism, it is affected by the target pixels on background estimation [8], [17], [18]. Therefore, the aim of this article is to use image inpainting technology to accurately reconstruct the background, and propose a new infrared small target detection algorithm to solve these shortcomings and weaknesses.

Image inpainting has been used for various applications, such as image restoration, object removal, disocclusion, loss concealment, and texture synthesis [19], [20]. It involves filling in missing parts or removing undesired objects from an image [21], [22]. Specifically, image inpainting can remove selected objects in an image, such as text, dates, or other selected targets, and then, recover the removed area using the information of the surrounding area [20], [23]. Most image inpainting algorithms are based on the following techniques: Copy-paste texture synthesis, coherence between neighboring pixels, and geometric partial differential equation (PDE). Fig. 1 shows an example of image inpainting [19].

Inspired by the idea of image inpainting, we propose a novel infrared small target detection method, in which a more accurate predicted background is obtained through a two-stage approach. In the first stage, the inner and outer window-based image inpainting (IOWII) is used to predict a rough background estimation, which aims to automatically calculate a mask of candidate small target region. In the second stage, the final background is reconstructed by mask-based image inpainting (MII), in which only pixels of candidate target regions are replaced with the estimated background pixels. It only uses the background part pixels to reconstruct the final background, avoiding the target pixels to participate in the calculation of background reconstruction. Finally, a target saliency map is calculated by subtracting the final background from the original

infrared image, and a simple adaptive threshold is adopted to segment the targets.

The main contributions of this work can be summarized as follows.

- 1) To the best of our knowledge, it is the first time using image inpainting technology for infrared small target detection.
- 2) A novel image inpainting-based small target detection (IISTD) method is proposed, where IOWII is used to obtain a mask matrix indicating the target pixels and the background pixels, and MII reconstructs the final accurate background image from the background pixels only.
- 3) Experiments show that the proposed method is simple and effective, has strong robustness and good real-time performance, and outperforms other state-of-the-art methods.

The rest of this article is organized as follows. Section II briefly reviews some related work. The proposed IISTD method is described in Section III. Experiments on real infrared images and sequences are presented in Section IV. Finally, Section V concludes the article.

II. RELATED WORK

In recent decades, researchers have proposed a large number of small target detection methods for infrared search and track systems. Roughly speaking, they can be simply categorized into two groups: 1) single frame-based methods [5], [7], [10], [13]; and 2) multiframe-based methods [4], [9], [11], [24], [25]. Given prior knowledge of target velocity and trajectory, multiframe-based methods process a sequence of frames to detect targets. The so-called tracking-before-detection method accumulates the target energy along its trajectory to enhance the target for reliable detection results [3]. 3-D matched filtering is a classical multiframe-based method that must be properly matched to the target feature and its velocity [4]. Some multiframe-based methods first process a single frame to enhance the target and suppress the background, and then use the accumulation or correlation between frames to obtain the detection results [9]. For static background or consistent targets in adjacent frames, multiframe-based methods perform well. However, in some practical applications, the background changes quickly with high maneuvering targets, and the prior knowledge is unavailable. The aforementioned issues can may cause severe performance degradation of multiframe-based methods. In addition, the algorithm complexity and the long computation time of multiframe-based methods cannot meet the real-time requirement of some practical applications. Therefore, single frame methods are essential and appropriate in such applications, and have attracted a lot of attention in recent decades [5], [7], [8], [26], [27], [28], [29]. In this article, we mainly consider small target detection in a single infrared image. We simply divide single frame-based methods into background estimation-based methods [8], [9], [17], [30], local information-based methods [7], [16], [18], [28], [31], decomposition-based methods [5], [12], [32], [33], [34], learning-based methods [11], [14], [15], [35], and other methods [36], [37], [38].

Background estimation-based methods include TopHat morphological filter [26], [30], max-mean and max-median filter [9], and 2-D least mean square filter [39], [40]. The residual image

is obtained by subtracting the predicted background from the original image, and then small targets are detected by using some postprocessing methods, such as segmentation or accumulation of residual images [8], [9]. Recently, Hu et al. [41] applied the nonlocal means filter to predict the background. Han et al. [17] proposed an adaptive background estimation method using eight multidirectional templates. After calculating all eight estimates, they chose the one closest to the original grayscale as the predicted background for that pixel. This will cause a problem that when estimating the background gray value of the bright target position, the final background estimation value will be brighter and, thus, inaccurate. In [8], a three-layer estimation window was proposed, which calculated the average of the active pixels in the eight directions of the surrounding layer, and then, selected the one closest to the central value as the final background estimation. It also leads to an inaccurate background estimation, which is close to the center value instead of the true background gray value.

Local information-based methods are the most effective approaches for target enhancement, which detect small targets by evaluating the discontinuity and contrast between the target area and its neighbors. The evaluation measures are the local contrast measure (LCM) [7], the improved LCM [42], the local difference measure (LDM) [43], the weighted LDM [28], the multiscale patch-based contrast measure (MPCM) [44], the ratio-difference joint LCM [17], the multiscale trilinear LCM [18], the halo structure prior-based LCM [16]. All these methods can effectively enhance the small targets in infrared images and obtain excellent detection results.

Decomposition-based methods distinguish the target from the background according to their different structural characteristics, which are also known as data structure-based methods [13]. Most of them are based on low-rank and sparse decomposition, transforming the target detection problem into a robust principle component analysis (RPCA) optimization problem [5], [32]. Based on the infrared patch-image (IPI) model [5] or infrared patch-tensor model [10], [32], the new reconstructed image or tensor is decomposed into low-rank background component and sparse target component using various RPCA algorithms, such as the alternating direction method of multipliers [45], [46]. Zhao et al. [34] utilized Tucker decomposition to estimate the major principal components that are mainly background, and then eliminated them to obtain the remaining minor components containing target information.

Recently, learning-based methods have received more attention from researchers, and some results have been obtained for infrared small target detection [11], [14], [15], [35]. Du et al. [11] used a neural network to extract spatiotemporal features in infrared sequences for small target detection, where interframe energy accumulation was used to suppress background and enhance the target. Dai et al. [14] attempted to combine the feature learning capabilities of deep networks and the physical mechanisms of model-driven methods into an end-to-end network to handle the detection of infrared small targets, in which a traditional patch-based local contrast measure is modified and inserted into the attentional local contrast network. In [35], they decomposed the small target detection problem

into two subtasks of suppressing miss detection and false alarm, and then proposed a conditional generative adversarial network consisting of two generators and a discriminator. Chen et al. [15] proposed a local patch network with global attention capability, which can focus on the local saliency of small target features.

III. IMAGE INPAINTING-BASED INFRARED SMALL TARGET DETECTION

In this section, we propose a novel small target detection method based on image inpainting technique. The key idea of the proposed method is how to accurately reconstruct the background image from the original infrared image, and then achieve the purpose of separating the target from the background. The new method obtains the background image by reconstructing the background twice, i.e., rough estimation and fine estimation. The rough background estimation adopts the IOWII method, and the fine background reconstruction is obtained by the MII method, in which the candidate target region needs to be provided. The proposed IISTD method is presented in Fig. 2.

