Article

Weighted Local Ratio-Difference Contrast Method for Detecting an Infrared Small Target against Ground–Sky Background

Hongguang Wei [1], Pengge Ma [1,*], Dongdong Pang [2], Wei Li [2], Jinwang Qian [1] and Xingchen Guo [1]

1   School of Intelligent Engineering, Zhengzhou University of Aeronautics, Zhengzhou 450015, China
2   Beijing Key Laboratory of Fractional Signals and Systems, School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
*   Correspondence: mapenge@163.com; Tel.: +86-136-7336-3271

Abstract: Fast and robust detection of infrared small targets in a single image has always been challenging. The background residue in complex ground–sky background images leads to high false alarm rates when traditional local contrast methods are used because of the complexity and variability of the ground–sky background imaging environment. A weighted local ratio-difference contrast (WLRDC) method is proposed in this paper to address this problem and detect infrared small targets in the ground–sky background. First, target candidate pixels are obtained using a simple facet kernel filter. Second, local contrast saliency maps and weighted mappings are calculated on the basis of the local ratio-difference contrast and the spatial dissimilarity of the target, respectively. Third, the final weighted mapping can be obtained through the multiplication fusion strategy. Finally, a simple threshold segmentation method is employed to extract the target. Experimental results on six real ground–sky infrared scenes showed that the proposed method outperforms existing state-of-the-art methods.

1. Introduction

Infrared small target detection is an important component of infrared search and tracking systems that has been utilized extensively in space surveillance, remote sensing, missile tracking, and other domains [1–3]. Fast and reliable detection of infrared small targets is crucial in these applications. However, small targets usually contain only a few pixels and lack detailed shape and texture feature information due to long-range target imaging [4–6]. Noise disturbances, including grassland, poles, trees, buildings, and high-brightness clutter on the ground, commonly found in the complex and changeable imaging environment of the ground–sky scene typically result in targets with extremely low signal-to-noise ratios (SNRs). Very low SNRs of targets increase their susceptibility to background interference from nearby ground objects. Therefore, detecting infrared small targets in ground–sky scenes is challenging.

Many different methods have been proposed for detecting infrared small targets in complex backgrounds. Existing methods can be typically divided into single-frame and sequential detection methods [7]. The sequential detection method requires reference to temporal characteristics of multiple successive frames. Background interference clutter is eliminated on the basis of consistency of motion of the moving target and high correlation of the background in adjacent frames [8]. However, background variation is easily caused by the jitter of the photoelectric tower itself in practice, thereby affecting the performance of the sequential detection method. Meanwhile, the computationally complex sequential detection method and requirement of high hardware performance limit its applicability to practical infrared search systems. Compared with the sequential detection method, the

single-frame detection method can more easily obtain satisfactory real-time performance. Traditional single-frame detection methods include Top-Hat [9,10], Max-mean and Max-median filtering [11], two-dimensional least mean square (TDLMS) [12,13], and multi-scale directional filtering approaches [14]. Although these morphology filtering-based methods generally demonstrate satisfactory detection performance in infrared images with uniform background and are easy to implement, they usually cause high rates of false alarms in complex backgrounds. Notably, the method based on a human visual system (HVS) is introduced into the field of infrared small target detection. The background is usually a uniform area with some noise in the local region, while the target and the adjacent region are concentrated and discontinuous. Therefore, the region of interest can be extracted from the difference between features of the target and the background in the local region, and the most appropriate among the responses can be obtained using multi-scale calculations [15]. Chen et al. [16] proposed the initial local contrast method (LCM) for detecting infrared small targets. A series of extended LCM methods, such as an improved LCM (ILCM) [17], a novel local contrast measure (NLCM) [18], a multi-scale relative local contrast measure (RLCM) [19], and a multi-scale patch-based contrast measure (MPCM) [20], was subsequently proposed. Cui et al. [21] put forward a high-speed local contrast infrared small target detection method that achieves fast detection of infrared small targets using local contrast combined with a machine learning classifier to solve the problem of calculation inefficiency of LCM-based methods. Xia et al. [22] designed a multi-scale local contrast measure based on a local energy factor (LEF). Han et al. [23] subsequently proposed a weighted strengthened local contrast measure (WSLCM). Hence, these methods primarily utilize the gray difference between the target and the background in a local region to measure infrared small targets.

Some scholars have improved the window detection framework in LCM-based methods, as shown in Figure 1. For example, Han et al. [24] proposed a multi-scale triple-layer local contrast measurement (TLLCM) method utilizing a new window, and its triple-layer window detection framework is shown in Figure 1b. Wu et al. [25] put forward a double-neighborhood gradient measure for the detection of infrared small targets of different sizes through a designed double-neighborhood window and effective avoidance of the "expansion effect" of traditional multi-scale LCM-based methods. The double-neighborhood window is presented in Figure 1c. Lu et al. [26] devised a weighted double local contrast measure utilizing a new sliding window that further subdivides the central block in the window detection framework and fully considers the contrast information within the central block. This new sliding window is shown in Figure 1d. The improvement of the window detection framework allows the enhanced capture of energy in the central region and contrast with the adjacent background while avoiding multi-scale calculations and improving computational efficiency.
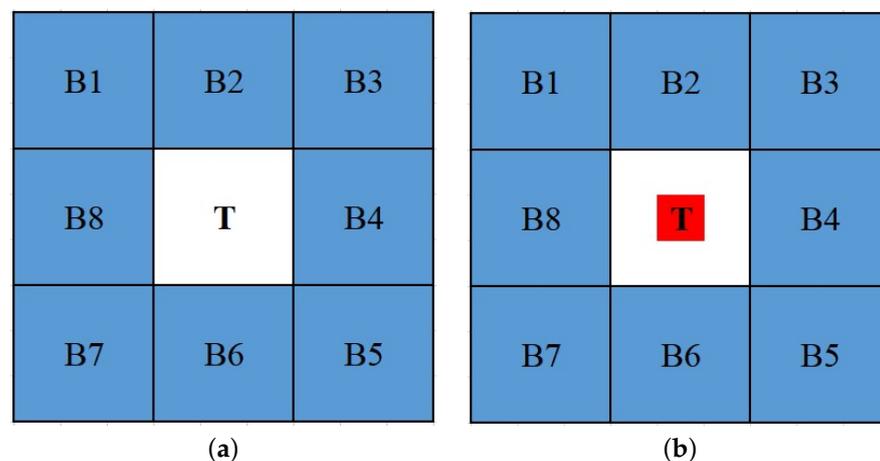

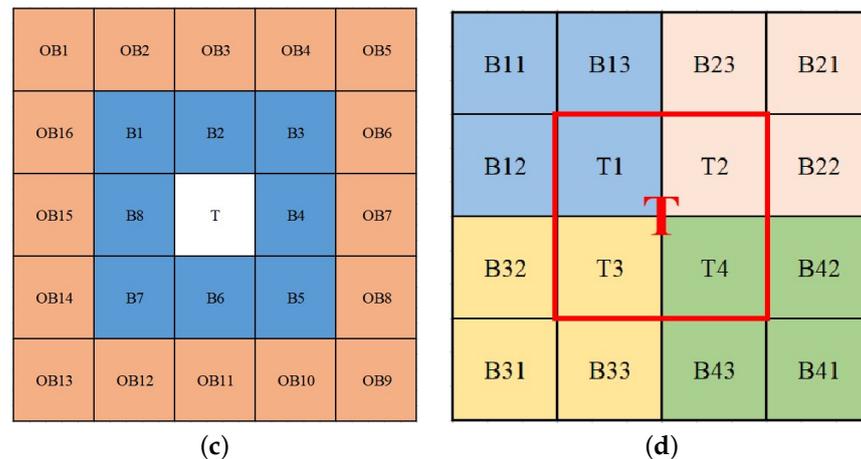(a)   (b)

Figure 1. *Cont.*

OB1	OB2	OB3	OB4	OB5
OB16	B1	B2	B3	OB6
OB15	B8	T	B4	OB7
OB14	B7	B6	B5	OB8
OB13	OB12	OB11	OB10	OB9

(c)

B11	B13	B23	B21
B12	T1	T2	B22
B32	T3	T4	B42
B31	B33	B43	B41

(d)

Figure 1. Window detection framework: (**a**) traditional window detection framework; (**b**) triple-layer window detection framework; (**c**) double neighbor window detection framework; (**d**) double-layer nested window.

At present, methods based on matrix decomposition can successfully improve the detection performance of infrared small targets in complex scenes. The core idea of these kinds of methods is to transform the small target detection problem into a low-rank and sparse optimization problem according to sparse a priori characteristics of targets. Gao et al. [27] established an infrared patch image (IPI) model for infrared small target detection, expressed the small target detection as an optimization problem of low-rank sparse matrix recovery, and solved it effectively using stable principal component tracking. Wang et al. [28] then created a novel stable multi-subspace learning (SMSL) method by analyzing the background data structure. On the basis of the IPI model, Dai and Wu [29] designed an infrared patch tensor (IPT) model, exploited the target sparse prior and background non-local self-correlation prior, and modeled the target–background separation as a robust low-rank tensor recovery problem. On the basis of an infrared patch tensor (IPT) model, a novel non-convex low-rank constraint named partial sum of tensor nuclear norm (PSTNN) [30] with joint weighted ℓ_1 norm was employed to suppress the background and preserve the target efficiently. However, these methods are inefficient in the face of complex backgrounds with multiple or structural sparse targets and commonly result in high false alarm rates.

