



# Article New Results on Small and Dim Infrared Target Detection

Hao Wang<sup>1</sup>, Zehao Zhao<sup>1,2</sup>, Chiman Kwan<sup>3</sup>, Geqiang Zhou<sup>4</sup> and Yaohong Chen<sup>1,\*</sup>

- 1 Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China; wanghao@opt.ac.cn (H.W.); zhaozehao19@ucas.ac.cn (Z.Z.)
- 2 University of Chinese Academy of Sciences, Beijing 100049, China
- 3 Signal Processing, Inc., Rockville, MD 20850, USA; chiman.kwan@signalpro.net
- 4 Astronaut Selection and Training Center, Beijing 100094, China; zhgeqiang@163.com \*
  - Correspondence: chenyaohong@opt.ac.cn

Abstract: Real-time small infrared (IR) target detection is critical to the performance of the situational awareness system in high-altitude aircraft. However, current IR target detection systems are generally hardware-unfriendly and have difficulty in achieving a robust performance in datasets with clouds occupying a large proportion of the image background. In this paper, we present new results by using an efficient method that extracts the candidate targets in the pre-processing stage and fuses the local scale, blob-based contrast map and gradient map in the detection stage. We also developed mid-wave infrared (MWIR) and long-wave infrared (LWIR) cameras for data collection experiments and algorithm evaluations. Experimental results using both publicly available datasets and image sequences acquired by our cameras clearly demonstrated that the proposed method achieves high detection accuracy with the mean AUC being at least 22.3% higher than comparable methods, and the computational cost beating the other methods by a large margin.

Keywords: IR target detection; real-time detection; imaging processing



Citation: Wang, H.; Zhao, Z.; Kwan, C.; Zhou, G.; Chen, Y. New Results on Small and Dim Infrared Target Detection. Sensors 2021, 21, 7746. https://doi.org/10.3390/s21227746

Academic Editor: Kenneth Schepler

Received: 26 September 2021 Accepted: 19 November 2021 Published: 21 November 2021

Publisher's Note: MDPI stavs neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/)

# 1. Introduction

High-altitude aircraft has great potential in early warning and detection, space offense and defense, and electronic countermeasures, which become increasingly significant in a modern battlefield. The ability to detect long-distance dim and small targets is essential for high-altitude aircraft to obtain real-time battlefield information and perceive threats during missions. A high-performance infrared (IR) detection system for dim and small targets is an indispensable instrument onboard high-altitude aircraft due to its key advantages including long detection range and all-weather surveillance capability [1].

The imaging and detection of dim and small infrared targets in high-altitude scenarios have distinctive characteristics. First, the target size is typically less than  $3 \times 3$ , or even only 1 pixel due to long imaging distance. The optical path attenuation and large field of view (FOV) result in low signal-to-noise ratio (SNR) of the output image without any contour and textual features, making it difficult to detect by learning based methods [2–4]. Second, the background of the image is dominated by clouds, which are vulnerable to the negative impacts caused by irradiation and reflection of sunlight or moonlight, and these impacts further increase the difficulty of detection. Third, with the rapid development of infrared detector technology and performance, the latest IR imaging systems have shown a trend of "three highs" (high resolution, high frame rate, and high dynamic range). Consequently, these imaging systems put forward extremely high requirements on the computational efficiency of data processing methods, and a large number of conventional algorithms cannot meet the real-time requirements [5]. Finally, due to the difficulty of data acquisition for research and experiments, there are only a few high-performance IR detection techniques for dim and small targets in high-altitude scenarios.

Due to the large FOV of the situational awareness system, multiple targets with different scales may appear within the effective detection range, and the scale of the target changes as the distance between detection system and target changes. Moreover, fast flying targets have large displacement between adjacent frames, while target flying along the optical axis has small displacement between adjacent frames, making it more difficult to establish a robust sequence-based detection method. Therefore, the single-frame detection algorithm is more suitable for such applications.

Previous works on single-frame detection can be generally divided into three categories: filtering-based methods, algorithms based on image data structure, and local feature-based approaches. The filtering-based methods typically use spatial or frequency domain filters to suppress the background or enhance the targets and then further distinguish the targets from the background. Deshpande et al. [6] proposed to suppress the clutters by max-mean/max-medium filter. Cao et al. [7] improved the two-dimensional least mean square (TDLMS) filter and incorporated neighborhood analysis and data fusion. Bai et al. [8] detected the small target by gradually decreasing the structure element in Top-hat filter. Bae et al. [9] first introduced the bilateral filter to small IR target detection. However, all of the aforementioned algorithms cannot achieve promising results in complex backgrounds.