A. Inner and Outer Windows-Based Image Inpainting

The key issue that needs to be considered is how to accurately estimate the background of the infrared image, especially the region of small target and its neighborhood. Therefore, we designed an inner and outer windows-based background estimation method, which is suitable for both the background area and the target area. Our method first digs out the pixel, and then, uses its neighborhood to recover it through image inpainting approaches. Fig. 3 shows the inner and outer windows used to calculate the gray value of a pixel in the background image. The background estimation is calculated from top to bottom and from left to right. In the ideal case, the size of the inner window is the target size, which plays a role in preventing the bright target pixels from participating in the estimation of the background gray value. We hope that the pixels in the outer window (excluding the inner window) always belong to the background part, so we can only use these pixels for background reconstruction. In order to reduce the complexity and computation, we use the average value of these pixels as the estimated gray value of the pixel. Assuming that the pixel is located at (x, y) , the corresponding estimated value in the first rough background B_1 is

$$B_1(x, y) = \frac{1}{c_1} \sum_{l=-r_1}^{r_1} \sum_{k=-r_1}^{r_1} W(l, k) I(x+l, y+k), \quad (2)$$

$$W(l, k) = \begin{cases} 0, & \text{if } (l, k) \in \Omega_2, \\ 1, & \text{if } (l, k) \in \Omega_1 \text{ and } (l, k) \notin \Omega_2, \end{cases}$$

$$c_1 = (2r_1 + 1)^2 - (2r_2 + 1)^2 \quad (3)$$

where the sizes of the outer window Ω_1 and inner window Ω_2 are $(2r_1 + 1) \times (2r_1 + 1)$ and $(2r_2 + 1) \times (2r_2 + 1)$, respectively. r_1 and r_2 are the radius of windows Ω_1 and Ω_2 , respectively. W is a matrix that labels the active pixels involved in the background estimation. For example, when $r_1 = 2$, $r_2 = 1$, the

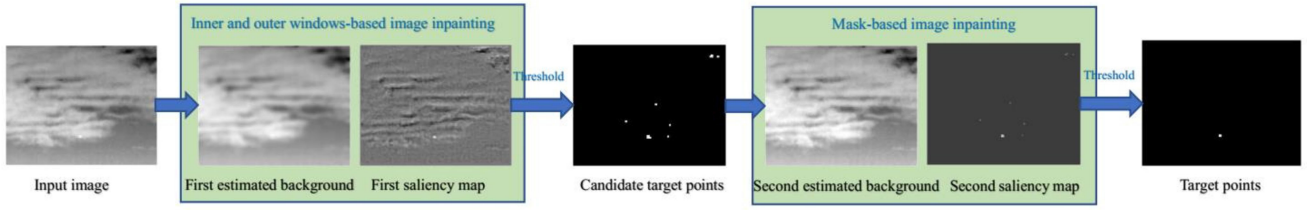


Fig. 2. Overview of the proposed image inpainting-based infrared small target detection method.

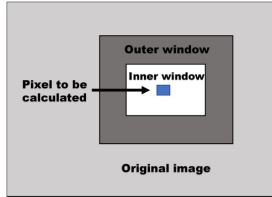


Fig. 3. Schematic diagram of the inner and outer windows used to calculate a pixel gray value in rough background estimation.

constant $c_1 = 16$ and matrix W is

$$W = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \quad (4)$$

Next, we consider the following three cases to illustrate the effectiveness of the IOWII method.

- 1) When the calculated pixel belongs to the background area of the original infrared image, the predicted background value can be accurately computed by (2), where we assume that the background is locally consistent.
- 2) When the calculated pixel belongs to the target area, (2) can also correctly obtain the pixel gray value of the estimated background, because most of the target pixels have been excluded from the background estimation process.
- 3) When the calculated pixel belongs to the edge of the target area, the estimated value obtained by (2) is affected by the gray value of the target pixels, which has little effect on the whole target detection algorithm and can be eliminated or mitigated by other methods described later.

In addition, we consider the effect of the window size. If the inner window is too small, the protection mechanism of the inner window does not work, causing bright target pixels to participate in the background reconstruction. On the other hand, if the inner window is too large, the estimated background is too smooth. In our experiments, the inner window size can be 3×3 , or 5×5 , and the outer window size is 5×5 , or 7×7 .

After the first background reconstruction is completed, the first target saliency map S_1 is gained by subtracting the estimated background image from the original image

$$S_1(x, y) = I(x, y) - B_1(x, y). \quad (5)$$

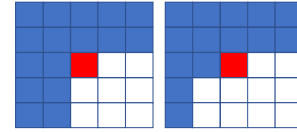


Fig. 4. Example of fine background estimation using the background pixels of the 5×5 neighborhood. The center pixel is estimated as the average value of the blue background pixels (left), and then be updated to the background pixels. As a new background pixel, it participates in the reconstruction of the next pixel to its right (right).

B. Candidate Target Region

In this section, we introduce the selection of candidate target regions, which provide a mask for the fine background estimation in the next step. It should be noted that our method can handle both bright targets and dark targets. Without loss of generality, only the bright target situation will be discussed here. In the first target saliency map, the bright point is most likely to be a target. Therefore, we use a simple threshold segmentation method to segment the candidate target region. The threshold T_1 is given as follows:

$$T_1 = m_1 + k_1 \sigma_1 \quad (6)$$

where k_1 is a constant parameter, m_1 and σ_1 are the mean and the standard deviation of the first target saliency map S_1 . Generally speaking, the parameter k_1 can be selected manually and experimentally, and the range is from 3 to 10 in our experiments. A pixel at (x, y) belongs to candidate target pixels if $S_1(x, y) > T_1$, and a pixel at (x, y) belongs to background pixels if $S_1(x, y) \leq T_1$. The mask matrix is defined as follows:

$$M(x, y) = \begin{cases} 0, & \text{if } S_1(x, y) > T_1, \\ 1, & \text{if } S_1(x, y) \leq T_1. \end{cases} \quad (7)$$

In order to select all the small target pixels, the candidate target region in the mask matrix needs to be expanded by several pixels in the case of a prior with a larger target, which can further avoid the influence of the target pixels on the background reconstruction process. For example, when dealing with the large target in image VII of Fig. 5, it is sufficient to expand the candidate target region by 2 pixels. In other experiments, mask expansion was not required.

C. Mask-Based Image Inpainting

After the mask of the candidate target area is automatically provided in the first stage, the original infrared image is divided into target pixels and background pixels, and then, MII method

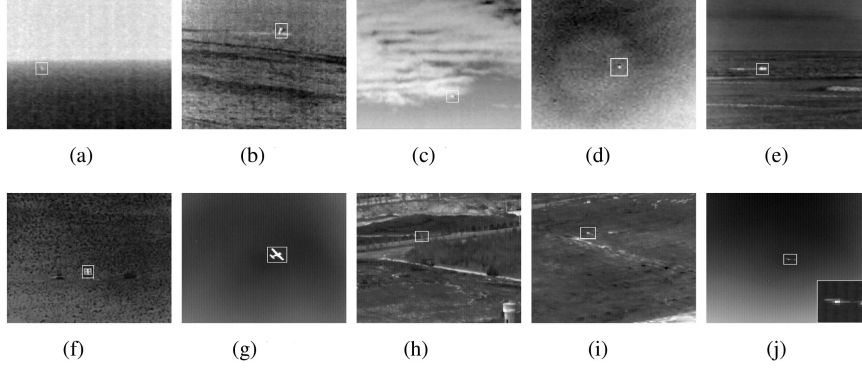


Fig. 5. Ten typical real infrared images. Small targets are marked by white boxes. (a) Image I. (b) Image II. (c) Image III. (d) Image IX. (e) Image V. (f) Image VI. (g) Image VII. (h) Image VIII. (i) Image IX. (j) Image X.

is proposed to reconstruct the accurate background using the background pixels in the fine estimation stage.