Many sequential detection methods based on multiple frames have been proposed to detect infrared small targets and achieve enhanced detection results in complex backgrounds with the improvement of computer performance. For instance, Deng et al. [31] and Zhao et al. [32] realized the detection of motion point targets by fusing the spatial and temporal local contrast information. Du et al. [33] proposed a new spatio-temporal local difference measurement method. Liu et al. [34] put forward an infrared video small target detection method based on the spatio-temporal tensor model. Pang et al. [35] established a novel spatio-temporal saliency method for detecting low-altitude slow infrared small targets in image sequences. These sequential detection methods use the information from anterior-posterior multiple frames to enhance the suppression of background clutter and extraction of targets. However, the output results of these sequential detection methods are usually lagging.

With the development of deep learning (DL) technology, many DL-based methods have been put forward to detect infrared small targets. Wang et al. [36] designed a YOLO-based feature extraction backbone network for infrared small target detection. Subsequently, an infrared small target detection method based on Generative Adversarial Network (GAN) was proposed in [37]. Dai et al. [38] incorporated a visual attention mechanism into a neural network to improve target detection performance. Kim et al. [39] utilized a GAN framework to obtain synthetic training datasets, and the detection performance was effectively improved. Zuo et al. [40] designed the attention fusion feature pyramid network

specifically for small infrared target detection. The method based on deep learning has achieved some relatively good detection results in infrared small target detection. However, the above methods usually require rich infrared image data. Regrettably, there are few publicly available datasets to support the research of these methods.

A weighted local ratio-difference contrast (WLRDC) method is proposed in this paper to detect infrared small targets and improve the robust detection performance in ground–sky complex backgrounds. Specifically, the innovation of the proposed WLRDC method is as presented follows:

1. A local ratio-difference contrast (LRDC) method that can simultaneously enhance the target and suppress complex background clutter and noise is proposed by combining local ratio information and difference information. LRDC uses the mean of the *Z* max pixel gray values in the center block to effectively solve the problem of poorly enhancing the target at low contrast when the traditional LCM-based method is applied.
2. A simple and effective strategy of block difference product weighted (BDPW) mapping is designed on the basis of spatial dissimilarity of the target to improve the robustness of the WLRDC method. BDPW can further suppress background clutter residuals without increasing the computation complexity given that this strategy is also calculated using the gray of the center and adjacent blocks.

The rest of the paper is organized as follows: in Section 2, we present the related work. The proposed WLRDC method is described in Section 3. In Section 4, we conduct extensive experiments in various scenes to verify the effectiveness of the proposed method. Finally, the paper is discussed and summarized in Sections 5 and 6, respectively.

2. Related Work

In recent years, some infrared small target detection methods based on LCM have been extensively investigated. Existing local contrast calculation types can usually be divided into three types: ratio, difference, and ratio-difference (RD) form. We briefly review the RD-based form of infrared small target detection method in this section because the proposed method is based on the RD form.

RLCM [19] is a typical method of the RD form method that uses the average gray of the K_1 and K_2 max pixels in the central and adjacent cell blocks for calculation to suppress the interference of various noises effectively. Guan et al. [41] utilized a multi-scale Gaussian filter combined with RD contrast and proposed an infrared small target detection method. Although the RD-based contrast method achieved relatively excellent detection performance compared to the traditional LCM-based method, the detection capability was weak in complex backgrounds. Moreover, the weighted function was developed to improve the detection capability and robustness of the method further. For instance, Han et al. [42] put forward a multi-directional two-dimensional least mean square product weighted RD-LCM method for infrared small target detection, which has better detection performance for different types of backgrounds and targets. Subsequently, Han et al. [43] also offered a weighted RD local feature contrast method with improved robustness to noise. With the combination of weighting function with the RD contrast, the detection performance of these methods was effectively improved.

The RD form of the contrast calculation method integrates the advantages of both the ratio and difference forms, enhances the target, and effectively suppresses background clutter. Although the weighting function is an excellent method for improving the detection performance, existing weighting functions are overly complex and increase the computational complexity. Therefore, rapidly and reliably detecting infrared dim and small targets in complex backgrounds remains an important challenge to overcome.

3. Materials and Methods

The overall target detection process of the proposed WLRDC method is shown in Figure 2. This process is mainly divided into the following stages: (1) preprocessing (facet kernel filtering and square calculation are used to enhance the target) and (2) detection

(LRDC is first calculated from the local ratio-difference contrast and on the basis of spatial dissimilarity of the target to measure the BDPW) stages. WLRDC mapping is obtained from the fusion multiplication strategy of the LRDC and BDPW to extract the target easily using a simple threshold segmentation method.

3.1. Preprocessing: Target Enhancement

3.1.1. Facet Kernel Filtering

The target presents low contrast in the ground–sky scene due to the existence of high-brightness buildings, grass, roads, and other strong interference clutter in the ground–sky background. Here, facet kernel [44] filtering is employed to highlight the target area significantly different from the adjacent background by traversing the whole image. In this way, targets of different sizes and shapes can be effectively enhanced. The literature demonstrates the effectiveness of the facet kernel [45,46]. The target-enhanced image H is expressed as follows:

$$H = I(x, y) * F \tag{1}$$

where $*$ represents the convolution operation, and F denotes the facet kernel.
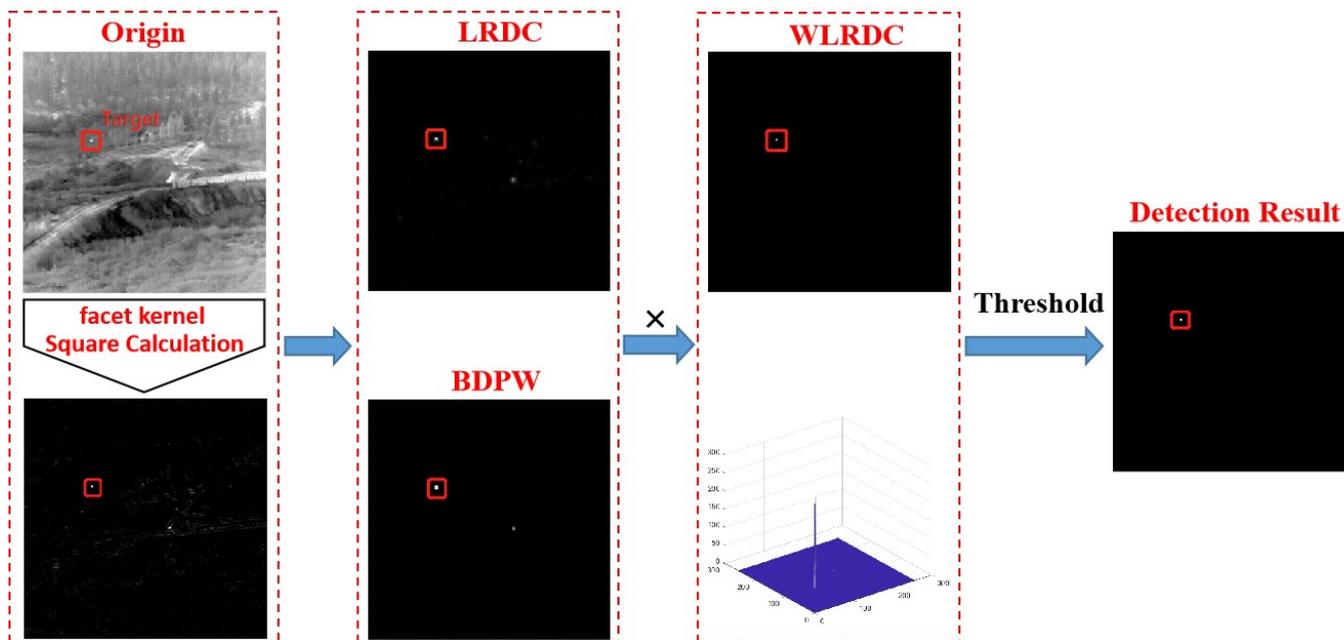

Figure 2. The overall target detection process of the proposed WLRDC method. The detection process includes facet kernel filtering, calculation of LRDC and BDPW, and WLRDC obtained by fusing the LRDC and BDPW.

The facet kernel model uses a bivariate cubic function to approximate the gray intensity surface formed by gray values of all pixels in its neighborhood. Here, the size used is 5×5, and its two-dimensional discrete orthogonal polynomial set Pi can be expressed as follows:

$$P_i \in \left\{1, r, c, r^2 - c, rc, c^2 - 2, r^3 - \frac{17}{5}r, (r^2 - 2)c, r(c^2 - 2), c^3 - \frac{17}{5}c\right\} \tag{2}$$

where $r \in \{-2, -1, 0, 1, 2\}$ and $c \in \{-2, -1, 0, 1, 2\}$ are the respective row and column coordinates in the 5×5 neighborhood. The pixel surface function $f(r,c)$ in the neighborhood is expressed as follows:

$$f(r, c) = \sum_{i=1}^{10} K_i \cdot P_i(r, c) \tag{3}$$

where K_i is the fitting coefficient of the polynomial that can be derived using the least square method as follows:

$$K_i = \frac{\sum_r \sum_c P_i(r,c) f(r,c)}{\sum_r \sum_c P_i^2(r,c)} \tag{4}$$

K_i can be calculated using the convolution of $f(r,c)$ and W_i. W_i is expressed as follows:

$$W_i = \frac{P_i(r,c)}{\sum_r \sum_c P_i^2(r,c)} \tag{5}$$

According to Equations (2) and (3), second-order partial derivatives of the central pixel $(0,0)$ of the $r \times c$ window along the row and column directions can be obtained as follows:

$$\frac{\partial^2 f(r,c)}{\partial r^2} = 2K_4, \frac{\partial^2 f(r,c)}{\partial c^2} = 2K_6 \tag{6}$$

Equation (6) shows that the sum of K_4 and K_6 can be calculated to enhance the target region effectively. K_4 and K_6 can be obtained from the convolution of W_i with $I(x,y)$. Notably, the relative results of filtering for each region of the image remains unaffected because of the denominator term in Equation (7). Thus, we can further eliminate denominator coefficients and substitute $P_i(r,c)$ into Equation (5) to obtain kernel coefficients.