Detection methods based on image data structure include low-rank sparse decomposition, tensor recovery, robust principal component analysis (RPCA), etc. Gao et al. [10] convert the infrared patch-image (IPI) model into an optimization problem based on sparse matrix. Zhang et al. proposed to detect the small IR target by non-convex rank approximation minimization (NRAM) [11] and partial sum of the tensor nuclear norm (PSTNN) [12]. Dai et al. [13] used a priori information to adaptively assign weights to each column based on the IPI model and proposed a weighted infrared patch image (WIPI) model. Huang et al. [14] proposed to detect targets by density peaks searching and maximum-gray region growing. Qin et al. [15] developed an algorithm based on facet kernel and random walker (FKRW). However, the aforementioned methods have difficulty in achieving robust detection performance in complex and changing scenarios.

The key idea of the local feature-based approaches is to detect targets by exploiting the different characteristics between the targets and the background. Kim et al. [16] and Wang et al. [17] first introduced the Laplacian of Gaussian (LoG) filter and the difference of Gaussian (DoG) filter for IR target detection, respectively. However, the above blob detection filters are susceptible to spot-like backgrounds and blinking pixels. Chen et al. [18] developed local contrast measure (LCM), which enhanced the target significantly but has difficulty in distinguishing targets and strong clutters. Wei et al. [19] developed a multiscale patch-based contrast measurement (MPCM) to improve the processing speed of LCM. Deng et al. [20] proposed a weighted local difference measure (WLCM) to enhance the target and suppress the background simultaneously. Han et al. [21] developed a relative local contrast measure (RLCM). Although a large number of methods are derived based on the idea of LCM, single feature detection has limited performance on low SNR dataset. Consequently, a combination of multiple features such as local intensity and gradient (LIG) [22] properties has become a new trend.

In the application of long-distance imaging systems, previous work generally has the problems of rapidly degrading performance under complex cloud backgrounds and difficulty in meeting the real-time application requirements. The filtering-based methods including TDLMS and Top-hat filter cannot handle complex scenarios [21]. The computational cost of the methods based on image data structure such as IPI and WIPI increases exponentially with the image resolution, while with the NRAM and PSTNN methods it is easy to regard some clutter in complex scenes as the target. Methods based on local features, especially local contrast methods, typically have difficulties in distinguishing the targets from the clutter, especially hot blinking-pixels with strong gray value.

Inspired by [16,17,22], a few of significant features can be used for the small IR target detection. For one thing, the small targets (less than  $3 \times 3$ ), which exhibit pop-out behavior and blob-like morphology, can be remarkably enhanced by the blob detection filter. For another, there are huge local contrast and gradient differences between the target pixels

and background pixels. Consequently, the pop-out behavior and blob-like morphology enable efficient candidate extraction, while the local contrast and gradient properties have advantages indistinguishing real targets from clutters effectively. A recent paper by us [23] exploited the above IR target characteristics and proposed a new detection algorithm for small and dim target detection.

In this paper, we present new results of applying the efficient and effective detection method combining local contrast, blob-like morphology and scale and gradient features in [4]. The key contributions of the proposed method are described as follows:

To further improve the detection performance in real imaging system, the earlier framework for small and dim target detection method in [23] was fine-tuned. Although the computational time is a little more than the earlier version, the background clutter is better suppressed.

To the best of our knowledge, we are the first to develop a high-performance MWIR camera to evaluate the single-pixel target detection performance, which plays a significant role in long range imaging systems.

We present new target detection results using both publicly available datasets and datasets acquired by our own cameras, meaning that the proposed method has promising application value.

The computational cost of our method is still much less than other state-of-the-art approaches.

This paper is organized as follows: Section 2 describes the details of the proposed method. Section 3 summarizes the experimental results and comparison with other methods. Finally, Section 4 concludes the paper with a few future directions.

## 2. Detection Based on the Local Contrast and Gradient Feature

The incorporation of local contrast, blob-like morphology, and scale and gradient features for effective small target detection was proposed in our recent paper [23]. In this section, we include more details of the various steps of target detection. We also fine-tuned our algorithm with more emphasis on background clutter suppression.

Figure 1 illustrates the framework of the proposed method. In the pre-processing stage, we apply a DoG filter to the input image and introduce the cumulative distribution function (CDF) to generate a binarized mask, which contains a small proportion of candidate pixels. The mask enables a fast and accurate overall target detection system. In the detection stage, we calculate the local contrast map by blob-like morphology feature and local gradient map by four quadrant analysis. We also fuse the local contrast map and gradient map. Finally, we accurately extract the targets by performing segmentation.



Figure 1. Framework of the proposed IR small target detection method.

In the next few subsections, we will describe the details of pre-processing stage, detection procedure, and segmentation.

#### 2.1. Pre-Processing and Binary Mask Generation

Figure 2 demonstrates the 3-D and 1-D analysis for small IR targets with different sizes in the sky-cloud scenarios. It is clear that all of the targets including that on the edge of strong clouds (the third column of the Figure 2a) show the blob-like characteristics in the 3-D mesh (Figure 2b) and 1-D cross-section profile analysis (Figure 2c), meaning that the blob detection filters can be used to enhance the target effectively.