First, the fine background image B_2 is initialized as

$$B_2(x, y) = M(x, y)I(x, y). \quad (9)$$

Then, from the top left to the bottom right, we estimate the pixel gray values at the candidate target locations only using the background pixels of their neighborhoods. When using a local window of size $(2r_3 + 1) \times (2r_3 + 1)$, the background prediction for pixel (x, y) is as follows:

$$B_2(x, y) = \frac{1}{c_2} \sum_{l=-r_3}^{r_3} \sum_{k=-r_3}^{r_3} M(x+l, y+k) B_2(x+l, y+k), \quad (10)$$

$$c_2 = \sum_{l=-r_3}^{r_3} \sum_{k=-r_3}^{r_3} M(x+l, y+k). \quad (11)$$

Fig. 4 shows an example of fine background estimation using the background pixels of the 5×5 neighborhood. In the figure, the white pixels belong to the candidate target area, and the blue ones are the background pixels, and the red pixels in the center are the pixels that need to be calculated. For the sake of low complexity and high real-time performance, we use the average value of the background pixels (blue ones) in the local neighborhood as the gray value of the pixel (red ones). In addition, in order to ensure that there are background pixels in the local neighborhood every time, the background pixels are updated immediately after completing the estimation of a pixel. That is to say, after completing the background estimation of $B_2(x, y)$, $M(x, y)$ is updated to 1. As shown in Fig. 4, when the center pixel of the left figure is estimated, it will be updated to the background pixels and then participates in the calculation during the reconstruction of the next pixel to its right.

Moreover, it is also important to note the following.

- 1) The local neighborhood used for fine background estimation can be 7×7 , or other sizes.
- 2) Instead of the mean value, the background pixel value can also be estimated by weighted average, which may increase the amount of calculation.

- 3) The estimation order of candidate target pixel values, i.e., the order of image inpainting, can be further studied, such as from the edge of the region to the center.
- 4) It may be worth trying other image inpainting techniques to get a better background in the future, such as diffusion-based or patch-based algorithms, PDE-based algorithms, and learning-based methods [19], [20].

After using the MII method to complete the fine background reconstruction, the second estimated background is subtracted from the original image to obtain the second target saliency map S_2

$$S_2(x, y) = I(x, y) - B_2(x, y). \quad (12)$$

D. Whole Method of IISTD

The whole algorithm of image inpainting-based infrared small target detection is shown in Algorithm 1, which can be described as follows.

- 1) The IOWII method is used to calculate the first rough background estimation. Then, the first target saliency map S_1 is obtained by subtracting the first estimated background from the original image.
- 2) The threshold T_1 in the (6) is used to segment the first target saliency map to obtain the candidate target points.
- 3) The fine background estimation is reconstructed by the MII method, and then, subtracted from the original image to get the second target saliency map S_2 .
- 4) Small targets are detected via a simple threshold segmentation with

$$T_2 = k_2 \max(S_2) \quad (13)$$

where k_2 is a constant in the range of $(0.3, 0.9)$ and $\max(S_2)$ is the brightest pixel value of S_2 .

IV. EXPERIMENTAL RESULTS

In this section, we report the experimental results to demonstrate the detection performance of the proposed method. The experiments are conducted on a computer with 12-GB random access memory and 2.40-GHz Intel i5 processor, and all the code is implemented in MATLAB.

Algorithm 1: IISTD.

```

1: Input: Infrared image  $I$  with resolution  $M \times N$ ,
   parameter  $r_1, r_2, r_3, k_1, k_2$ .
2: for  $x = 1 : M$ 
3:   for  $y = 1 : N$ 
4:     Calculate the rough background  $B_1(x, y)$ 
       according to (2).
5:   end for
6: end for
7:  $S_1 = I - B_1$ 
8:  $T_1 = m_1 + k_1\sigma_1$ 
9: Calculate the mask  $M$  according to (7) and (8).
10: Initialize the fine background  $B_2 = M \cdot I$ 
11: while  $M(x, y) = 0$  do
12:   Calculate the fine background  $B_2(x, y)$  according to
       (10).
13:    $M(x, y) = 1$ 
14: end while
15:  $S_2 = I - B_2$ 
16:  $T_2 = k_2 \max(S_2)$ 
17: Output: Target location and other information.

```

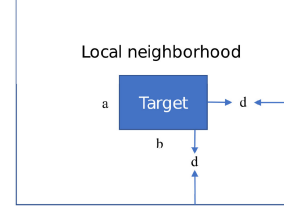


Fig. 6. Schematic diagram of target region and its local neighborhood for calculating the signal-to-clutter ratio (SCR).

signal-to-clutter ratio gain (SCRG), which is defined as follows:

$$\text{SCRG} = \frac{\text{SCR}_{out}}{\text{SCR}_{in}} \quad (15)$$

where the subscripts *out* and *in* represent the obtained target saliency maps before target segmentation and the original infrared images, respectively.

On the other hand, background suppression ability is also an important consideration to evaluate the performance of the algorithm. Thus, we define the background suppression factor (BSF) as follows [46]:

$$\text{BSF} = \frac{\sigma_{in}}{\sigma_{out}}. \quad (16)$$

Finally, the ROC curve can also be used to evaluate the performance of small target detection methods. The ROC curve reflects the relationship between the probability of detection (Pd) and false alarm rate (Fa), which are defined by

$$\text{Pd} = \frac{\text{NCT}}{\text{NT}}, \text{Fa} = \frac{\text{NFP}}{\text{NP}} \quad (17)$$

where NCT is the number of correctly detected targets, NT is the number of targets, NFP is the number of false pixels that are determined as targets, and NP is the total number of pixels. In our experiments, we change the threshold during the target segmentation process to calculate a series of Pd and Fa, and then, plot them to obtain the ROC curve. The closer the curve is to the upper left corner, the better the performance.

A. Datasets

First, ten real infrared images are used to test and verify the target enhancement ability and background suppression effect of the proposed algorithm. As shown in Fig. 5, the background consists of sea scene, sky scene, and ground scene. Each of the first nine images contains only one target, and image X contains two UAV targets. For better visualization, the small target area is enlarged several times to be displayed in the lower-left or lower-right corner of the image if necessary.

Then, we consider three infrared image sequences to compare our method with other baseline methods. The detailed descriptions are shown in Table I. One frame of Sequence 1 is image III in Fig. 5(c). Images VIII and IX in Fig. 5 are the first frame of Sequences 2 and 3, respectively.

B. Evaluation Metrics

To evaluate the detection performance of different methods, we introduce several metrics in this section. First, the SCR is adopted to evaluate the relative contrast between the target signal and clutter, which also reflects the difficulty of target detection. The definition of SCR in this article is as follows:

$$\text{SCR} = \frac{|m_t - m_b|}{\sigma_b} \quad (14)$$

where m_t denotes the maximum gray value of the target, m_b and σ_b denote the mean value and the standard deviation of the local neighborhood of the target, respectively. Some researchers exploited the average value of the target to calculate SCR, where the definition of target pixels has a great impact on it. Therefore, we use the target maximum value to calculate the SCR. Fig. 6 shows the schematic diagram of the target area and its local neighborhood, whose size is $a \times b$ and $(a + 2d) \times (b + 2d)$, respectively. In the following experiments, we set d as 20 pixels.

