



OPEN ACCESS

EDITED BY

Rajib Biswas,
Tezpur University, India

REVIEWED BY

Yongling Ren,
University of Western Australia, Australia
Peng Zhenming,
University of Electronic Science and
Technology of China, China

*CORRESPONDENCE

Dongdong Pang,
✉ dongdong.pang@foxmail.com

RECEIVED 07 August 2024

ACCEPTED 22 October 2024

PUBLISHED 10 April 2025

CITATION

Xu D, Pang D, Gou S, Liu Z and Zhuo Z (2025)
A novel local contrast method based on
third-order central moments for infrared
small target detection.
Front. Phys. 12:1477253.
doi: 10.3389/fphy.2024.1477253

COPYRIGHT

© 2025 Xu, Pang, Gou, Liu and Zhuo. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with
these terms.

A novel local contrast method based on third-order central moments for infrared small target detection

Danyang Xu¹, Dongdong Pang^{2*}, Shimiao Gou³, Zhaoyu Liu³
and Zhihai Zhuo²

¹School of Information Engineering, Henan University of Science and Technology, Luoyang, China,

²School of Information and Communication Engineering, Beijing Information Science and
Technology University, Beijing, China, ³Henan Key Laboratory of General Aviation Technology, School
of Electronics and Information, Zhengzhou University of Aeronautics, Zhengzhou, China

As local contrast mechanisms are extensively utilized in infrared small target detection. However, the performance of existing local contrast-based methods is often compromised in complex backgrounds. This study presents a novel local contrast method based on third-order central moments to address the above challenges. Initially, the infrared image undergoes top-hat transformation to mitigate most background clutter and highlight potential target pixels. Then, a local contrast description operator based on third-order central moments is defined to characterize the grayscale changes in different regions of the preprocessed image, enhance the target and suppress the background. Finally, the target is extracted by using an adaptive threshold segmentation operation. The experimental results in six real-life scenarios demonstrate that the proposed method occupies the best detection index compared to other similar technologies.

KEYWORDS

small infrared (IR) target detection, local contrast, third-order central moments, image sequence, low-altitude moving background

1 Introduction

Target detection is a pivotal technology in infrared search and track (IRST) systems [1]. In practical applications such as guidance, early warning, and surveillance/monitoring, whether airborne, space-borne or in anti-UAV (unmanned aerial vehicle) operations, which have garnered significant attention recently, detecting targets presents substantial challenges. These challenges stem from the inherent characteristics of the targets, which typically exhibit weak brightness and small size, while lacking unique shape, texture, and color information. Consequently, rapid and robust detection of small infrared targets has always been an unresolved issue in the field of object detection.

Currently, the relevant infrared small target detection technologies are mainly classified into two types: sequential detection approaches and single-frame detection approaches. Sequential detection technology typically capitalizes on the continuous motion of the target contrasted against the random motion of noise in sequential imagery. This type of technology facilitates small target detection and random noise

elimination by identifying potential motion trajectories, albeit requiring substantial prior information. Conversely, the single-frame detection method relies on the characteristic feature information of dim and small targets within a single image to accomplish detection. This type of technology offers several advantages over sequential detection, including simpler computations, faster processing speeds, higher real-time performance, and broader applicability.