**Figure 2.** Demonstration of the 3-D and 1-D analysis for small IR targets with different sizes: (**a**) local patches contain targets; (**b**) 3-D mesh; (**c**) 1-D cross-section profile analysis.

Since the previous application of the Laplacian of Gaussian (LoG) filter [16] and the difference of Gaussian (DoG) filter [17], which are widely used in blob-target detection, we propose to apply the DoG filter to the input image because the DoG is more hardware-friendly than the LoG filter [24]. Let  $\sigma_1$  and  $\sigma_2$ , which are customer-selected parameters, denote the standard deviations of the Gaussian functions; the DoG filter can be explicitly expressed by

$$DoG(i, j, \sigma_1, \sigma_2) = \frac{1}{2\pi\sigma_1^2} e^{-(i^2 + j^2)/2\sigma_1^2} - \frac{1}{2\pi\sigma_2^2} e^{-(i^2 + j^2)/2\sigma_2^2},$$
(1)

The DoG filter is almost the same as the LoG filter [24], which is one of the most popular blob detectors if  $\sigma_2/\sigma_1 = 1.6$ . Since the single-pixel target shows blob-like morphology due to the atmospheric turbulence and the point spread function (PSF) of optical lens, a single-pixel target affects about  $3 \times 3$  pixels in the image. Given that the target sizes typically range from  $3 \times 3$  to  $5 \times 5$  in our applications and the LoG filter has the highest response with a target diameter of  $2\sqrt{2\sigma}$  [25], we set  $\sigma_1 = 3/2\sqrt{2} = 1.06$  and  $\sigma_2 = 1.6 \times \sigma_1 = 1.70$ , respectively. We also found that a minor adjustment of  $\sigma_1$  and  $\sigma_2$  will not significantly affect the results.

The filtering results of the targets in Figure 2 are shown in Figure 3. Let  $I_D$  represent the filtering result of the DoG filter. There is no doubt that the targets are enhanced remarkably in  $I_D$ . However, the DoG filter is susceptible to flickering noise and spot-like backgrounds, which means that high performance detection cannot be obtained by using only the DoG filter. Therefore, we propose to extract candidate target pixels as a binarized mask by calculating the cumulative distribution function (CDF) of  $I_D$  and extracting the pixels with the highest intensity. Let the adjustable parameter *p* represent the proportion of candidate pixels; we found that p = 0.2% works excellent inour experiments. The pseudocode for the

binarized mask generation is illustrated in Algorithm 1. Let  $I_M$  denote the binarized mask, all the subsequent processing stages are guided by  $I_M$ , meaning that the proposed method has great advantage in computational cost.

#### Algorithm 1. Pseudocode for Binarized Mask Calculation.

**Input:** The DoG filtering result  $I_D$ , p = 0.2%, gray dynamic range N **Output:** The binarized mask  $I_M$ 

1:  $[row, col] = size(I_D);$ 2:  $I_M = zeros(row, col);$ 3: **for** *i* = 1 : *row* 4: for j = 1 : col5:  $h(I_D(i,j)+1)++;$ 6: end 7: end  $P(k) = h(k) / (row \times col), k \in [1, N];$ 8: 9: CDF(k) = P(k) + CDF(k-1);10:  $I_M(find(CDF(I_D)) >= (1-p)) = 1$ 



Figure 3. The filtering results of the targets in the Figure 2.

#### 2.2. Local Blob-like Contrast Map

The targets are typically enhanced remarkably by the DoG filter, while the surrounding areas of the targets are drastically suppressed, as shown in Figure 3. Inspired by the characteristics of  $I_D$ , we propose to calculate the local blob-like contrast map by

$$I_{C}(i,j) = (I_{D}(i,j) - \mu_{t}) \times (\mu_{t}/\mu_{b}),$$
(2)

where  $\mu_t$  and  $\mu_b$  are the mean intensity of candidate pixels in the target and surrounding region. The contrast map takes into account both intensity difference and ratio difference between the local target and the surrounding region.

Here, we should state explicitly that the local mean intensity calculation is guided by connected components in the binarized mask. We find all the connected non-zero regions in the mask and define a surrounding region, which is represented by the convolution result of the local operation region and a binarized filter kernel. Let *L* denote a local operation region, which is a rectangle containing the non-zero regions in the mask, and the binarized filter kernel is defined as follows:

$$w(i,j) = if(\sqrt{(i - c_{row})^2 + (j - c_{col})^2} <= r, 1, 0),$$
(3)

where  $(c_{row}, c_{col})$  is the center coordinate of the filter kernel, (i, j) represents the pixel location in the kernel, r is a customer-selected parameter, and 2 is a proper value that has been demonstrated to work well in our experiments. The input image and the binarized

mask are shown in Figure 4a,b, respectively, and Figure 4c represents the candidate target region. Here we regard the combination of the non-zero pixels in the convolution result of the candidate target region and binarized filter kernel w as the surrounding region, which is illustrated in Figure 4e. It should be noted that even if the clutters such as cloud edge may be denoted by the mask with large scale, the parameters of the DoG filter determine that a small target with a size ranging from  $3 \times 3$  to  $5 \times 5$  will achieve a higher enhancement than these clutters, as shown in Figure 4a,b.