In order to further compare the target enhancement capabilities of different algorithms, we also introduce the concept of

C. Baseline Methods and Parameter Setting

In this article, we compare the proposed method with six baseline methods. They are as follows:

- 1) the TopHat filter [30];
- 2) the local contrast measure (LCM) [7];
- 3) the multiscale patch-based contrast measure (MPCM) [44];
- 4) the facet kernel and random walker model (FKRW) [36];
- 5) the infrared patch-image model (IPI) [5];
- 6) the partial sum of tensor nuclear norm model (PSTNN) [10].

The parameter settings are summarized in Table II. In the experiments, we choose the inner window to be 3×3 (i.e., $r_2 = 1$) and the outer window to be 5×5 (i.e., $r_1 = 2$) in the first rough background estimation stage. The local window size is 5×5 (i.e., $r_3 = 2$) in the second fine background estimation stage. Specifically, in the experiments of Sequences 1, 2, and 3, k_1 are set to 5, 8, and 9, respectively.

TABLE I
DETAILED DESCRIPTIONS OF THREE INFRARED IMAGE SEQUENCES

	Frame number	Size	Target type	Background description	Target description
Sequence 1	30	256×200	Airplane	Sky scene with banded cloud, slowly changing	Small, size variation
Sequence 2	120	256×256	UAV	Ground scene with trees, roads and towers, slowly changing	Very small, $1 \sim 3$ pixels
Sequence 3	120	256×256	UAV	Ground scene with trees and roads, rapidly changing	Small, fast moving

TABLE II
DETAILED PARAMETER SETTINGS OF THE BASELINE METHODS.

Methods	Acronyms	Parameter
TopHat method [30]	TopHat	Structure shape: Square, structure size: 4×4
Local contrast measure [7]	LCM	$L=4$, local window size: 3,5,7,9
Multiscale patch-based contrast measure [44]	MPCM	$L=4$, local window size: 3,5,7,9
Facet kernel and random walker model [36]	FKRW	$K=4$, $p=6$, $\beta = 200$, window size: 11
Infrared patch-image model [5]	IPI	Patch size: 50×50 , sliding step: 10
Partial sum of tensor nuclear norm model [10]	PSTNN	Patch size: 40×40 , sliding step: 40
The proposed method	IISTD	Inner window size: 3×3 , outer window size: 5×5 , local window size: 5×5

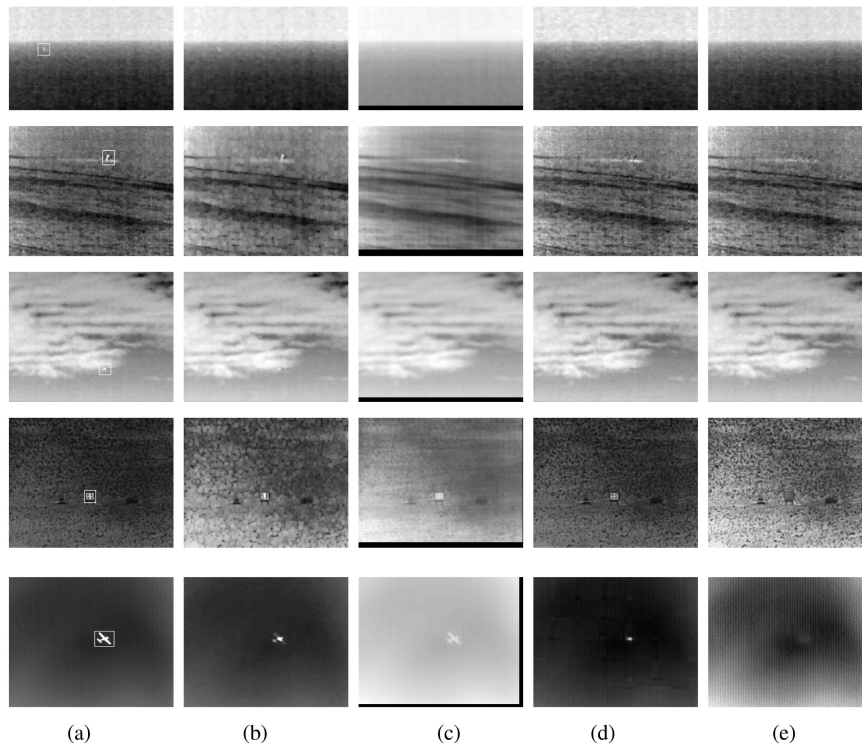


Fig. 7. Original infrared images and the estimated background images by different methods. (a) Original. (b) TopHat. (c) IPI. (d) PSTNN. (e) IISTD.

D. Comparison With Baseline Methods

In this section, we first compare the results of background reconstruction, then show the visual and numerical results of different images using different methods, then compare the detection rate and Fa through ROC curves, and finally compare their computational time.

1) *Background Reconstruction:* As described in the previous sections, our method is based on background reconstruction, and it can estimate the background image more accurately. We show this through Fig. 7, in which we compare the estimated background images obtained by several methods including TopHat [30], IPI [5], PSTNN [10], and our IISTD method. From the figure, it is easy to find that the background images obtained

by IISTD are the most accurate one, and there is almost no target residue. In contrast, the background images estimated by other methods contain target residues, especially for the last two images VI and VII.

2) *Visual Results:* The first nine infrared images and the corresponding processing results of different methods before target segmentation are presented in Fig. 8. Each image has only one target in Fig. 8. Fig. 9 shows the processing results of different methods when there are two targets in image X. From these figures, we can find the following results.

1) Compared with TopHat, LCM, and MPCM, our IISTD method, as well as FKRW, IPI, and PSTNN, can remove most clutter and noise for different backgrounds, which is the key to reduce the Fa.