$$W_4 = \frac{1}{70} \begin{bmatrix} 2 & 2 & 2 & 2 & 2 \\ -1 & -1 & -1 & -1 & -1 \\ -2 & -2 & -2 & -2 & -2 \\ -1 & -1 & -1 & -1 & -1 \\ 2 & 2 & 2 & 2 & 2 \end{bmatrix}, W_6 = W_4^T \tag{7}$$

Therefore, $F = -2(W_4 + W_6)$. The weighted kernel coefficients are inverted to detect bright targets in the ground–sky background where the central region presents greater gray areas than the adjacent regions. Enabling the detection of bright targets where the central gray is greater than the adjacent gray leads to the positive weight of the central coefficients. The final facet kernel is expressed as follows:

$$F = \begin{bmatrix} -4 & -1 & 0 & -1 & -4 \\ -1 & 2 & 3 & 2 & -1 \\ 0 & 3 & 4 & 3 & 0 \\ -1 & 2 & 3 & 2 & -1 \\ -4 & -1 & 0 & -1 & -4 \end{bmatrix} \tag{8}$$

3.1.2. Square Calculation

The square of H is calculated to improve the energy of the target further and obtain the enhanced target response map R as follows:

$$R = H^2 \tag{9}$$

Figure 3 shows that the target energy enhances but the noise also enhances after the calculation of the square for each pixel in H. However, the target is maximally salient in H. The square calculation will obtain greater energy relative to the noise.
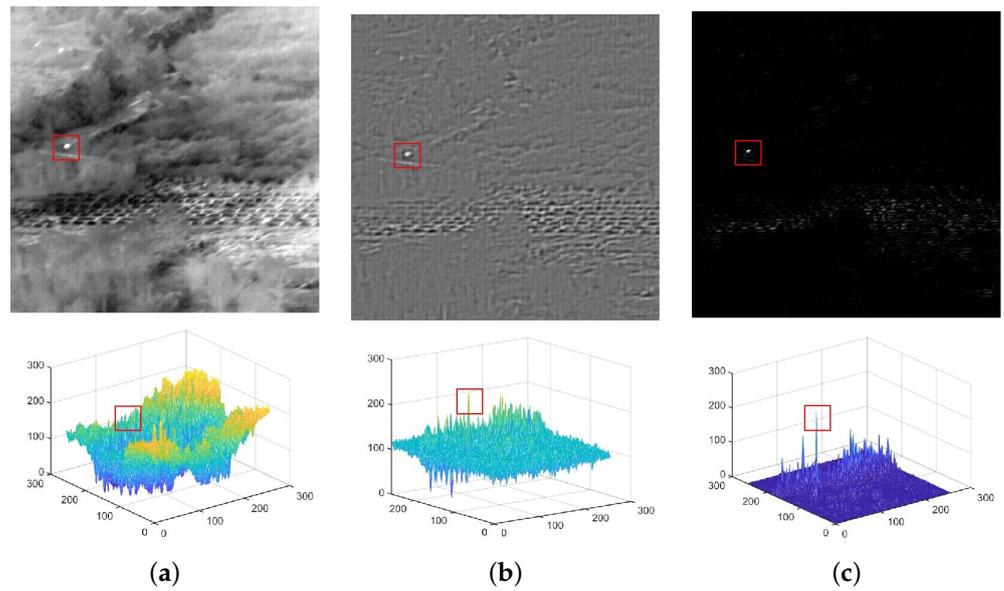
Figure 3. Target enhancement process. (**a**) original image; (**b**) result of facet kernel filtering; (**c**) result of square calculation. Red boxes indicate the targets.

3.2. Calculation of LRDC

We utilize local features of the target and background to calculate the local ratio-difference contrast; that is, the ratio and difference information of the central and eight directional adjacent blocks within the local region are applied to calculate the LRDC saliency map. Here, a sliding window with nine unit blocks is used to traverse the whole image, as shown in Figure 4. The LRDC is expressed as:

$$LRDC(x,y) = \frac{S_Z}{m_{\max}}(D_{\max} - D_{\min}) \tag{10}$$

where S_Z is the mean of the Z maximum pixel gray values in the center block. $D_{\max}$ and $D_{\min}$ are the maximum and minimum values of D_i, respectively; $m_{\max}$ is the maximum of m_i. Values of Z on different scales, which will be discussed in Section 5, can be obtained as follows:

$$S_Z = \frac{1}{Z} \sum_{j=1}^{Z} G_T^j \tag{11}$$

$$m_i = \frac{1}{L^2} \sum_{j=1}^{L^2} I_i^j, (i = 1, 2, \ldots, 8) \tag{12}$$

$$D_i = (S_Z - m_i)^2, (i = 1, 2, \ldots, 8) \tag{13}$$

where Z, L, G_T^j, and I_i^j indicate the number of maximum pixel gray values in the center block, length of the block, jth maximum gray value of the center block T, and gray value of the jth pixel in the ith block, respectively.
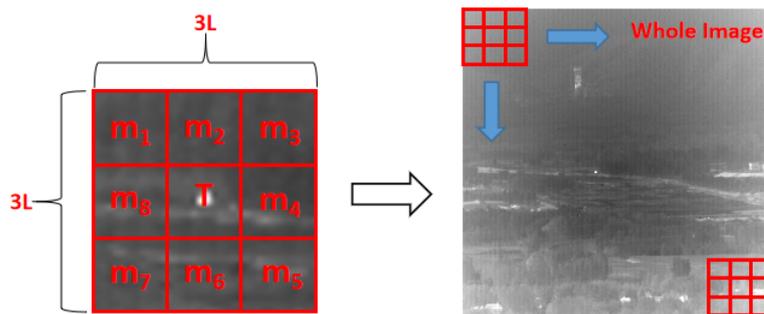
Figure 4. Structure of the nested sliding window, where *L* is the length of the block.

We will now analyze the motivation for the design of the LRDC. The local ratio-difference contrast combination can simultaneously enhance the target and suppress complex backgrounds. Different from traditional local contrast methods, the proposed method utilizes the S_Z/m_i in the ratio calculation and is mainly based on the following facts:

1. False alarms easily occur when pixel-sized noises with high brightness (PNHB) appear in the background, the maximum gray value of the center block is used in the ratio calculation, and PNHB is easily taken as the target. S_Z improves the accurate representation of gray features of the central block and avoids the weighting of PNHB in the calculation of LRDC.

2. Compared with the method that only uses the gray mean of the central block, our method uses the mean of Z maximum gray values in the center block S_Z to expand the contrast between the target and the background further as well as enhance the target.

The difference calculation presents the advantage of effectively eliminating high-brightness backgrounds. The contrast is calculated using the difference between S_Z of the central block and m_i of adjacent blocks to reduce the effect of high-brightness backgrounds effectively. Section 3.1.2 demonstrated that the squared calculation enhances the energy of the target. Therefore, $(S_Z - m_i)^2$ is used to calculate D_i. Meanwhile, background clutter can be further attenuated with $D_{\max} - D_{\min}$.

3.3. Calculation of BDPW

The weighted mapping mainly exploits spatial features of the target and the background. Targets exhibit discontinuities of features with surrounding backgrounds in local regions and minimal similarities to the background given that they usually satisfy a Gaussian distribution. Hence, the increased dissimilarity between the central and adjacent regions indicate a high probability that the central region is the target. The spatial dissimilarity between the central and adjacent blocks is used as the weighting function for the LRDC to suppress noise residuals in the LRDC further. Weighted mapping is defined as:

$$BDPW(x,y) = \prod_{i=1}^{8}(S_Z - m_i) \tag{14}$$

We will discuss the weighted enhancement mechanism of BDPW in different local regions.

1. If the central block is the target, then we can easily obtain the following because the target is the most significant in the local region:

$$S_Z - \max m_i > 0, \ (i = 1, 2, \ldots, 8) \tag{15}$$

$$BDPW > 0 \tag{16}$$

2. If the central block is the background, then we can easily obtain the following because the background is a uniform area with some noise in the local region:

$$S_Z - \max m_i \leq 0, \ (i = 1, 2, \ldots, 8) \tag{17}$$

$$BDPW \leq 0 \tag{18}$$

As shown in Equations (15) and (17), if the central region is the target, then the LRDC will obtain a large weight value. If the central region is the background, then the LRDC will obtain a negative weight value. Therefore, BDPW can realize the weighted enhancement of LRDC.

3.4. Multi-Scale Calculation of WLRDC

The local ratio-difference contrast (LRDC) and the block difference product weight (BDPW) are obtained by calculations. WLRDC is defined as:

$$WLRDC(x, y) = LRDC(x, y) \times BDPW(x, y) \tag{19}$$

The obtained WLRDC saliency map must be normalized to facilitate the subsequent threshold segmentation. The normalization result is expressed as follows:

$$WLRDC(x, y) = \frac{LRDC(x, y) \times BDPW(x, y)}{\max\{LRDC(x, y) \times BDPW(x, y)\}} \tag{20}$$

The sliding detection window is the vital component of the LCM-based method. Ideally, the detection window should be the same size as the target size for a more accurate measurement of local contrast. In practical applications, the dataset may contain targets of multiple sizes. However, the size of the detection window is fixed. To solve this problem, a multi-scale calculation method was adopted in [16–20]. Typically, the detection window is set to 3×3, 5×5, 7×7 and 9×9. Multi-scale operation is necessary because the size of targets in infrared images cannot be determined in advance. Multi-scale WLRDC is defined as:

$$WLRDC = \max\{WLRDCn(x, y)\}, n = 1, 2, \ldots, s \tag{21}$$

where n and s represent the nth scale and the total number of scales, respectively.