**Figure 4.** Illustrations of (**a**) input image, (**b**) binarized mask, (**c**) local operation region, (**d**) binarized filter kernel when r = 2, and (**e**) the surrounding region represented by the mask.

## 2.3. Local Gradient Map

Since the detection method based on a single feature makes it difficult to maintain reliable detection performance in various scenarios, we proposed to improve the detection performance by using the fusion of the local contrast and gradient. In the first step, we detect all the pixels in the mask with non-zero values and generate an adaptive scale for each region. We apply all-ones matrix  $O_m$  to each candidate pixel, with *m* ranging from 2 to the maximum morphological size of the connected region, which consists of non-zeros pixel clusters. The adaptive scale factor *s* of each region is calculated according to the convolution result of all-ones matrix and connected region in the mask, and the scale for the region is assigned to *m* when the maximum value of the convolution is greater than m(m - 1). In order to facilitate the subsequent step, we propose to assign the scale factor 3 to the isolated pixels and clusters with a size less than 3 and adjust the scale factor to an even value by adding 1 if the scale factor is an odd integer.

Inspired by the characteristics described in the introduction, the gradient distribution of an infrared target is typically a few higher than the surrounding areas [16]. We propose to divide the operation region into four quadrants and estimate the gradient features for each quadrant by performing different filters, as shown in Figure 5a. Since filtering results of the directions including 180°, 225°, 270° and 315° can be calculated by inverting the results of directions with 0°, 45°, 90° and 135°, we only perform gradient filtering on four directions, as shown in Figure 5b. Let  $S_{ij}$  represent the sum of elements in matrix obtained by gradient filtering, with the value of *i* is 1 to 4 denoting the four quadrants and the value

of *j* is 1 to 8 denoting  $0^{\circ}$  to  $315^{\circ}$ , respectively. The gradient score of each quadrant can be calculated by

$$G_{Q1} = a \times S_{16} + b \times (S_{15} + S_{17}), \tag{4}$$

$$G_{O2} = a \times S_{28} + b \times (S_{21} + S_{27}), \tag{5}$$

$$G_{O3} = a \times S_{32} + b \times (S_{31} + S_{33}), \tag{6}$$

$$G_{Q4} = a \times S_{44} + b \times (S_{43} + S_{45}), \tag{7}$$



**Figure 5.** (a) Illustration for the operation region for the gradient analysis; (**b**–**e**) gradient filters for different directions.

Both *a* and *b* in Equation (4) to Equation (7) are adjustable parameters and subject to a + 2b = 1. We set a = 0.5 and b = 0.25 in our experiments because  $225^{\circ}$ ,  $315^{\circ}$ ,  $45^{\circ}$  and  $135^{\circ}$  represent the dominant directions of quadrant 1 to 4, respectively. Let  $\mu_{GQ}$  denote the mean gradient feature value of the four quadrants, the intensity of each candidate pixel can be calculated by

$$I_G(i,j) = \frac{\mu_{GQ}}{(1+p_1)} \times p_2,$$
(8)

where and  $p_2$  are penalty factors for the suppression of the strong clutter. The penalty factor  $p_1$  is defined by

1

$$v_1 = \sum d_{Qi}, \quad i \in [1, 4],$$
(9)

where  $d_{Qi}$  denotes the index distance between the index of the dominant direction and the index with the maximum gradient feature value in quadrant 1~4. The parameter  $p_2$  is a binarized parameter used to suppress strong edge interference, and  $p_2$  equals 0 when both the four quadrants has the same maximum directions or at least two quadrants obtain negative scores in the main direction because these two situations are highly related to the appearance of the strong edge. Figure 6a illustrates the index distance  $p_1$  calculation for quadrant 1 with different situations. Figure 6b,c demonstrates the two situations with penalty factor  $p_2$  equals zero.

## 2.4. Fusion and Segmentation

To generate the detection result, we fuse the local contrast map and the local gradient map by simply multiplying  $I_C$  and  $I_G$ :

$$I_{out} = I_{\rm C} \times I_{\rm G}.\tag{10}$$

In the segmentation stage, we propose to segment the detection result by threshold [18]:

$$TH = \mu_O + \mathbf{k} \times \sigma_O, \tag{11}$$

where  $\mu_O$  and  $\sigma_O$  are the mean and standard deviation of the  $I_{out}$ , and k is a constant ranging from 3 to 5. It should be emphasized that the calculation of the  $\mu_O$  and  $\sigma_O$  is only performed for the candidate pixels indicated by the mask.