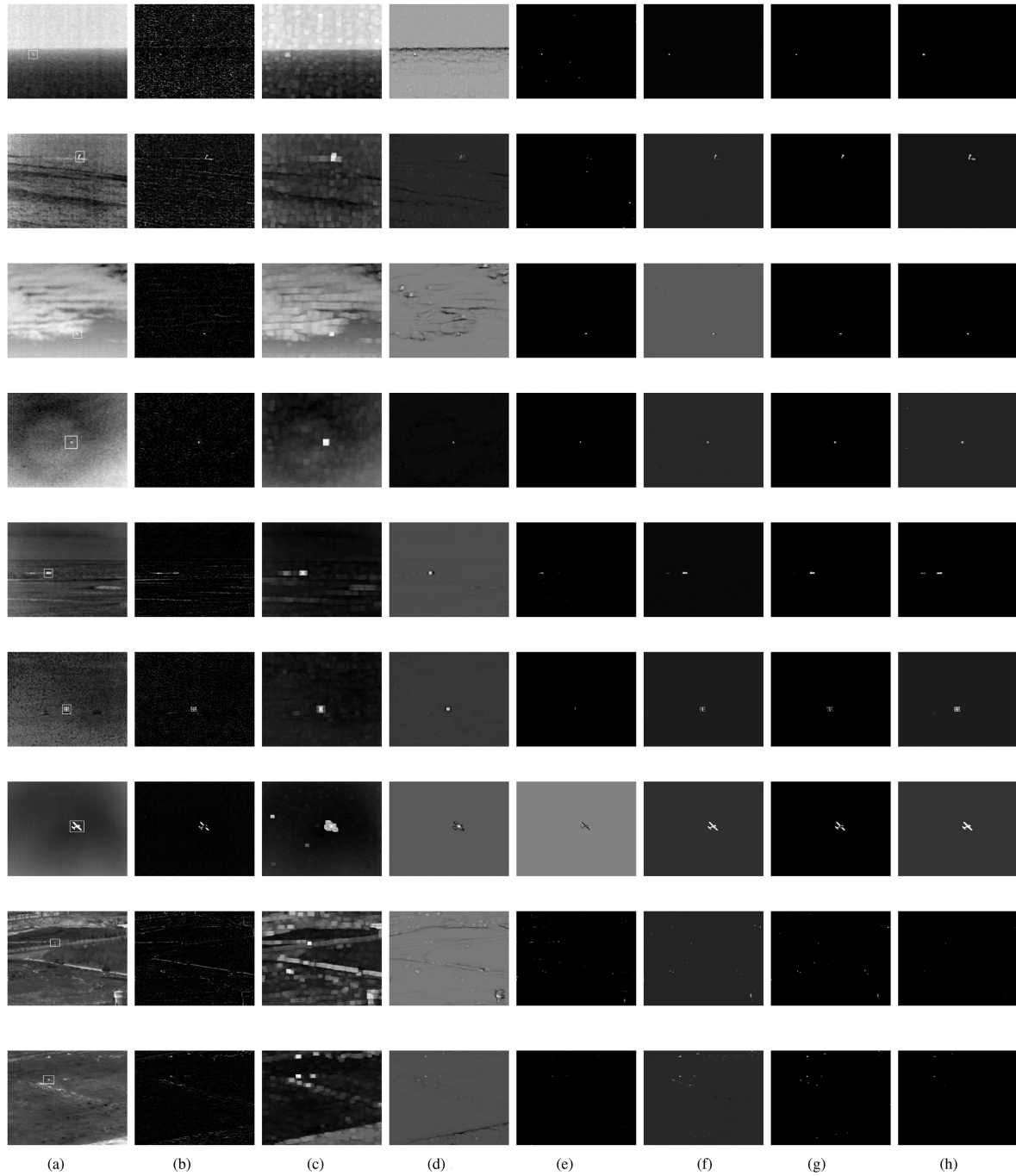


Fig. 8. Nine infrared images and the corresponding processing results of different methods before target segmentation. (a) The original infrared images. (b)–(g) The filtered results of the TopHat, LCM, MPCM, FKRW, IPI, and PSTNN methods, respectively. (h) The filtered results of the proposed IISTD method. (a) Raw image. (b) TopHat. (c) LCM. (d) MPCM. (e) FKRW. (f) IPI. (g) PSTNN. (h) IISTD.

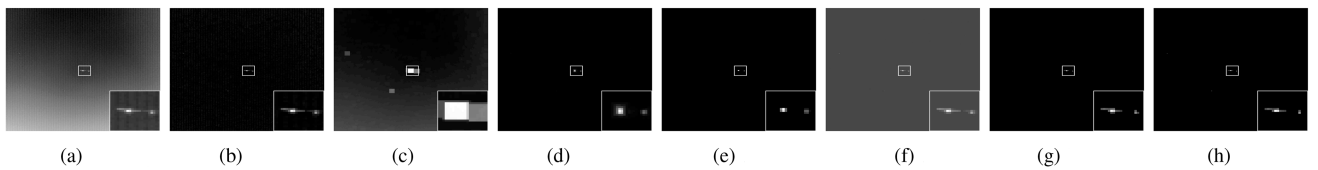


Fig. 9. Detecting results of different methods when there are two targets in image X. (a) Raw image. (b) TopHat. (c) LCM. (d) MPCM. (e) FKRW. (f) IPI. (g) PSTNN. (h) IISTD.

TABLE III
COMPARISON OF SCRG OBTAINED BY DIFFERENT METHODS FOR NINE TYPICAL INFRARED IMAGES

	Image I	Image II	Image III	Image IV	Image V	Image VI	Image VII	Image VIII	Image IX
TopHat [30]	7.8847	1.5346	10.4118	4.1515	0.9845	1.5260	1.2279	2.9979	2.3817
LCM [7]	1.3575	1.6594	1.2516	2.0866	1.3907	2.1482	0.4613	0.6962	0.5652
MPCM [44]	4.9209	10.4281	4.5820	19.4188	12.1474	9.7173	55.3531	0.7457	7.8015
FKRW [36]	Inf	4.3203	Inf	Inf	1.7173	Inf	-	1.2310	18.6490
IPI [5]	Inf	Inf	39.6940	194.9341	281.6875	Inf	18.4170	19.3566	5.1719
PSTNN [10]	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	4.0279
IISTD	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf

Notes: The bold values represent the best ones.

TABLE IV
COMPARISON OF BSF OBTAINED BY DIFFERENT METHODS FOR NINE TYPICAL INFRARED IMAGES

	Image I	Image II	Image III	Image IV	Image V	Image VI	Image VII	Image VIII	Image IX
TopHat [30]	14.9839	2.4848	16.5496	4.6798	2.9301	1.9759	1.3171	2.9915	2.5816
LCM [7]	0.8922	0.6119	0.7782	0.7978	0.4340	0.5711	0.2190	0.3195	0.2151
MPCM [44]	0.8307	0.2624	0.2996	0.6105	0.6415	0.2904	1.4159	0.2301	0.1393
FKRW [36]	Inf	725.0372	Inf	Inf	1567.3	Inf	Inf	434.1471	2503.7
IPI [5]	Inf	Inf	54.9083	256.3064	433.9241	Inf	20.3309	26.6310	6.3964
PSTNN [10]	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	4.1719
IISTD	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf	Inf

Notes: The bold values represent the best ones.

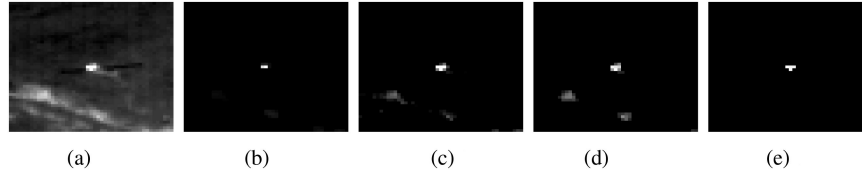


Fig. 10. Local neighborhoods of the target in image IX obtained by the FKRW, IPI, PSTNN, and IISTD methods, which are also used to calculate the local SCR and BSF. (a) original. (b) FKRW. (c) IPI. (d) PSTNN. (e) IISTD.

- 2) The target saliency images obtained by our IISTD method are the best, the most consistent with human vision, and can accurately maintain the original shapes of the targets, especially for images VI and VII.
- 3) LCM and MPCM cannot preserve the shapes of the targets due to their calculation formulas of the target saliency maps and the magnification effect. They are designed for the detection of point source targets, and are not sufficiently adaptable when dealing with other types of targets. In the result of LCM in Fig. 9, two UAV targets have been connected together, which is difficult to distinguish.
- 4) Most of the target saliency maps obtained by IPI and PSTNN methods are excellent, only in the cases of images VI and VII, the original shapes of the targets cannot be completely maintained. Note that you can zoom in on these figures to see more clearly through the electronic version of the article.