3.5. Target Extraction

The real target is properly enhanced according to the previous calculations and is most significant in the WLRDC saliency map. Thus, the target can be extracted using the following threshold segmentation operation:

$$Th = \mu + k\sigma \tag{22}$$

where μ and σ are the mean and standard deviation of the WLRDC saliency map, respectively, and k is an adjustable parameter.

In summary, the whole process of WLRDC calculation is described in Algorithm 1.

Algorithm 1 Detection steps of the proposed WLRDC method.

　Input: Infrared image I, size of sliding window block $L{\times}L$, facet kernel, and parameters Z, k and s
　Output: Detection result

　1:　Calculate the facet kernel filter map H using (1).
　2:　Calculate the enhanced target response map R using (9).
　3:　**for** 1 to s **do**
　4:　　Calculate the S_Z, m_i and D_i by (11)–(13), respectively.
　5:　　Calculate the LRDC saliency map using (10).
　6:　　Calculate the weighted mapping BDPW using (14).
　7:　　Obtain WLRDC mapping by fusing LRDC and BDPW.
　8:　**end for**
　9:　Obtain maximum pooling for multi-scale WLRDC using (21).
　10:　Obtain the final detection result using (22).

3.6. Complexity Analysis

We briefly analyze the computation complexity of the proposed method in this section. Suppose the size of the input infrared image is M × N. The preprocessing process consists of facet kernel filtering (kernel operator size is q × p) and square calculation with a computation complexity of $\mathcal{O}(qpMN)$ and $\mathcal{O}(MN)$, respectively. The saliency map needs to be computed pixel-by-pixel in the target detection stage, and the computation complexity of the whole process is $\mathcal{O}(n^2MN)$, where $n(n = 1, 2, \ldots, s)$ is the scale of the sliding window. The total computational complexity of the multi-scale target detection stage is $\mathcal{O}(s^3MN)$. In conclusion, the whole computation complexity of the proposed method is $\mathcal{O}(qpMN + s^3MN)$.

4. Experimental Results and Analysis

A series of evaluation metrics is used to evaluate the detection performance of different methods in this section to validate the effectiveness and robustness of the proposed method. Six real scenes were utilized in the experiments, and the proposed method was compared with the baseline methods. We adjust the parameters of the baseline method in the next experiments to achieve the optimal detection results. Finally, we analyze the computational efficiency of the proposed and baseline methods.

4.1. Experimental Setup

4.1.1. Datasets

In this paper, we use the public dataset from the ATR Key Laboratory of the National University of Defense Technology and the 25th Institute of the Second Research Institute of China Aerospace Science [47]. The dataset contains multiple image sequences with complex and variable imaging backgrounds, including cluttered grass, high-brightness ground, and forests. The sensor used for dataset acquisition was a cooled mid-wave infrared camera. We choose six typical ground–sky scenes for analysis to validate the efficiency and robustness of the proposed method. Note that the background in Scene 1 contains high-brightness dotted noise similar to the real target. Some targets in Scenes 4 and 5 are mixed with the high-luminance ground background, thereby significantly increasing the difficulty of detection. A detailed description of all datasets is presented in Table 1. Each scene contains one target and is marked with a red box.

Table 1. Details of the six test datasets.

Data	Number of Frames	Image Resolution	Background Description	Target Type
Scene 1	259	256 × 256	Ground–sky background, high voltage towers, and strong radiation buildings	UAV
Scene 2	151	256 × 256	Ground–sky background, high-brightness roads, and forests	UAV
Scene 3	131	256 × 256	Ground–sky background, grasslands, and strong radiation ground	UAV
Scene 4	75	256 × 256	Ground–sky background, trees, and high-brightness ground	UAV
Scene 5	100	256 × 256	Ground–sky background, telegraph poles, and high-brightness ground	UAV
Scene 6	150	256 × 256	Ground–sky background, forests, and strong ground disturbance clutter	UAV

4.1.2. Evaluation Criteria

Several evaluation metrics commonly used in the field of infrared small target detection are introduced to evaluate the performance of different infrared small target detection methods quantitatively. Signal-to-noise ratio gain (SNRG) is usually used to describe the target enhancement ability of methods and related to the SNR before and after image processing. SNR is expressed as follows:

$$SNR = (I_{max} - I_{mean})/\sigma \tag{23}$$

where I_{max} is the maximum gray value of the image, I_{mean} is the mean value of the image, and σ is the standard deviation. The SNRG is defined as follows:

$$SNRG = 20 \times log_{10}(SNR_{out}/SNR_{in}) \tag{24}$$

where SNR_{out} and SNR_{in} represent the SNR of the original image and the output result, respectively.

Background suppression factor (BSF) is used to describe the background suppression ability of the corresponding method as follows:

$$BSF = \frac{C_{in}}{C_{out}} \tag{25}$$

where C_{in} and C_{out} represent the standard deviation of the original image and the output result, respectively. Receiver operating characteristic (ROC) curve is used as a common evaluation index to quantify the effectiveness of methods at the pixel level. The ROC curve represents the relationship between the probabilities of detection (P_d) and false alarm (P_f). P_d and P_f are expressed as follows:

$$P_d = \frac{n_t}{N_t}, P_f = \frac{n_f}{N} \tag{26}$$

where n_t, N_t, n_f, and N represent the number of true pixels detected in the target image, number of target pixels in the original image, number of false alarm pixels, and the total number of pixels, respectively.

4.1.3. Baseline Methods

We choose eight existing advanced approaches as baseline methods to demonstrate the effectiveness and robustness of the proposed method. Among them, Top-Hat [9] is

based on filtering methods. The HVS-based methods include a multi-scale relative local contrast measure (RLCM) [19], a multi-scale patch-based contrast measure (MPCM) [20], a tri-layer local contrast measure (TLLCM) [24], a multi-scale local contrast measure based on local energy factor (LEF) [22], and a weighted strengthened local contrast measure (WSLCM) [23]. Infrared patch image (IPI) [27] and partial sum of tensor nuclear norm (PSTNN) [30] are low-rank sparse decomposition-based methods. Parameter settings of the baseline method were obtained from experimental analysis and discussions in the paper of the original authors. Parameter settings of all methods are listed in Table 2.

Table 2. Parameter settings of the different methods.

Methods	Parameter Settings
Top-Hat [9]	Structure size: square, local window size: 3×3
RLCM [19]	$(k_1, k_2) = (2,4),(5,9)$ and $(9,16)$
MPCM [20]	Local window size:$N = 3,5,7,9$. mean filter size: 3×3
IPI [27]	Patch size:50×50, sliding step: 10, $\lambda = 1/\sqrt{min(m,n)}$, $\varepsilon = 10^{-7}$
TLLCM [24]	Window size: 3×3, $s = 5,7,9$
LEF [22]	$P = 1,3,5,7,9$, $\alpha = 0.5$, and h = 0.2
PSTNN [30]	Patch size: 40×40, sliding step: 40, $\lambda = 1/\sqrt{min(m,n)}$, $\varepsilon = 10^{-7}$
WSLCM [23]	$K = 9$, $\lambda = 0.6{\sim}0.9$
Proposed	Local window size:$L = 3,5,7,9$, $K = 4,9,11$