**Figure 6.** (a) Illustration of index distance calculation for quadrant 1 in different situations; (b) example of the four quadrants having the same maximum gradient directions; (c) example of dominant direction of two quadrants having a negative gradient score.

## 3. Experiments and Discussions

## 3.1. Motivation and Preparation for the Experiments

The IR targets in public datasets are typically bigger than  $3 \times 3$ , and it is hard to acquire images with single-pixel or small-scale target, which are quite significant in the application of the long-distance imaging systems. To evaluate the all-aluminum unobscured two-mirror freeform imaging lens [26] and thermal characteristics of the high-performance camera, we developed a  $3.7 \sim 4.8 \mu$ m MWIR camera and a  $7.7 - 9.7 \mu$ m LWIR camera with a resolution of  $1024 \times 1024$  and  $512 \times 512$ , respectively. The focal lengths of MWIR and LWIR cameras are 21 and 75 mm, respectively. Our dual-band camera is also a good platform for a better implementation and evaluation of the proposed method. Since the detailed hardware description is not the focus of this paper, we mainly introduce the test experiments and results. Given that the LWIR camera is not suitable for single-pixel target detection due to the long focal length, we only use the MWIR camera for the experiment.

We made a target board (description in Figure 7a), which contains one, two, three, and four holes in corresponding quadrants, respectively. We then integrated the target board with a blackbody, a collimator with a focal length of 3 m, and our MWIR camera to simulate the imaging and detection performance of single-pixel target, as shown in Figure 7b. We also collected data under different weather conditions, as shown in Figure 7c.

To evaluate the performance of the proposed method, we applied it to both publicly available datasets and images acquired by our cameras. Given that the filtering-based methods have poor detection performance on data sets with complex background, we regard the local feature-based approaches and image data structure-based method including IPI [10], PSTNN [12], FKRW [15], LCM [18], MCPM [19], and LIG [22] as the baseline. All the methods were performed by running them on a laptop with Intel i5-10500 and 16GB RAM and with MATLAB 2020a.



**Figure 7.** (a) Illustration of the target board, (b) single-pixel target acquisition, and (c) sequence acquisition with cloudy weather.

# 3.2. Comparison of Single-Pixel Target Detection Performance

Since it is quite difficult to acquire images with single-pixel targets, we use target board and collimator to simulate single-pixel target detection. We set the temperature difference between the black body and the ambient temperature to 20 °C and apply a low integration time (1ms for MWIR camera) to simulate low SNR imaging. It should be emphasized that only MWIR cameras can simulate single-pixel imaging and detection due to its large FOV  $(41^{\circ} \times 41^{\circ})$  and small focal length (f = 21 mm). The detection performance comparison of different methods is shown in Figure 8, and the precision-recall curve (PRC) [27] of different methods on single-pixel target detection are shown in the Figure 8h. The target enhancement performance of the LCM is highly related to the target brightness, but the LCM performs poor in terms of clutter suppression. The MPCM improves the LCM's clutter suppression performance at the cost of the target enhancement ability. The data structure-based methods (IPI, PSTNN, and FKRW) have advantages in clutter suppression, but all of these methods are prone to regard the dim target as the background. It is clear that the proposed method enhances the target and suppresses the background clutters simultaneously with all the targets detected, especially for those challenging targets in quadrant three and quadrant four (shown in Figure 8a).

To further compare the robustness of the various methods, we added 20 hot defective pixels with random locations in the single-pixel targets simulation image (Figure 9a) and compared the performances of different methods. The detection performance of all methods has decreased significantly, especially LCM, LIG, and PSTNN have collapsed, which means that these methods are quite sensitive to defective pixel noise. Although our method cannot detect some dim targets, it still maintains an effective suppression of defective pixels.

#### 3.3. Comparison of Sequence Detection Performance

The details of the test sequences are described in Table 1. It should be noted that S1 and S2 are open datasets, while S3 and S4 are acquired by our MWIR and LWIR cameras, respectively. Since the proposed method only performs on a few candidate pixels, it is unfair to use the signal-to-clutter ratio gain (SCRG) and the background suppression factor (BSF) as the evaluation metrics [11]. Recent research showed that for detection application on an imbalanced dataset, the precision–recall curve (PRC) has better performance metrics than the receiver operating characteristic (ROC) curve [28,29], so we used the PRC as the evaluation metrics.



**Figure 8.** The performance comparison of different methods on single-pixel target detection: (a) target image; (b–h) normalized 3-D mesh obtained by the LCM [18], IPI [10], MCPM [19], LIG [22], PSTNN [12], FKRW [15], and the proposed method; and (i) PRC comparison. The red arrows indicate the targets, and the green arrows indicate the noisy clutters.