3) *Numerical Results:* Table III lists the SCRG obtained by different methods for the first nine typical infrared images. The BSF obtained by different methods for these images are listed in Table IV. From these tables, we can find some similar observations, which are closely related to the conclusions of the aforementioned visual results. Compared with the TopHat, LCM, and MPCM methods, our IISTD method and IPI, PSTNN, and FKRW have a stronger ability to enhance small targets and suppress background, resulting in higher SCRGs and BSFs. It can be seen from (14) that when the standard deviation of the local neighborhood of the target tends to zero, SCR tends to

infinity (since 0 is used as the denominator), and then, an infinite SCRG is obtained, denoted by **Inf**. When σ_{out} in (16) tends to zero, BSF also tends to infinity. Among these methods, the SCRGs and BSFs obtained by our IISTD method are all infinite, which proves that our method has the strongest ability of target enhancement and background suppression. Fig. 10 displays the local neighborhoods of the target in image IX obtained by the FKRW, IPI, PSTNN, and IISTD methods, which are also used to calculate the local SCR and BSF. There are some clutter and noise in the neighborhoods of FKRW, IPI, and PSTNN, so their SCRs, SCRGs, and BSFs will not reach infinity.

In addition, we also show the SCRG and BSF values obtained by different methods in three real image sequences in Figs. 11 and 12. For the convenience of display, the maximum value is set to 10^4 , i.e., **Inf** is displayed as 10^4 . It can be found that the SCRG and BSF values of the IISTD method can reach the maximum value in most cases, except for a few frames in Sequence 2, and some frames in Sequence 3. The IISTD method has strong capabilities in background suppression and target enhancement.

4) *ROC Curves:* We show the ROC curves of the proposed IISTD method and other baseline methods for three real image sequences in Fig. 13. These real infrared image sequences are described in detail in Table I. In the ROC curves, the vertical axis is the Pd and the horizontal axis is the logarithm of the Fa. For infrared image Sequence 1, the background is not too complicated, and the SCRs is not very low, so the detection rate of these methods is always higher than 0.7 in Fig. 13(a). Our IISTD method, like IPI, PSTNN, and FKRW methods,

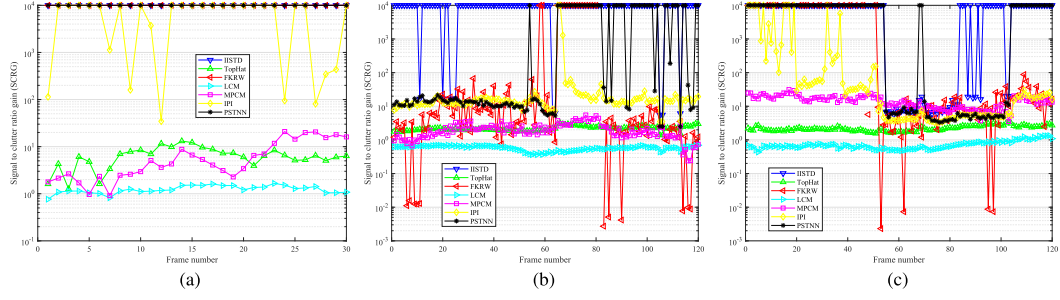


Fig. 11. Signal to clutter ratio gain (SCRG) values obtained by different methods in three real image sequences. For the convenience of display, the maximum value is set to 10^4 , i.e., **Inf** is displayed as 10^4 . (a) Sequence 1. (b) Sequence 2. (c) Sequence 3.

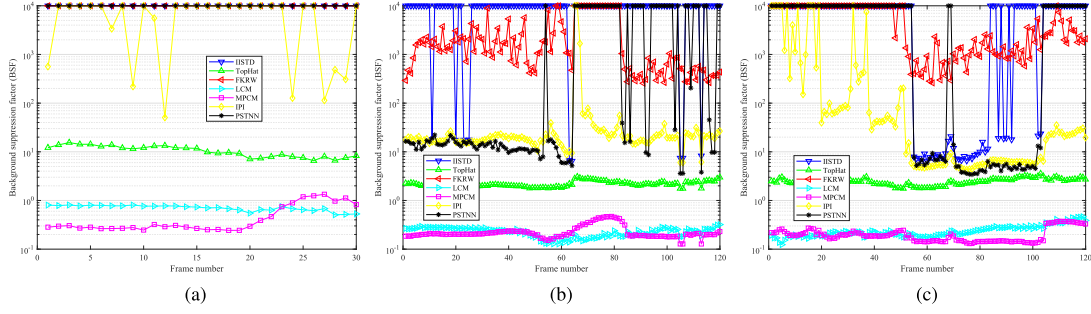


Fig. 12. Background suppression factor (BSF) values obtained by different methods in three real image sequences. For the convenience of display, the maximum value is set to 10^4 , i.e., **Inf** is displayed as 10^4 . (a) Sequence 1. (b) Sequence 2. (c) Sequence 3.

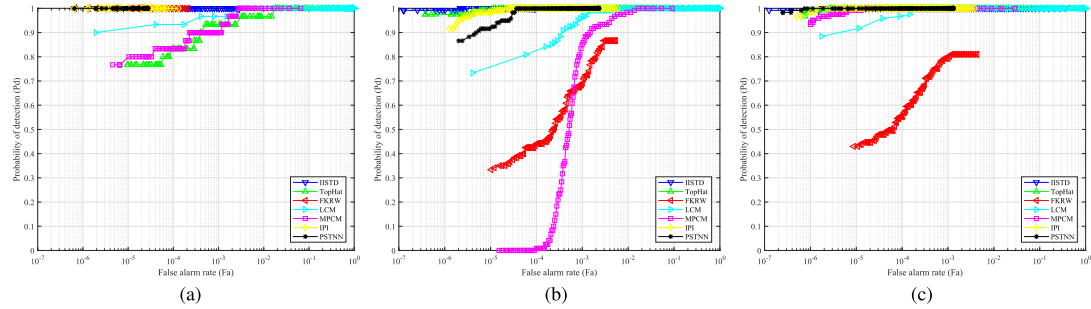


Fig. 13. Receiver operating characteristic (ROC) curves of seven infrared small target detection methods for three real image sequences. (a) Sequence 1. (b) Sequence 2. (c) Sequence 3.

can achieve a detection rate of 100% without false alarms. For infrared image Sequence 2, the background is a complex ground scene including trees, roads and towers, and the target is very small (less than 3 pixels). One frame of Sequence 2 (image VIII) and its processing results of seven methods can be seen in Fig. 8. Among the seven methods, the performance of the MPCM method is the worst. When the false alarm rate is about 10^{-4} , the detection rate is still zero. The performance of the FKRW method is also poor. When the false alarm rate has reached 0.005, the detection rate has not yet reached 0.9. However, under the condition of very low false alarm rate (less than 3×10^{-6}), the detection rate of our IISTD method has reached 1. The ground scene of infrared image Sequence 3 is also very complex. It contains some high brightness ground objects, which easily cause false detection. Its first frame is image IX in Fig. 5, and Fig. 8 also presents the filtered results of seven methods. It can be seen from Fig. 13(c) that our IISTD method is the best, followed

by the PSTNN method, and the worst is FKRW. To summarize, the proposed IISTD method has the best performance among these seven methods.