The current single-frame detection technology mainly includes the following two categories: non-local information-based methods and local information-based methods. Non-local information methods distinguish the target image from the entire original image by leveraging differences in frequency bands and data spaces between the target and background/noise images. These methods include frequency domain techniques, classifier methods, overcomplete sparse representation, and sparse low-rank decomposition. To illustrate, Gregoris et al. [2] integrated wavelet transform into infrared small target detection. Wu et al. [3] developed a method based on the contour transform that suppresses background frequencies while enhancing those of the target within the transform domain to bolster detection capabilities. Kong et al. [4] utilized diagonal detail information from Haar wavelet decomposition to aid in detecting weak infrared targets against sea-sky backgrounds. Wang et al. [5] applied dual-tree complex wavelets for decomposing the original image and employed a top-hat operator to filter the low-frequency sub-bands, subsequently using local image entropy to weight the reconstructed image for extracting infrared small targets. Zheng et al. [6] initially employed background estimation and differencing to pinpoint suspicious pixels, followed by clustering analysis to identify potential small targets. Dong et al. [7] extracted interest points from the image and introduced a novel R-mean clustering method to categorize these points into targets and backgrounds based on their associative patterns. Wu et al. [8] utilized support vector machines to identify the optimal hyperplane in kernel space that segregates targets from backgrounds, thus classifying pixels into these categories. Jiang et al. [9] merged the correlation between the observation matrix and the sparse matrix with the concept of gradient descent to devise an adaptive gradient descent method. Gao et al. [10] posited in their IPI model that the background in infrared images constitutes a low-rank component and the target a sparse component, transforming the detection of infrared small targets into a recovery problem for low-rank and sparse matrices, followed by decomposing image blocks to more effectively separate targets from backgrounds. Building on this framework, Zhang et al. [11] proposed a method that involves partial tensor nuclear norm sums and incorporates target edge information into the model to enhance its ability to suppress edge clutter.

The local information class method posits that the grey level of the background pixel point in an infrared image is typically similar to that of its local surrounding pixel points, whereas the target pixel point exhibits a more pronounced disparity in its grey level relative to its peripheral pixel points. By extracting the difference information between each pixel point in the image and its neighborhood reference pixel point, the target can be successfully filtered. For instance, Shao et al. [12] utilized a Laplacian of Gaussian (LoG) filter template characterized by a positive central coefficient and negative surrounding coefficients. During the convolution filtering of the original image with this template, the differences between

each pixel and its neighbors are accentuated. From the standpoint of local image segmentation, Yao et al. [13] developed a small target detection model based on facet kernel and random walker (FKRW), which effectively mitigates edge noise. Chen et al. [14] introduced the Local Contrast Measure (LCM) method to enhance small target detection performance, although it suffered from a pronounced “blocking effect.” Building on LCM, a series of improved measures were subsequently proposed. Such as, Han et al. [15] introduced an Improved LCM (ILCM) method, incorporating the average value of sub-blocks as a parameter to better suppress random point noise, yet it tended to smooth out small targets when they were diminutive. Qin et al. [16] proposed a novel LCM (NLCM), that averages only the largest number of pixels in each sub-block, thus better addressing the issue of small targets being smoothed while still retaining noise suppression capabilities. Han et al. [17] initially defined Refined LCM (RLCM) and made significant enhancements in the design of the filtering template [18], the selection of background references [19], and the introduction of a weighting function [20]. This concept has been widely recognized and adopted in the field. For example, Wei et al. [21] proposed a target enhancement method based on multi-scale patch contrast measurement (MPCM) for infrared small target detection, although this method struggled with retaining target shape and edge information and resulted in numerous false alarms. Fu et al. [22] combined an adaptive filter with a probabilistic Hough transform to enhance local contrast, effectively distinguishing targets from backgrounds and accelerating the target detection process, albeit with reduced accuracy in highly complex backgrounds. Notably, recent studies have shown a trend where many researchers are combining local contrast with other types of algorithms to achieve superior detection results. Such as, Cui et al. [23, 24] integrated local contrast with support vector machines. Deng et al. [25–27] applied information entropy to weight local contrast, thoroughly analyzing the shape and size of the local information entropy window. Chen et al. [28, 29] utilized the local signal-to-clutter ratio (SCR) to weight local contrast or combined the local contrast with frequency domain concepts. Du et al. [30] employed local smoothness to weight local contrast. Xiong et al. [31] initially calculated the local gradient of the original image and then assessed the local contrast of the gradient map. Han et al. [32] merged local contrast with TDLMS adaptive background estimation. Additionally, Dai et al. [33] integrated both local and non-local prior information to propose the RIPT model, which effectively suppresses interference factors and enhances the accuracy of target detection in specific scenarios. However, in complex environments, this method is prone to interference from background elements, resulting in lower detection precision. Pang et al. [34] proposed a low-rank and sparse decomposition method based on greedy bilateral decomposition for infrared dim and small target detection. This method can detect the target quickly and stably in complex scenes with a low signal-to-noise ratio, but the detection effect performs unsatisfactorily in the background with significant changes between different frames.