4.2. Comparison with State-of-the-Art Methods

The results of different methods were processed independently in six groups of real ground–sky backgrounds to more clearly illustrate the detection capability of the proposed method clearly. The findings of the visualization of different methods are illustrated in Figures 5–7. Targets are marked with a red box, unmarked areas indicate that no target is detected, and a close-up depiction of the target is shown in the bottom right corner of the image. The corresponding three-dimensional (3D) display diagram is presented below the detection result of each method to show the effectiveness and robustness of the proposed method intuitively. As shown in Figures 5–7, the traditional Top-Hat method properly suppresses regions with a uniform background but poorly suppresses regions with strong interference background, thereby resulting in poor detection results. LCM-based methods focus on increasing the target background contrast and suppressing background clutter. However, accurately segmenting targets is difficult when targets are mixed in the strong interference background clutter. The method based on low-rank sparse decomposition focuses on how to separate the target from the background and can easily lead to false alarms when suspicious target areas exist in the background. The proposed method suppresses suspicious target regions through spatial dissimilarity weighting of the local contrast and achieves accurate segmentation of the target.
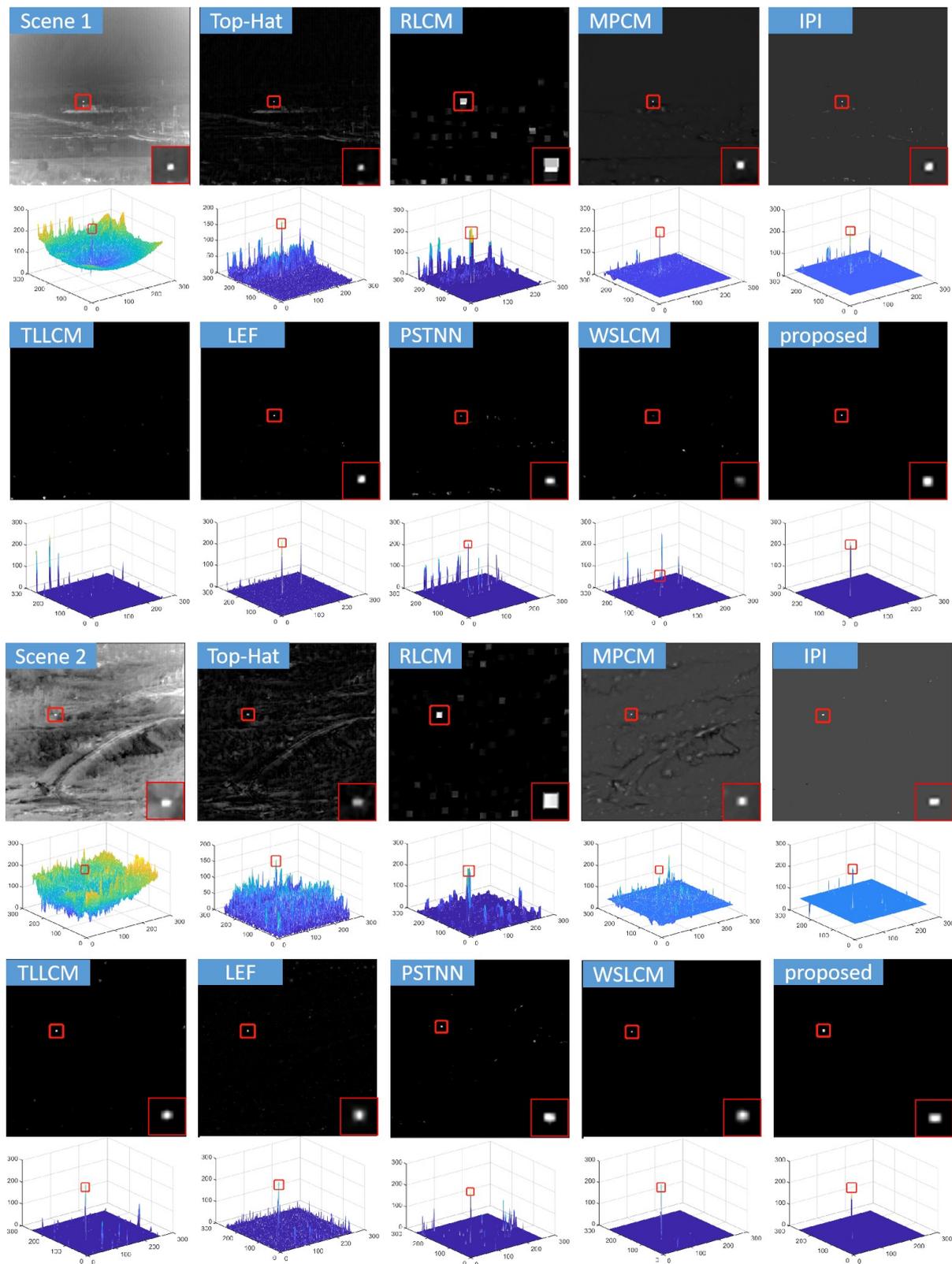
Figure 5. Detection results and 3D distribution of different methods in Scenes 1 and 2. The target is marked with a red box, and a close-up of the target is shown in the lower right corner. Unmarked areas indicate the absence of detection of the target.
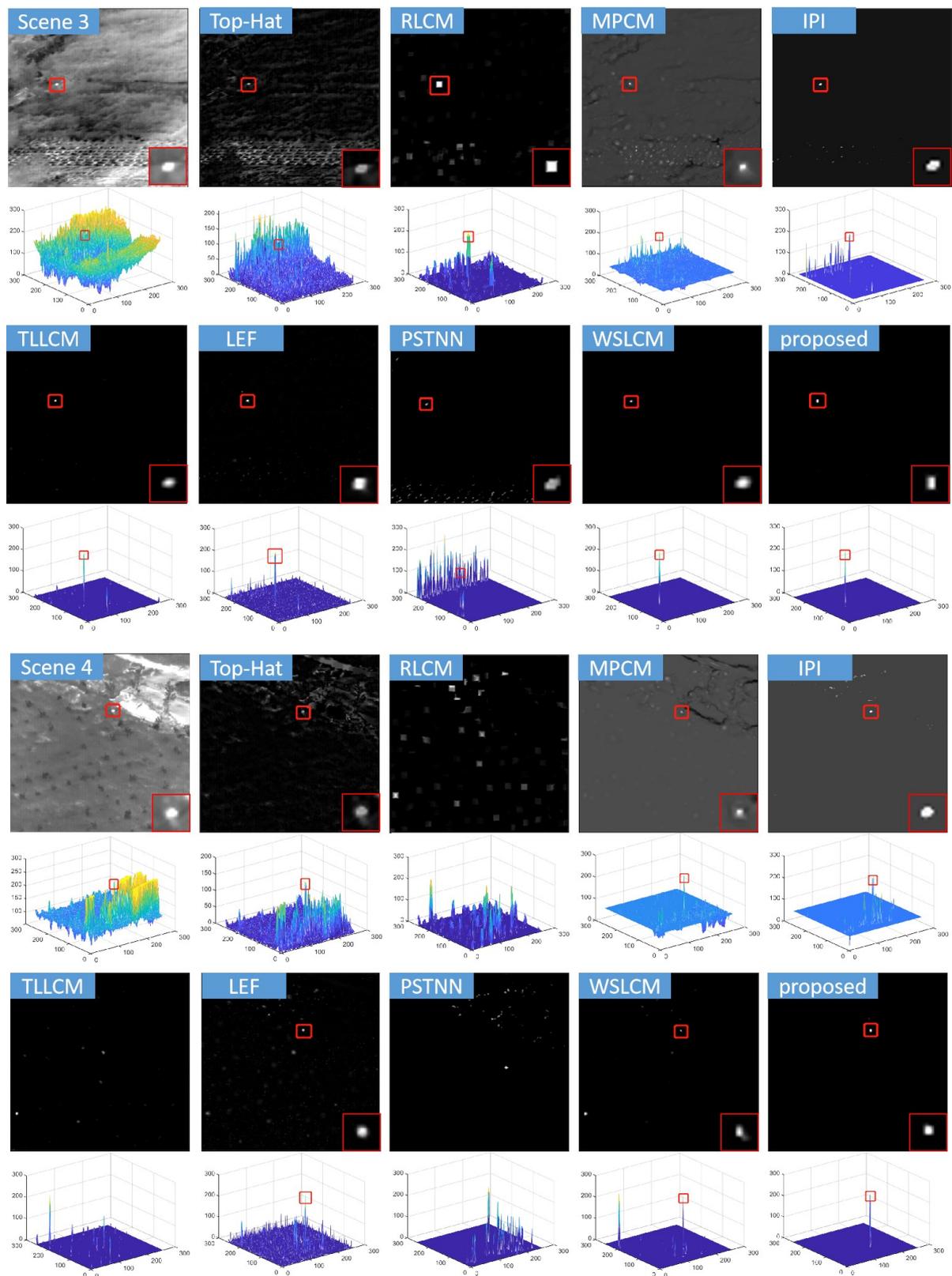
Figure 6. Detection results and 3D distribution of different methods in Scenes 3 and 4. The target is marked with a red box, and a close-up of the target is shown in the lower right corner. Unmarked areas indicate the absence of detection of the target.

Figure 7. Detection results and 3D distribution of different methods in Scenes 5 and 6. The target is marked with a red box, and a close-up of the target is shown in the lower right corner. Unmarked areas indicate the absence of detection of the target.
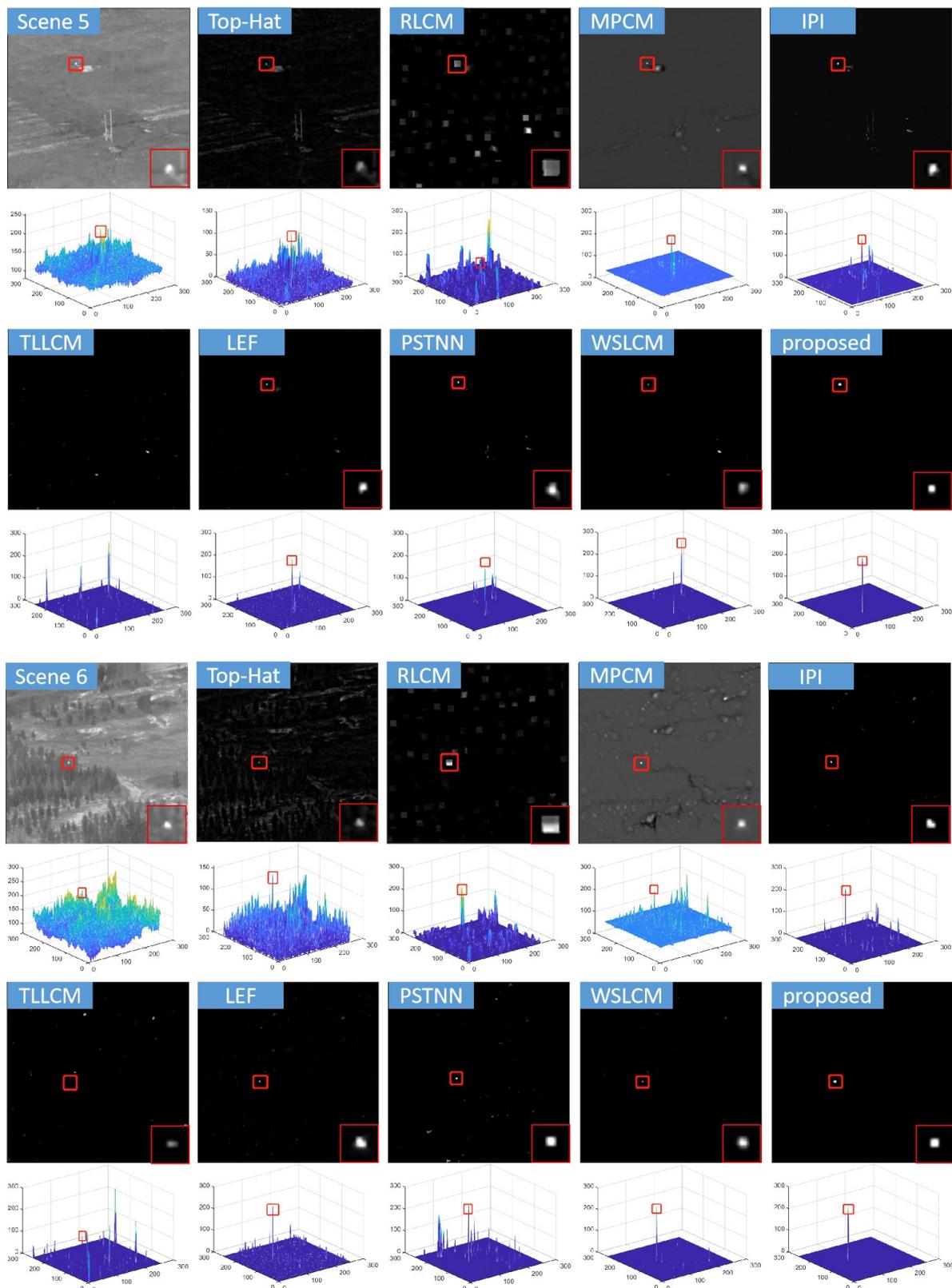
Strong interference clutter in Scene 1 is widely distributed, and the detection results of baseline methods all contain background noise residuals. Among them, the TLLCM method