**Figure 9.** The performance comparison of different methods on single-pixel image with defective pixels: (a) target image; (b–h) normalized 3-D mesh obtained by the LCM [18], IPI [10], MCPM [19], LIG [22], PSTNN [12], FKRW [15], and the proposed method; and (i) PRC comparison.

Sequences	Frame Number	Resolution	Background Description	Target Characteristics
S1	100	$320 \times 256$	Strong cloud	Aircraft with a size of $4 \times 3$
S2	100	$320 \times 256$	Strong cloud with higher gray level	Dim aircraft with a size of $4 \times 2$
S3	50	1024  imes 1024	Strong cloud with irregular shape	Aircraft with a size of $4 \times 3$
S4	50	$640 \times 512$	Cloudy background	birds with size of around $3 \times 3$

Table 1. The details of the test sequences.

The detection results of different methods on sequence 1 are shown in Figure 10, and the target is marked by the red circle. The 3-D mesh demonstrated that the proposed method yielded the best detection performance by enhancing the target and suppressing the clutters simultaneously. The IPI, LIG, PSTNN, and FKRW could not distinguish the target from the defective pixel clusters, while the LCM and MPCM failed in suppressing the clutters.



**Figure 10.** The detection results of different methods on S1: (**a**) the 49th image in S1 and (**b**–**h**) normalized 3-D mesh obtained by the LCM [18], IPI [10], MCPM [19], LIG [22], PSTNN [12], FKRW [15], and the proposed method, respectively.

Figure 11 shows the detection result of a challenging sequence, in which background clutters have a much higher gray value than the target. The LCM was collapsed on S2, while the IPI, MPCM, LIG, and PSTNN cannot distinguish the target from the background because the target has lower intensity than some strong clutters in the background. Although the FKRW can enhance the target effectively, it is clear that our method has a better performance in terms of clutter suppression.



**Figure 11.** The detection results of different methods on S2: (a) the 17th image in S2 and (b–h) normalized 3-D mesh obtained by the LCM [18], IPI [10], MCPM [19], LIG [22], PSTNN [12], FKRW [15], and the proposed method, respectively.

The S3 sequence was acquired by our MWIR camera in good weather conditions with a small aircraft and with low cloud intensity. The results obtained by different methods are shown in Figure 12. The IPI and PSTNN, which are based on sparse representation and low-rank decomposition, have remarkable advantages over the LCM, MPCM, and LIG in terms of clutter suppression. The FKRW suppresses most of the clutter effectively, but still regards a small cloud as the target. The proposed method works well on S3, beating the other methods with a large margin.



**Figure 12.** The detection results of different methods on S3: (**a**) the 1st image in S3 and (**b**–**h**) normalized 3-D mesh obtained by the LCM [18], IPI [10], MCPM [19], LIG [22], PSTNN [12], FKRW [15], and the proposed method, respectively.

The S4 sequence was acquired by our LWIR camera at night, and the number of the birds in S4 ranges from 3 to 5. The bird (target 5) in the lower left corner has small scale and quite low intensity, making it difficult to detect. The results in Figure 13 demonstrated that most of the methods have good detection performance on targets 1 to 4, and only the FKRW and the proposed method can distinguish the target 5 from the clutters effectively.



**Figure 13.** The detection results of different methods on S4: (**a**) the 79th image in S4 and (**b**–**h**) normalized 3-D mesh obtained by the LCM [18], IPI [10], MCPM [19], LIG [22], PSTNN [12], FKRW [15], and the proposed method, respectively.

The precision–recall curves of the four different sequences are shown in the Figure 14. Since the test datasets are challenging, the overall performance of noise-sensitive methods including LCM, MPCM, and LIG is poor, while the performance of the methods based on data structure such as IPI and PSTNN are not robust. It is clear that the proposed method yields the most robust performance on these sequences. Although the PRC of FKRW on S4



is a bit better than our method, FKRW cannot obtain reliable results on S1 and S3. Moreover, our method has a better performance in terms of clutter suppression.

**Figure 14.** The PRC on different sequences: (a) PRC for S1, (b) PRC for S2, (c) PRC for S3, and (d) PRC for S4.

The comparison of the area under the curve (AUC) among different methods is given in Table 2. Although our proposed method does not achieve the largest AUC on the S4 sequence, the overall performance is the most stable, and the mean AUC is at least 22.3% higher than other methods (calculated based on the results of PSTNN).

		LCM	IPI	MPCM	LIG	PSTNN	FKRW	Proposed
Area Under Curve (AUC)	S1 S2 S3 S4 mean	$\begin{array}{c} 0.5763 \\ 0.0001 \\ 0.5032 \\ 0.7892 \\ 0.4672 \end{array}$	0.5442 0.3059 0.8780 0.7075 0.6089	0.3461 0.2168 0.0706 0.6716 0.3263	0.1493 0.1610 0.2014 0.7179 0.3074	0.5963 0.5058 0.8819 0.8082 0.6981	0.2678 0.7339 0.3773 <b>0.9050</b> 0.5710	0.8532 0.8112 0.8835 0.8678 0.8539

 Table 2. The comparison of the AUC among different methods.