In general, local information class algorithms focus on a limited number of pixel points within a local area when calculating each pixel point. This results in a relatively small computational volume, which may have the potential for real-time processing when engineering optimisation techniques such as parallel acceleration and pipelined architecture are employed. The local contrast

algorithm is a relatively simple and straightforward approach that aligns well with the infrared image model. By employing a carefully designed contrast formula, it is possible to enhance the visibility of a target while simultaneously suppressing complex background noise. However, the effectiveness of this method hinges on the premise that the target must be the most prominent locally. In practice, this may not always be the case, particularly in scenarios with highly complex backgrounds. In a real scene, if the background is highly complex, the target may be in close proximity to an extremely bright background, which may overwhelm it. This makes it challenging to detect the target using local information, which in turn leads to a degradation in the detection performance of local information-based algorithms in complex real backgrounds. The non-local information class makes full use of all the information in the frame to detect weak targets, irrespective of the prominence of the target in the local region. Non-local information algorithms leverage all available information within the frame to detect weak targets, independent of the target's prominence within the local region. Even when a target becomes less discernible due to its proximity to a highlighted background, it can still be successfully separated, presenting a significant advantage over local information-based classification algorithms. However, many non-local information-based algorithms proposed to date exhibit certain shortcomings that warrant further research and improvement. For instance, frequency domain methods posit that the frequency bands occupied by the target and background are distinct. However, in scenarios with more complex backgrounds, these frequency bands often overlap, complicating accurate differentiation in the frequency domain. Furthermore, most classifier methods require a substantial number of training samples, which can be challenging to obtain in the field of infrared weak target detection, especially in the presence of non-cooperative targets. The efficacy of the hyper-complete sparse representation method relies on the accuracy of the ultra-complete dictionary. Nevertheless, constructing an ultra-complete dictionary that encompasses all potential scenarios is impractical in practice. The sparse low-rank decomposition method assumes that the target is sparse while the background is low-rank. However, in complex backgrounds, sparse information may also be present at the edges of the background and at noise locations, which is susceptible to generating false alarms.

Due to the presence of various types of noise and complex background interference, the aforementioned methods are likely to result in false positives and missed detections. In practical applications, it is necessary to adjust and optimize these methods based on specific scenarios and data characteristics. Therefore, to address the issue of detecting small infrared targets in complex backgrounds, this paper proposes a novel target detection method based on the contrast of local third-order central moments. Specifically, the innovative aspects of the proposed method are as follows:

1. This study introduces the third-order central moment to characterize the fluctuation properties of image gray levels in different regions for the first time.
2. Utilizing the gray level fluctuation properties across different image regions, a local contrast descriptor based on the third-order central moment is designed to enhance targets and suppress backgrounds.

3. Extensive experiments have been conducted, and multiple evaluation metrics have been employed to validate the effectiveness and superiority of the proposed method.

The other sections of the paper are organized as follows: [Section 2](#) details the method proposed in this paper; [Section 3](#) presents comparative experiments with six baseline methods in real infrared scenes, provides experimental results, and uses a series of evaluation metrics to verify the effectiveness of the proposed method; finally, [Section 4](#) concludes this article.

2 Materials and methods

[Figure 1](#) illustrates the detection workflow of the proposed method. Initially, the infrared image is preprocessed using a top-hat operator to suppress most of the background clutter and extract candidate target pixels. Subsequently, the contrast of the third-order central moments in the local areas of each candidate target pixel is calculated to enhance the targets and suppress clutter, resulting in a saliency detection map. Finally, the targets are accurately segmented and extracted using an adaptive threshold.