5) Running Time: In order to evaluate the performance of target detection algorithm, it is also necessary to consider its real-time performance, especially for some important practical applications. In most cases, the real-time performance of the detection algorithm and its accuracy and robustness are difficult to balance. Table V lists the average calculation time of these seven methods for different infrared images. We compare their computational efficiency under relatively fair conditions as much as possible. For each image, the simulation was repeated 100 times, and then the average was calculated. The bold values represent the best ones and the underlined values represent the second. It is easy to find that the TopHat filtering method is the fastest, followed by our IISTD method, and IPI method is the most time consuming. The average computing time of TopHat,

TABLE V
COMPARISON OF THE COMPUTATIONAL TIME (IN SECONDS) OBTAINED BY DIFFERENT METHODS FOR TEN TYPICAL INFRARED IMAGES

	Image I	Image II	Image III	Image IV	Image V	Image VI	Image VII	Image VIII	Image IX	Image X
TopHat [30]	0.0037	0.0036	0.0040	0.0026	0.0033	0.0036	0.0035	0.0036	0.0034	0.0035
LCM [7]	0.2636	0.2231	0.2548	0.1575	0.2361	0.2442	0.2566	0.2617	0.2584	0.2551
MPCM [44]	0.3034	0.2501	0.2973	0.1640	0.2870	0.2812	0.3147	0.3067	0.3031	0.3005
FKRW [36]	0.1378	0.1593	0.1648	0.1143	0.2298	0.1451	0.1726	0.3031	0.2761	0.1243
IPI [5]	20.1491	8.6750	20.1208	3.7282	15.1065	8.3735	19.0782	16.6014	19.0917	18.2594
PSTNN [10]	0.1837	0.2920	0.2317	0.0839	0.2007	0.3034	0.3147	0.5982	0.6981	0.1851
IISTD	<u>0.0087</u>	<u>0.0052</u>	<u>0.0076</u>	<u>0.0045</u>	<u>0.0076</u>	<u>0.0068</u>	<u>0.0077</u>	<u>0.0094</u>	<u>0.0098</u>	<u>0.0045</u>

Notes: The bold values represent the best ones and the underlined values represent the second.

TABLE VI
COMPARISON OF THE COMPUTATIONAL TIME (IN SECONDS) OBTAINED BY DIFFERENT METHODS FOR THREE REAL IMAGE SEQUENCES

	Sequence 1	Sequence 2	Sequence 3
TopHat [30]	0.1078	0.4400	0.3917
LCM [7]	6.3594	31.6406	32.2969
MPCM [44]	7.0469	36.1719	36.3125
FKRW [36]	4.0313	39.1406	32.7031
IPI [5]	419.7656	2015.3124	2350.9873
PSTNN [10]	6.1719	101.5625	81.3906
IISTD	<u>0.2500</u>	<u>1.2188</u>	<u>1.0000</u>

Notes: The bold values represent the best ones and the underlined values represent the second.

LCM, MPCM, FKRW, IPI, PSTNN, and our IISTD method is 0.0035, 0.2410, 0.2808, 0.1827, 14.9180, 0.3091, and 0.0072 s, respectively.

Furthermore, we also list the computational time (in seconds) obtained by different methods for three real image sequences in Table VI. Obviously, the TopHat method is also the fastest, followed by our IISTD method. In particular, in Sequence 3, the proposed IISTD method takes only one second in total to process 120 frames of images. The proposed method has high real-time performance.

V. CONCLUSION

In this article, a fast and effective infrared small target detection method motivated by image inpainting technique is presented. The main idea of the proposed method is to use rough estimation and fine estimation to reconstruct a more accurate background. First, IOWII is used for rough background estimation. Then, the original image is divided into background components and candidate target components by calculating and evaluating the difference between the original image and the rough background. After removing the candidate target area, only the information of the surrounding background pixels is used to reconstruct the second background, avoiding the target pixels to participate in the fine background estimation. At last, a target saliency map is obtained by subtracting the final estimated background from the original image, and a simple adaptive threshold is adopted to segment the target. The experimental results on infrared images and real image sequences demonstrate that the proposed method outperforms other state-of-the-arts. It is simple and effective, with strong robustness and good real-time performance.

Although the novel method mainly focuses on single-frame detection, it is easy to generalize and apply to sequence-based method because it is well suited for multiframe accumulation and association. We will investigate this issue in the future work.

REFERENCES

- [1] M. Uzair, R. S. Brinkworth, and A. Finn, "Detecting small size and minimal thermal signature targets in infrared imagery using biologically inspired vision," *Sensors*, vol. 21, no. 5, pp. 1–18, 2021.
- [2] D. Pang, T. Shan, P. Ma, W. Li, S. Liu, and R. Tao, "A novel spatiotemporal saliency method for low-altitude slow small infrared target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, Art. no. 7000705.
- [3] P. Wei, J. Zeidler, and W. Ku, "Analysis of multiframe target detection using pixel statistics," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, no. 1, pp. 238–247, Jan. 1995.
- [4] I. S. Reed, R. M. Gagliardi, and L. B. Stotts, "Optical moving target detection with 3-D matched filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 327–336, Jul. 1988.
- [5] C. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4996–5009, Dec. 2013.
- [6] Y. Dai, Y. Wu, F. Zhou, and K. Barnard, "Asymmetric contextual modulation for infrared small target detection," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2021, pp. 950–959.
- [7] C. L. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2014.
- [8] J. Han, C. Liu, Y. Liu, Z. Luo, X. Zhang, and Q. Niu, "Infrared small target detection utilizing the enhanced closest-mean background estimation," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 645–662, 2021.
- [9] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," *Proc. SPIE*, vol. 3809, pp. 74–83, 1999.
- [10] L. Zhang and Z. Peng, "Infrared small target detection based on partial sum of the tensor nuclear norm," *Remote Sens.*, vol. 11, no. 4, 2019, Art. no. 382.
- [11] J. Du et al., "A spatial-temporal feature-based detection framework for infrared dim small target," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 3000412.
- [12] T. Liu et al., "Nonconvex tensor low-rank approximation for infrared small target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–18, 2022, Art. no. 5614718.
- [13] M. Zhao, W. Li, L. Li, J. Hu, P. Ma, and R. Tao, "Single-frame infrared small-target detection: A survey," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 2, pp. 87–119, Jun. 2022.
- [14] Y. Dai, Y. Wu, F. Zhou, and K. Barnard, "Attentional local contrast networks for infrared small target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 11, pp. 9813–9824, Nov. 2021.
- [15] F. Chen et al., "Local patch network with global attention for infrared small target detection," *IEEE Trans. Aerosp. Electron. Syst.*, early access, Mar. 15, 2022, doi: [10.1109/TAES.2022.3159308](https://doi.org/10.1109/TAES.2022.3159308).
- [16] J. Liu, H. Wang, L. Lei, and J. He, "Infrared small target detection utilizing halo structure prior-based local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, Art. no. 6508205.
- [17] J. Han, S. Liu, G. Qin, Q. Zhao, H. Zhang, and N. Li, "A local contrast method combined with adaptive background estimation for infrared small target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 9, pp. 1442–1446, Sep. 2019.
- [18] J. Han, S. Moradi, I. Faramarzi, C. Liu, H. Zhang, and Q. Zhao, "A local contrast method for infrared small-target detection utilizing a tri-layer window," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 10, pp. 1822–1826, Oct. 2020.
- [19] A. Bugeau, M. Bertalmio, V. Caselles, and G. Sapiro, "A comprehensive framework for image inpainting," *IEEE Trans. Image Process.*, vol. 19, no. 10, pp. 2634–2645, Oct. 2010.