failed to detect the target. The background of Scenes 2–4 depicts messy grass and high-brightness ground. The 3D distribution demonstrates that a small amount of background noise is still retained in the WSLCM detection results. The increased background clutter in the detection results lead to a high false alarm rate although other baseline methods can detect the target. In particular, the target in Scene 4 is undetected by RLCM, TLLCM, and PSTNN due to the close proximity of the target to the high-brightness background, while the proposed method accurately detects the target. The background in Scene 5 is uniform despite the presence of suspicious targets. The suspicious target also exists although the real target is identified in the detection results of the baseline method. Specifically, the TLLCM only retains the false target and misses the real target. The background in Scene 6 contains trees and strong-interference ground clutter. Moreover, the 3D distribution shows that that IPI and PSTNN result in high false alarm rates when strong edges or sparse points exist in the background. Although LCM-based methods can successfully detect targets, many significant background residuals are observed in the detection results because the local contrast measurement poorly describes the suspected non-target region and results in a high false alarm rate in the presence of strong interference. These visualization results indicate that background clutter is nearly nonexistent in the detection results of the proposed method, and the target is accurately extracted compared with those of baseline methods. Moreover, experiments based on six groups of ground–sky scenes verified the robustness of the proposed method.

SNRG and BSF were used to evaluate the target enhancement and background suppression abilities of the proposed and baseline methods. High values of SNRG and BSF indicate the enhanced performance of the corresponding method. Tables 3 and 4 show the SNRG and BSF of different methods in six different ground–sky backgrounds, where numbers in bold font indicate the maximum value of SNRG and BSF in each scene and underlined numbers denote the second highest values. The proposed method achieved the maximum SNRG and BSF in Scenes 2–5. Notably, high SNRG and BSF values are mainly concentrated in TLLCM, LEF, WSLCM, and the proposed method. These methods can improve the target enhancement and background suppression by combining local contrast measurements with weighting functions. The LEF and IPI methods present the maximum SNRG and BSF values in Scenes 1 and 6, respectively, but the background clutter that still exists in the detection results leads to inaccurate segmentation of the target.

Table 3. Average SNRG values of different methods in six real scenes.

Methods	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5	Scene 6
Top-Hat [9]	30.321	23.568	20.404	10.105	6.059	15.975
RLCM [19]	26.046	30.587	28.373	21.395	8.476	18.487
MPCM [20]	30.796	38.533	37.942	24.485	18.996	22.243
IPI [27]	35.465	38.994	34.064	23.766	16.049	**31.107**
TLLCM [24]	33.302	41.840	39.910	30.099	17.976	29.228
LEF [22]	**38.445**	40.831	37.878	28.637	18.818	30.620
PSTNN [30]	35.573	38.407	29.373	26.285	17.017	24.576
WSLCM [23]	36.353	<u>42.122</u>	<u>40.456</u>	<u>31.293</u>	<u>19.414</u>	30.478
Proposed	<u>38.023</u>	**42.847**	**42.550**	**33.966**	**20.636**	<u>30.793</u>

Table 4. Average BSF values of different methods in six real scenes.

Methods	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5	Scene 6
Top-Hat [9]	24.167	7.264	5.468	2.920	2.009	5.958
RLCM [19]	15.735	15.615	13.402	10.747	2.719	7.868
MPCM [20]	28.897	45.049	43.421	15.294	9.957	12.717
IPI [27]	43.060	41.818	25.871	13.589	6.414	**33.793**
TLLCM [24]	33.967	57.377	50.029	27.964	8.159	28.127
LEF [22]	**60.306**	51.201	39.703	23.685	8.875	32.058
PSTNN [30]	43.872	39.179	16.295	18.037	7.146	16.031
WSLCM [23]	50.165	59.105	53.180	32.643	9.443	32.262
Proposed	57.183	**64.301**	**67.949**	**43.330**	**10.753**	32.355

Figure 8 shows the ROC curves of different methods in six real groups of ground–sky scenes. A method typically exhibits satisfactory detection performance when the ROC curve is close to the upper left corner. The proposed method achieved the best detection performance in each of the six real scenes compared with the baseline method. Meanwhile, the detection performance of PSTNN is poor and its false alarm rate is high mainly due to the complex background of the dataset used in the experiments and the susceptibility of PSTNN to treat strong interference clutter and suspicious targets as sparse points. Notably, the classical Top-Hat method achieved satisfactory detection performance in the test due to the appropriate match achieved between the selected structural elements and the target. In addition, although all LCM-based methods present high detection rates, they also show high false alarm rates mainly due to the presence of some high-brightness spots in the ground background that exhibit high contrast in the local background, thereby creating a false perception of the target. The proposed method achieves high detection performance in six scenes, especially when weighted mapping is used to suppress false alarm targets further.
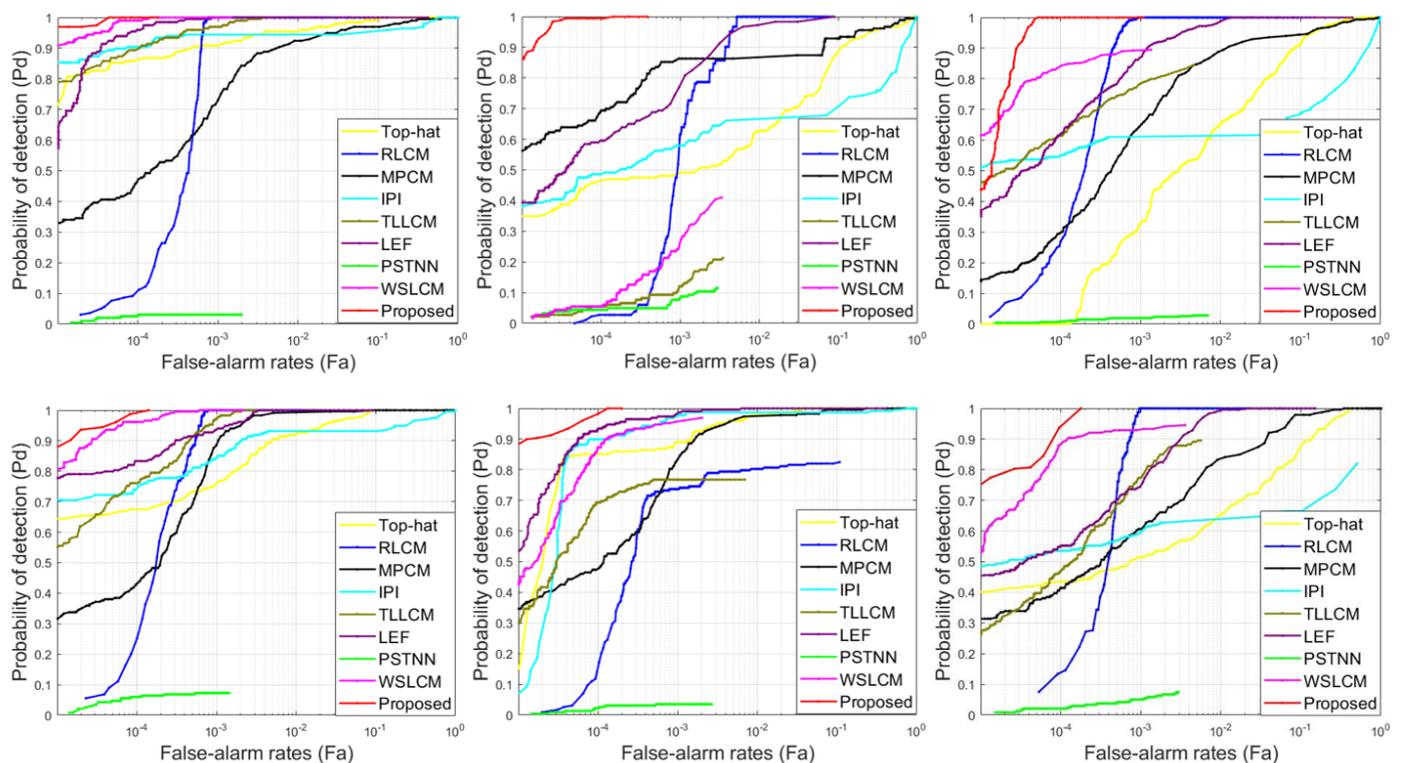
Figure 8. ROC curves of different methods in six real scenes.