We calculate the average time for one frame in different sequences and the comparison of the running time among different methods is shown in Table 3. Although the running time of all methods increases rapidly with the increasement of the resolution, it is clear that our method is the fastest and maintains a faster speed as the resolution increases. The computational cost of IPI and PSTNN increased dramatically due to the substantial increase in matrix size, while the running time of the LCM, MPCM, LIG, and the proposed method has an approximately linear relationship with the resolution. The computational cost of the FKRW is highly related to the data structure, but it is still far from real-time processing.

		LCM	IPI	MPCM	LIG	PSTNN	FKRW	Proposed
Running time (s)	S1 S2 S3 S4	$\begin{array}{c} 0.035 \\ 0.038 \\ 0.546 \\ 0.141 \end{array}$	9.227 7.913 854.3 374.6	0.048 0.047 0.734 0.201	1.726 1.702 24.52 7.386	0.028 0.027 2.153 0.108	$\begin{array}{c} 0.141 \\ 0.075 \\ 1.201 \\ 0.152 \end{array}$	0.024 0.025 0.262 0.079

Table 3. The comparison of the running time among different methods.

### 3.4. Merits and Limitations

The proposed method achieved stable detection performances for small-scale targets on the data set under the sky background. The computational cost of the proposed method is much less than that of the other baseline methods due to the fast extraction by the DoG filter. Moreover, the experimental results demonstrated that the combination of local contrast and gradient can effectively enhance small-scale targets and suppress background clutter. Comparing the results here with our earlier results in [23], the background clutter has been better suppressed at the expense of slightly more computational times.

However, the proposed method also has some limitations. First, missed detection can occur if the pre-processing stage cannot accurately extract the target due to the coarse-to-fine architecture. Second, the local contrast and gradient sometimes may not be good enough to meet the detection requirements of small and dim targets under all sky back-grounds, and it would be better if more local features can be utilized. Third, the proposed method cannot process the high-resolution image ( $1024 \times 1024$ ) in real-time.

## 4. Conclusions

In this paper, we proposed a novel and real-time method to detect a small and dim IR target in a sky background. The experimental results demonstrated that our proposed method achieves robust performance in terms of simultaneous target enhancement and background clutter suppression. More importantly, the proposed method runs much faster than the baseline methods.

Potential future directions include further improving the detection performance by utilizing more local features and accelerating the processing speed by introducing a parallel processing and pipeline architecture, which are essential for real-time processing of data in IR imaging systems.

**Author Contributions:** Conceptualization, Y.C. and C.K.; methodology, H.W.; software, H.W.; validation, Z.Z.; resources, G.Z.; funding acquisition, Y.C. All authors have read and agreed to the published version of the manuscript.

Funding: National Natural Science Foundation of China under Grant 51905529.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

**Data Availability Statement:** Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Acknowledgments: We would like to thank Jinhui Han for providing the open datasets.

Conflicts of Interest: The authors declare no conflict of interest.

## References

- 1. Rogalski, A. New material systems for third generation infrared detectors. *Ninth Int. Conf. Correl. Opt.* 2009, 7388, 73880.
- Dai, Y.; Wu, Y.; Zhou, F.; Barnard, K. Attentional local contrast networks for infrared small target detection. *IEEE Trans. Geosci. Remote Sens.* 2021, 59, 9813–9824. [CrossRef]
- 3. Zhao, B.; Wang, C.; Fu, Q.; Han, Z. A Novel Pattern for Infrared Small Target Detection With Generative Adversarial Network. *IEEE Trans. Geosci. Remote Sens.* 2020, 59, 4481–4492. [CrossRef]
- Hou, Q.; Wang, Z.; Tan, F.; Zhao, Y.; Zheng, H.; Zhang, W. RISTDnet: Robust Infrared Small Target Detection Network. *IEEE Geosci. Remote Sens. Lett.* 2021, 1–5. [CrossRef]