- [20] C. Guillemot and O. Le Meur, "Image inpainting: Overview and recent advances," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 127–144, Jan. 2014.
- [21] D. Ding, S. Ram, and J. J. Rodriguez, "Image inpainting using nonlocal texture matching and nonlinear filtering," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1705–1719, Apr. 2019.
- [22] D. Zhang, Z. Liang, G. Yang, Q. Li, L. Li, and X. Sun, "A robust forgery detection algorithm for object removal by exemplar-based image inpainting," *Multimedia Tools Appl.*, vol. 77, no. 10, pp. 11823–11842, 2018.
- [23] I. Katircioglu, H. Rhodin, V. Constantin, J. Sporri, M. Salzmann, and P. Fua, "Self-supervised human detection and segmentation via background inpainting," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Oct. 29, 2021, doi: [10.1109/TPAMI.2021.3123902](https://doi.org/10.1109/TPAMI.2021.3123902).
- [24] T. W. Bae, "Small target detection using bilateral filter and temporal cross product in infrared images," *Infrared Phys. Technol.*, vol. 54, no. 5, pp. 403–411, 2011.
- [25] D. Pang, T. Shan, W. Li, P. Ma, S. Liu, and R. Tao, "Infrared dim and small target detection based on greedy bilateral factorization in image sequences," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 13, pp. 3394–3408, 2020.
- [26] X. Bai and F. Zhou, "Analysis of new top-hat transformation and the application for infrared dim small target detection," *Pattern Recognit.*, vol. 43, no. 6, pp. 2145–2156, 2010.
- [27] S. Kim and J. Lee, "Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track," *Pattern Recognit.*, vol. 45, no. 1, pp. 393–406, 2012.
- [28] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Small infrared target detection based on weighted local difference measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 4204–4214, Jul. 2016.
- [29] Y. Dai and Y. Wu, "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.*, vol. 10, no. 8, pp. 3752–3767, Aug. 2017.
- [30] V. T. Tom, T. Peli, M. Leung, and J. E. Bondaryk, "Morphology-based algorithm for point target detection in infrared backgrounds," *Proc. SPIE*, vol. 1954, pp. 2–11, 1993.
- [31] W. Li, M. Zhao, X. Deng, L. Li, L. Li, and W. Zhang, "Infrared small target detection using local and nonlocal spatial information," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.*, vol. 12, no. 9, pp. 3677–3689, Sep. 2019.
- [32] Y. Dai, Y. Wu, Y. Song, and J. Guo, "Non-negative infrared patch-image model: Robust target-background separation via partial sum minimization of singular values," *Infrared Phys. Technol.*, vol. 81, pp. 182–194, 2017.
- [33] H. Zhu, S. Liu, L. Deng, Y. Li, and F. Xiao, "Infrared small target detection via low-rank tensor completion with Top-Hat regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1004–1016, Feb. 2020.
- [34] M. Zhao, W. Li, L. Li, P. Ma, Z. Cai, and R. Tao, "Three-order tensor creation and Tucker decomposition for infrared small-target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–16, 2022, Art. no. 5000216.
- [35] H. Wang, L. Zhou, and L. Wang, "Miss detection vs. false alarm: Adversarial learning for small object segmentation in infrared images," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 8508–8517.
- [36] C. Gao, L. Wang, Y. Xiao, Q. Zhao, and D. Meng, "Infrared small-dim target detection based on Markov random field guided noise modeling," *Pattern Recognit.*, vol. 76, pp. 463–475, 2018.
- [37] S. Yao, Y. Chang, and X. Qin, "A coarse-to-fine method for infrared small target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 2, pp. 256–260, Feb. 2019.
- [38] M. Zhao, L. Li, W. Li, R. Tao, L. Li, and W. Zhang, "Infrared small-target detection based on multiple morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 6077–6091, Jul. 2021.
- [39] M. M. Hadhoud and D. W. Thomas, "The two-dimensional adaptive LMS (TDLMS) algorithm," *IEEE Trans. Circuits Syst.*, vol. 35, no. 5, pp. 485–494, May 1988.
- [40] T. W. Bae, F. Zhang, and I. S. Kweon, "Edge directional 2D LMS filter for infrared small target detection," *Infrared Phys. Technol.*, vol. 55, no. 1, pp. 137–145, 2012.
- [41] J. Hu, Y. Yu, and F. Liu, "Small and dim target detection by background estimation," *Infrared Phys. Technol.*, vol. 73, pp. 141–148, 2015.
- [42] J. Han, Y. Ma, B. Zhou, F. Fan, K. Liang, and Y. Fang, "A robust infrared small target detection algorithm based on human visual system," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2168–2172, Dec. 2014.
- [43] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Entropy-based window selection for detecting dim and small infrared targets," *Pattern Recognit.*, vol. 61, pp. 66–77, 2017.
- [44] Y. Wei, X. You, and H. Li, "Multiscale patch-based contrast measure for small infrared target detection," *Pattern Recognit.*, vol. 58, pp. 216–226, 2016.
- [45] T. Zhang, H. Wu, Y. Liu, L. Peng, C. Yang, and Z. Peng, "Infrared small target detection based on non-convex optimization with Lp-norm constraint," *Remote Sens.*, vol. 11, no. 5, 2019, Art. no. 559.
- [46] Y. Sun, J. Yang, and W. An, "Infrared dim and small target detection via multiple subspace learning and spatial-temporal patch-tensor model," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 3737–3752, May 2021.



Deyong Lu received the B.S. degree in information and computing science from Xiamen University, Xiamen, China, in 2009, and the M.S. degree in computational mathematics from Lanzhou University, Lanzhou, China, in 2013. He is currently working toward the Ph.D. degree in information and communication engineering with the National University of Defense Technology (NUDT), Changsha, China.

He is currently an Assistant Researcher with the Computational Aerodynamics Institute, China Aerodynamics Research and Development Center, Mi-nyang, China. His research interests include image processing, small target detection, and numerical computing.



Qiang Ling received the B.Eng. degree in measurement engineering and the M.Eng. degree in control science and engineering from Air Force Engineering University, Shaanxi, China, in 2013 and 2016, and the Ph.D. degree in information and communication engineering from the National University of Defense Technology (NUDT), Changsha, China, in 2019.

He is currently a Lecturer with the College of Electronic Science and Technology, NUDT. His research interests include pattern recognition and hyperspectral image processing.



Yuanyuan Zhang received the M.S. degree in operations research and cybernetics from Lanzhou University, Lanzhou, China, in 2013.

She is currently a Lecturer with Tianfu College, Southwestern University of Finance and Economics, Chengdu, China. Her current research interests include image processing and optimization method.



Zaiping Lin received the Ph.D. degrees in information and communication engineering from the National University of Defense Technology (NUDT), Changsha, China, in 2012.

He is currently an Associate Professor with the College of Electronic Science and Engineering, NUDT. His current research interests include infrared image processing and signal processing.



Wei An received the Ph.D. degree in information and communication engineering from the National University of Defense Technology (NUDT), Changsha, China, in 1999.

She was a Senior Visiting Scholar with the University of Southampton, Southampton, U.K., in 2016. She is currently a Professor with the College of Electronic Science, NUDT. She has authored or coauthored more than 100 journals and conference publications. Her current research interests include signal processing and image processing.