All experimental programs were compiled using MATLAB R2018a and conducted on a computer with a 2.3 GHz Intel i5-GTX 950 M GPU and 8 G of memory. We tested the time consumption of different methods in six scenes to verify the computational efficiency of the proposed method. The time consumption of a single-frame for different methods in Scenes 1–6 are listed in Table 5 , where the numbers in bold font indicate the minimum time consumption (maximum calculation efficiency) and underlined numbers indicate the calculation efficiency that ranks second. Methods based on morphology and local contrast measures are more computationally efficient than those based on low-rank sparse decomposition because morphology, and LCM-based methods only compute the pixel gray matrix for local regions of the image without additional computational complexity. However, methods based on low-rank sparse decomposition lead to low computational efficiency because they require singular value decomposition in every iteration. Table 5 demonstrates that MPCM and LEF achieve the highest and lowest computational efficiencies, respectively. The MPCM calculates local contrast measurements by simultaneously traversing the image through eight patches, thereby minimizing the computation time. The LEF consumes a significant amount of time in calculating the local energy factor, thereby increasing the time consumption costs. LCM-based RLCM, TLLCM, and WSLCM methods are time-consuming due to their need to traverse the whole image when calculating local contrast measures. IPI and PSTNN are methods based on low-rank sparse decomposition, with PSTNN demonstrating the minimum consumption of time. The proposed method ranks second in terms of computational efficiency because WLRDC needs to traverse the image Z times when calculating S_Z and perform multi-scale calculations, thereby seriously affecting the computational efficiency.

Table 5. Average time consumption of a single frame for all methods in six real scenes. (Unit: s).

Methods	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5	Scene 6
Top-Hat [9]	0.594	0.564	0.544	0.544	0.542	0.547
RLCM [19]	4.297	4.423	4.478	4.387	4.536	4.371
MPCM [20]	**0.162**	**0.138**	**0.133**	**0.126**	**0.131**	**0.122**
IPI [27]	9.266	8.826	8.939	9.327	9.863	8.922
TLLCM [24]	1.535	1.277	1.432	1.476	1.285	1.233
LEF [22]	19.252	19.612	20.744	19.430	20.156	19.443
PSTNN [30]	0.299	0.257	0.302	0.315	0.267	0.324
WSLCM [23]	5.269	4.807	5.478	4.626	5.444	5.294
Proposed	<u>0.216</u>	<u>0.217</u>	<u>0.217</u>	<u>0.219</u>	<u>0.221</u>	<u>0.220</u>

5. Discussion

5.1. Discussion of Detection Performance

Because the imaging environment is complex and variable and the ground–sky background usually contains substantial interference clutter and noise, the detection of infrared small targets in the ground–sky scene is extremely difficult. Traditional Top-Hat filtering is extremely sensitive to edges and noise, resulting in inability to segment targets accurately, especially in complex ground–sky scenes. Strong interference clutter is widely distributed in the ground–sky scenes, and small targets are in a complex background leading to low contrast. Therefore, MPCM and RLCM methods that only use local contrast calculation are easy-to-miss targets. IPI and PSTNN are quite sensitive to point noise. For example, it incorrectly includes background clutter and sparse point noise into the target image in Figure 5, resulting in a high false alarm rate. The TLLCM, LEF. and WSLCM methods recently proposed combine local contrast and weighted mapping, and the anti-interference capability is significantly improved. However, the detection results of these methods are frequently full of noise and background clutter interference, leading to a huge challenge in separating the real target in postprocessing.

In this paper, we innovatively adopt S_Z/m_i to solve the problem of difficult target enhancement in low contrast conditions. Meanwhile, $D_{max} - D_{min}$ is used to eliminate

complex background clutter and noise. Based on the local dissimilarity of the targets, a weighting function is proposed, which further suppresses the residual noise and extracts the targets accurately by fusing with LRDC. Since the proposed method requires pixel-by-pixel computation of gray features, especially in terms of a large dataset, when the image resolution increases, the computation increases as well. We perform a complexity analysis of the proposed method in Section 3.6. The real-time performance is also related to the hardware of the computer. In addition, the proposed method has low computational complexity and can be accelerated by GPU or field-programmable gate array (FPGA).

5.2. Discussion of the Key Parameter Z

We briefly discuss the selection of the key parameter Z in the proposed method in this section. Z in Equation (10) is the key parameter that directly determines the quality of the generated target saliency map. Many simulation experiments were performed at different scales to select the optimal Z values separately. First, we selected images with different sizes of targets as the test dataset. Note that sizes of targets should be less than or equal to 3×3, 5×5, 7×7, and 9×9. Second, simulation experiments were carried out at different scales with the corresponding settings of different Z values. We used ROC curves to measure the detection performance of the proposed method at different Z values. Figure 9 shows the ROC curves of the target under different sliding windows.

As shown in Figure 9, for smaller targets (3×3), we recommend Z be set from 4 to 6; for medium targets (5×5), we recommend Z be set from 9 to 13, and for larger targets (7×7 or 9×9), we recommend Z be set from 11 to 15. A large Z value indicates large time consumption because the Z value directly affects the computational efficiency of the proposed method. Therefore, the Z values are set to 4 (sliding window size is 3×3), 9 (sliding window size is 5×5), and 11 (sliding window size is 7×7 and 9×9).
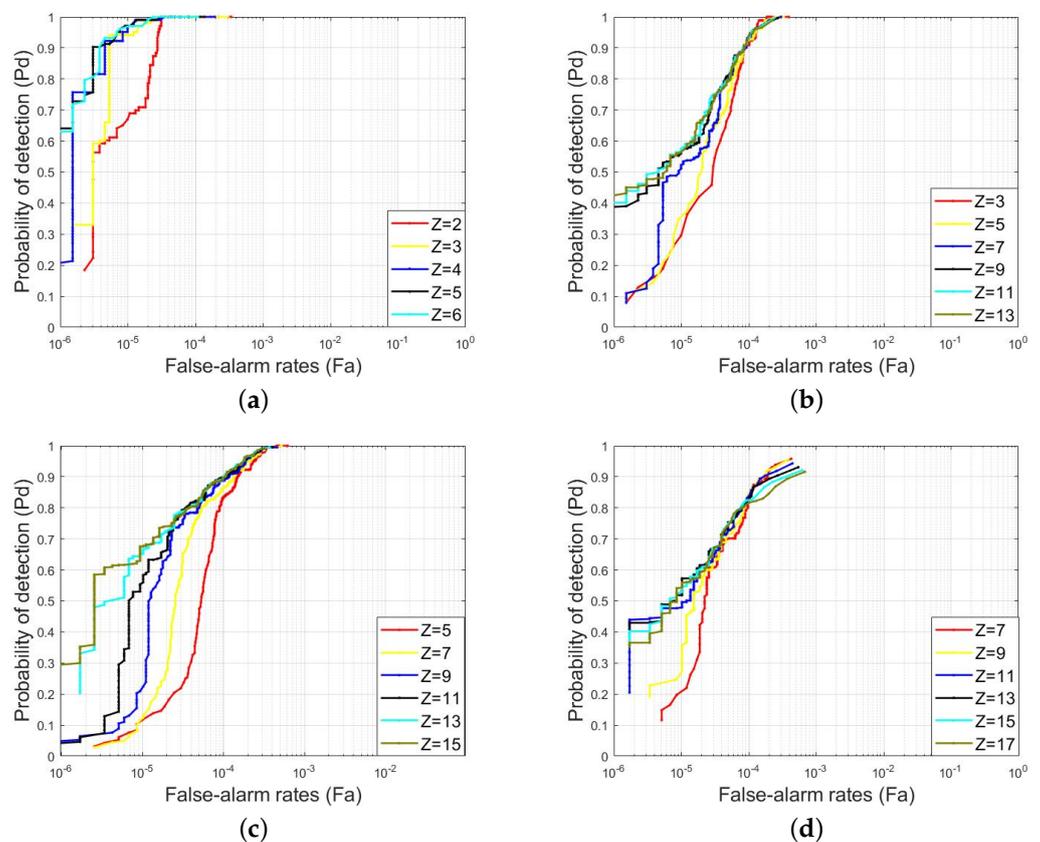

Figure 9. The ROC curves for different Z values in four windows: sliding window of (**a**) 3×3; (**b**) 5×5; (**c**) 7×7; and (**d**) 9×9.

6. Conclusions

In this paper, a weighted local ratio-difference contrast method was developed to detect infrared small targets in ground–sky background. First, facet kernel filtering and square calculation were used to obtain the enhanced target candidate pixels. Second, we used the local ratio-difference contrast and spatial dissimilarity of the target during the calculation of the WLRDC saliency map to suppress the complex background and enhance the real target. Finally, the effectiveness of the proposed method was verified in six real ground–sky scenes. The experimental results demonstrated that the proposed method can achieve efficient infrared small target detection in the ground–sky background and present clear advantages in a series of evaluation indexes.

However, the drawback of the proposed single-frame method is that the detection performance is poor when suspicious targets are presented in the background. In our future work, we will consider using temporal features of multi–frame images combined with local contrast to further suppress strong interference clutter in complex backgrounds. In addition, we will also verify the detection performance of the proposed method in multi–target scenarios.