- Reibel, Y.; Espuno, L.; Taalat, R.; Sultan, A.; Cassaigne, P.; Matallah, N. High performance infrared fast cooled detectors for missile applications. *Infrared Technol. Appl. XLII* 2016, *9819*, 98190I.
- Deshpande, S.D.; Er, M.H.; Venkateswarlu, R.; Chan, P. Max-mean and max-median filters for detection of small targets. In Signal and Data Processing of Small Targets 1999; SPIE Press: Bellingham, DC, USA, 1999; Volume 3809, pp. 74–83.
- Cao, Y.; Liu, R.; Yang, J. Small target detection using Two-Dimensional Least Mean Square (TDLMS) filter based on neighborhood analysis. Int. J. Infrared Millim. Waves 2008, 29, 188–200. [CrossRef]
- 8. Bai, X.; Zhou, F. Analysis of new top-hat transformation and the application for infrared dim small target detection. *Pattern Recognit.* **2010**, *43*, 2145–2156. [CrossRef]
- Bae, T.W.; Sohng, K.I. Small target detection using bilateral filter based on edge component. J. Infrared Millim. Terahertz Waves 2010, 31, 735–743. [CrossRef]
- 10. Gao, C.; Meng, D.; Yang, Y.; Wang, Y.; Zhou, X.; Hauptmann, A.G. Infrared patch-image model for small target detection in a single image. *IEEE Trans. Image Process.* **2013**, *22*, 4996–5009. [CrossRef]
- 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.* 2018, 10, 1821. [CrossRef]
- 12. Zhang, L.; Peng, Z. Infrared small target detection based on partial sum of the tensor nuclear norm. *Remote Sens.* **2019**, *11*, 382. [CrossRef]
- 13. Dai, Y.; Wu, Y.; Song, Y. Infrared small target and background separation via column-wise weighted robust principal component analysis. *Infrared Phys. Technol.* **2016**, *77*, 421–430. [CrossRef]
- 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.* 2019, 16, 1919–1923. [CrossRef]
- 15. Qin, Y.; Bruzzone, L.; Gao, C.; Li, B. Infrared small target detection based on facet kernel and random walker. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 7104–7118. [CrossRef]
- 16. Kim, S.; Yang, Y.; Lee, J.; Park, Y. Small target detection utilizing robust methods of the human visual system for IRST. *J. Infrared Millim. Terahertz Waves* **2009**, *30*, 994–1011. [CrossRef]
- 17. Wang, X.; Lv, G.; Xu, L. Infrared dim target detection based on visual attention. Infrared Phys. Technol. 2012, 55, 513–521. [CrossRef]
- Chen, C.L.P.; Li, H.; Wei, Y.; Xia, T.; Tang, Y.Y. A local contrast method for small infrared target detection. *IEEE Trans. Geosci. Remote Sens.* 2014, 52, 574–581. [CrossRef]
- 19. Wei, Y.; You, X.; Li, H. Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recognit.* **2016**, *58*, 216–226. [CrossRef]
- Deng, H.; Sun, X.; Liu, M.; Ye, C.; Zhou, X. Small Infrared Target Detection Based on Weighted Local Difference Measure. *IEEE Trans. Geosci. Remote Sens.* 2016, 54, 4204–4214. [CrossRef]
- Han, J.; Liang, K.; Zhou, B.; Zhu, X.; Zhao, J.; Zhao, L. Infrared Small Target Detection Utilizing the Multiscale Relative Local Contrast Measure. *IEEE Geosci. Remote Sens. Lett.* 2018, 15, 612–616. [CrossRef]
- 22. Zhang, H.; Zhang, L.; Yuan, D.; Chen, H. Infrared small target detection based on local intensity and gradient properties. *Infrared Phys. Technol.* **2018**, *89*, 88–96. [CrossRef]
- 23. Chen, Y.; Zhang, G.; Ma, Y.; Kang, J.U.; Kwan, C. Small Infrared Target Detection Based on Fast Adaptive Masking and Scaling With Iterative Segmentation. *IEEE Geosci. Remote Sens. Lett.* **2021**. [CrossRef]
- 24. Winnemöller, H.; Kyprianidis, J.E.; Olsen, S.C. XDoG: An eXtended difference-of-Gaussians compendium including advanced image stylization. *Comput. Graph.* **2012**, *36*, 740–753. [CrossRef]
- 25. Netsch, T.; Peitgen, H.O. Scale-space signatures for the detection of clustered microcalcifications in digital mammograms. *IEEE Trans. Med. Imaging* **1999**, *18*, 774–786. [CrossRef]
- 26. Xie, Y.; Mao, X.; Li, J.; Wang, F.; Wang, P.; Gao, R.; Li, X.; Ren, S.; Xu, Z.; Dong, R. Optical design and fabrication of an all-aluminum unobscured two-mirror freeform imaging telescope. *Appl. Opt.* **2020**, *59*, 833. [CrossRef]
- 27. Buckland, M.; Gey, F. The relationship between Recall and Precision. J. Am. Soc. Inf. Sci. 1994, 45, 12–19. [CrossRef]
- Davis, J.; Goadrich, M. The relationship between Precision-Recall and ROC curves. In Proceedings of the 23rd International Conference on Machine Learning, New York, NY, USA, 25–29 June 2006; pp. 233–240.
- 29. Saito, T.; Rehmsmeier, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE* 2015, *10*, e0118432. [CrossRef]