References

1. Zhang, T.; Wu, H.; Liu, Y.; Peng, L.; Yang, C.; Peng, Z. Infrared Small Target Detection Based on Non-Convex Optimization with Lp-Norm Constraint. *Remote Sens.* **2019**, *11*, 559. [CrossRef]
2. Chen, Y.; Xin, Y. An Efficient Infrared Small Target Detection Method Based on Visual Contrast Mechanism. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 962–966. [CrossRef]
3. Han, J.; Yong, M.; Huang, J.; Mei, X.; Ma, J. An Infrared Small Target Detecting Algorithm Based on Human Visual System. *IEEE Geosci. Remote Sens. Lett.* **2015**, *13*, 452–456. [CrossRef]
4. Pang, D.D.; Shan, T.; Li, W.; Ma, P.G.; Tao, R.; Ma, Y.R. Facet Derivative-Based Multidirectional Edge Awareness and Spatial–Temporal Tensor Model for Infrared Small Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–15. [CrossRef]
5. Li, Z.Z.; Chen, J.; Hou, Q.; Fu, H.X.; Dai, Z.; Jin, G.; Li, R.Z.; Liu, C.J. Sparse representation for infrared dim target detection via a discriminative over-complete dictionary learned online. *Sensors* **2014**, *14*, 9451–9470. [CrossRef] [PubMed]
6. Deng, H.; Wei, Y.; Tong, M. Small target detection based on weighted self-information map. *Infrared Phys. Technol.* **2013**, *60*, 197–206. [CrossRef]
7. Zhou, F.; Wu, Y.; Dai, Y.; Wang, P. Detection of Small Target Using Schatten 1/2 Quasi-Norm Regularization with Reweighted Sparse Enhancement in Complex Infrared Scenes. *Remote Sens.* **2019**, *11*, 2058. [CrossRef]
8. Lu, Y.; Dong, L.; Zhang, T.; Xu, W. A Robust Detection Algorithm for Infrared Maritime Small and Dim Targets. *Sensors* **2020**, *20*, 1237. [CrossRef] [PubMed]
9. Zhou, J.; Lv, H.; Zhou, F. Infrared small target enhancement by using sequential top-hat filters. *Proc. Int. Symp. Optoelectron. Technol. Appl.* **2014**, *9301*, 417–421.
10. Zeng, M.; Li, J.; Peng, Z. The design of top-hat morphological filter and application to infrared target detection. *Infrared Phys. Technol.* **2006**, *48*, 67–76. [CrossRef]
11. Deshpande, S.D.; Meng, H.E.; Ronda, V.; Chan, P. Max-mean and Max-median filters for detection of small-targets. *Proc. SPIE Int. Soc. Opt. Eng.* **1999**, *3809*, 74–83.

12. Fan, H.; Wen, C. Two-Dimensional Adaptive Filtering Based on Projection Algorithm. *IEEE Trans. Signal Process.* **2004**, *52*, 832–838. [CrossRef]

13. Zhao, Y.; Pan, H.; Du, C.; Peng, Y.; Zheng, Y. Bilateral two dimensional least mean square filter for infrared small target detection. *Infrared Phys. Technol.* **2014**, *65*, 17–23. [CrossRef]

14. Peng, L.B.; Zhang, T.F.; Liu, Y.H.; Li, M.H.; Peng, Z.M. Infrared dim target detection using shearlet's kurtosis maximization under non-uniform background. *Symmetry* **2019**, *11*, 723. [CrossRef]

15. Nie, J.Y.; Qu, S.C.; Wei, Y.T.; Zhang, L.M.; Deng, L.Z. An Infrared Small Target Detection Method Based on Multiscale Local Homogeneity Measure. *Infrared Phys. Technol.* **2018**, *90*, 186–194. [CrossRef]

16. Chen, C.L.; Li, H.; Wei, Y.T.; Xia, T.; Tang, Y.Y. A local contrast method for small infrared target detection. *IEEE Trans. Geosci. Remote Sens.* **2013**, *52*, 574–581. [CrossRef]

17. Han, J.H.; Ma, Y.; Zhou, B.; Fan, F.; Liang, K.; Fang, Y. A robust infrared small target detection algorithm based on human visual system. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 2168–2172.

18. Qin, Y.; Li, B. Effective infrared small target detection utilizing a novel local contrast method. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1890–1894. [CrossRef]

19. Han, J.H.; Liang, K.; Zhou, B.; Zhu, X.Y.; Zhao, J.; Zhao, L.L. Infrared small target detection utilizing the multi-scale relative local contrast measure. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 612–616. [CrossRef]

20. Wei, Y.T.; You, X.G.; Li, H. Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recogn.* **2016**, *58*, 216–226. [CrossRef]

21. Cui, Z.; Yang, J.; Jiang, S.; Li, J. An infrared small target detection algorithm based on high-speed local contrast method. *Infrared Phys. Technol.* **2016**, *76*, 474–481. [CrossRef]

22. Xia, C.Q.; Li, X.R.; Zhao, L.Y.; Shu, R. Infrared Small Target Detection Based on Multiscale Local Contrast Measure Using Local Energy Factor. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 157–161. [CrossRef]

23. Han, J.H.; Moradi, S.; Faramarzi, I.; Zhang, H.H.; Zhao, Q.; Zhang, X.J.; Li, N. Infrared Small Target Detection Based on the Weighted Strengthened Local Contrast Measure. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 1670–1674. [CrossRef]

24. Han, J.; Moradi, S.; Faramarzi, I.; Liu, C.; Zhang, H.; Zhao, Q. A local contrast method for infrared small-target detection utilizing a tri-layer window.*IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1822–1826. [CrossRef]

25. Wu, L.; Ma, Y.; Fan, F.; Wu, M.H.; Huang, J. A Double-Neighborhood Gradient Method for Infrared Small Target Detection. *IEEE Geosci. Remote Sens. Lett.* **2021**, *18*, 1476–1480. [CrossRef]

26. Lu, X.F.; Bai, X.F.; Li, S.X.; Hei, X.H. Infrared Small Target Detection Based on the Weighted Double Local Contrast Measure Utilizing a Novel Window. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]

27. Gao, C.Q.; Meng, D.Y.; Yang, Y.; Wang, Y.T.; Zhou, X.F.; Hauptmann, A.G. Infrared patch-image model forsmall target detection in a single image. *IEEE Trans. Image Process.* **2013**, *22*, 4996–5009. [CrossRef] [PubMed]

28. Wang, X.Y.; Peng, Z.M.; Kong, D.H.; He, Y.M. Infrared dim and small target detection based on stable multi-subspace learning in heterogeneous scenes. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5481–5493. [CrossRef]

29. Dai, Y.M.; Wu, Y. Reweighted Infrared Patch-Tensor Model With Both Nonlocal and Local Priors for Single-Frame Small Target Detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2017**, *10*, 3752–3767. [CrossRef]

30. Zhang, L.; Peng, Z. Infrared small target detection based on partial sum of the tensor nuclear norm. *Remote Sens.* **2019**, *11*, 382. [CrossRef]

31. Deng, L.Z.; Zhu, H.; Tao, C.; Wei, Y.T. Infrared moving point target detection based on spatial–temporal local contrast filter. *Infrared Phys. Technol.* **2016**, *76*, 168–173. [CrossRef]

32. Zhao, B.; Xiao, S.; Lu, H.; Wu, D. Spatial-temporal local contrast for moving point target detection in space-based infrared imaging system. *Infrared Phys. Technol.* **2018**, *95*, 53–60. [CrossRef]

33. Du, P.; Askar, H. Infrared Moving Small-Target Detection Using Spatial-Temporal Local Difference Measure. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1817–1821. [CrossRef]

34. Liu, H.K.; Zhang, L.; Huang, H. Small Target Detection in Infrared Videos Based on Spatio-Temporal Tensor Model. *IEEE Trans. Geosci. Remote Sens.* **2020**, *58*. 8689–8700. [CrossRef]

35. Pang, D.D.; Shan, T.; Ma, P.G.; Li, W.; Liu, S.H.; Tao, R. A Novel Spatiotemporal Saliency Method for Low-Altitude Slow Small Infrared Target Detection. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]

36. Wang, K.D.; Li, S.Y.; Niu, S.S.; Zhang, K. Detection of Infrared Small Targets Using Feature Fusion Convolutional Network. *IEEE Access.* **2019**, *7*, 146081–146092. [CrossRef]

37. Wang, H.; Zhou, L.; Wang, L. Miss Detection vs. False Alarm: Adversarial Learning for Small Object Segmentation in Infrared Images. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea, 27 October–2 November 2019; pp. 8508–8517.

38. Dai, Y.M.; Wu, Y.Q.; Zhou, F.; Barnard, K. Attentional Local Contrast Networks for Infrared Small Target Detection. *IEEE Trans. Geosci. Remote Sens*. **2021**, *59*, 9813–9824. [CrossRef]

39. Kim, J.H.; Hwang, Y. GAN-Based Synthetic Data Augmentation for Infrared Small Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–12. [CrossRef]

40. Zuo, Z.; Tong, X.; Wei, J.; Su, S.; Wu, P.; Guo, R.; Sun, B. AFFPN: Attention Fusion Feature Pyramid Network for Small Infrared Target Detection. *Remote Sens.* **2022**, *14*, 3412. [CrossRef]

41. Guan, X.W.; Peng, Z.M.; Huang, S.Q.; Chen, Y.P. Gaussian Scale-Space Enhanced Local Contrast Measure for Small Infrared Target Detection. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 327–331. [CrossRef]

42. Han, J.H.; Liu, S.B.; Qin, G.; Zhao, Q.; Zhang, H.H.; Li, N.N. A Local Contrast Method Combined With Adaptive Background Estimation for Infrared Small Target Detection.*IEEE Geosci. Remote Sens. Lett.* **2019**,*16*, 1442–1446. [CrossRef]

43. Han, J.H.; Xu, Q.Y.; Saed, M.; Fang, H.Z.; Yuan, X.Y.; Qi, Z.M.; Wan, J.Y. A Ratio-Difference Local Feature Contrast Method for Infrared Small Target Detection. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 1–5. [CrossRef]

44. Du, P.; Askar, H. *Infrared Small Target Detection Based on Facet-Kernel Filtering Local Contrast Measure*; Springer: Singapore, 2019.

45. Qi, S.; Xu, G.; Mou, Z.; Huang, D.; Zheng, X. A fast-saliency method for real-time infrared small target detection. *Infrared Phys. Technol.* **2016**, *77*, 440–450. [CrossRef]

46. Yang, P.; Dong, L.L.; Xu, W.H. Infrared Small Maritime Target Detection Based on Integrated Target Saliency Measure. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 2369–2386. [CrossRef]

47. Hui, B.W.; Song, Z.Y.; Fan, H.Q.; Zhong, P.; Hu, W.D.; Zhang, X.F.; Lin, J.G.; Su, H.Y.; Jin, W.; Zhang, Y.J.; et al. A dataset for infrared image dim-small aircraft target detection and tracking under ground/air background. *China Sci. Data* **2020**, *5*, 291–302.