2.1 Preprocessing

Due to the weak target signal and low target intensity, the target is easily masked under complex background or strong clutter interference. In order to enhance the target detection ability under various backgrounds, here the top-hat operator is used to do preprocessing on the whole image to suppress the background noise and improve the image signal-to-noise ratio. For the original image and the structure element two basic operations, namely dilation and erosion, are defined. The dilation operation makes the gray value of the image larger than the gray value of the input image due to the maxima operation, while the erosion operation makes the gray value of the image smaller than the gray value of the input image due to the minima operation. Thus, dilation results in increasing the size of the bright areas and decreasing the size of the dark areas. Erosion results in the opposite. They are denoted by and as [Equations 1, 2](#):

$$(f \oplus b)(x, y) = \max_{m, n} (f(x - m, y - n) + b(m, n)) \quad (1)$$

$$(f \ominus b)(x, y) = \min_{m, n} (f(x + m, y + n) - b(m, n)) \quad (2)$$

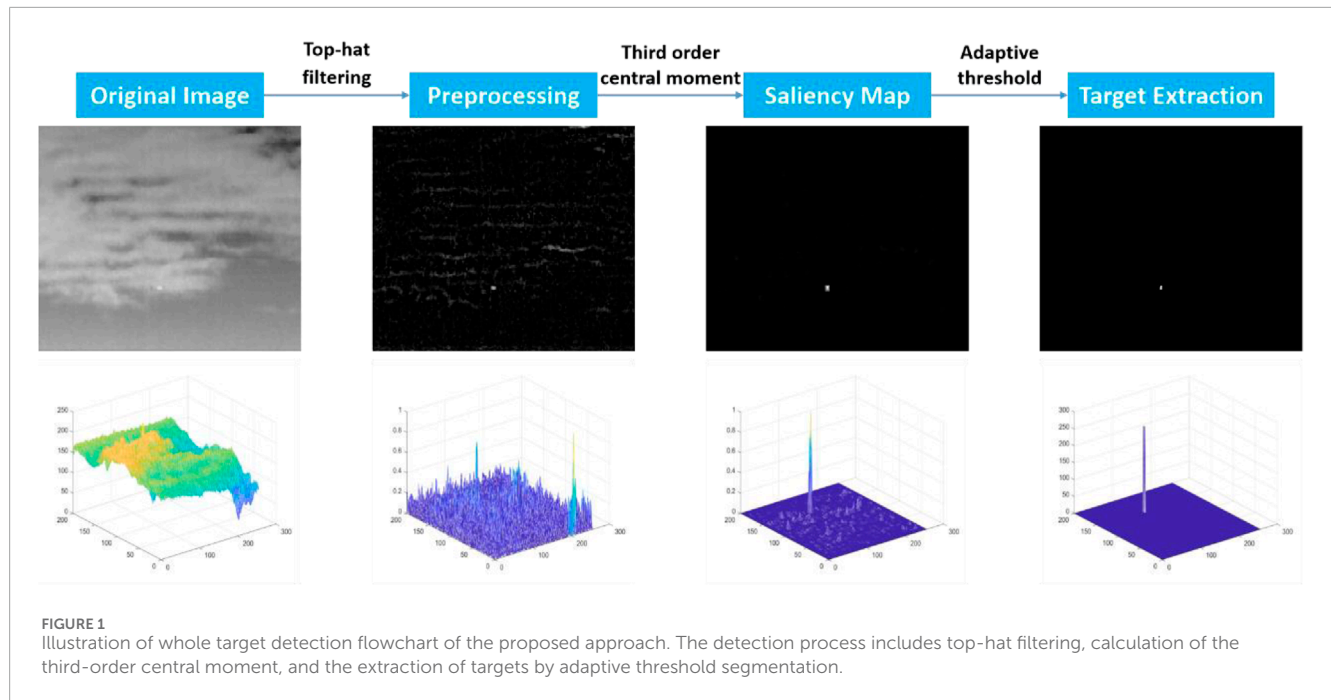
where x and y are the coordinates of the pixels in the image, m and n are the offsets of the coordinates of the pixels in the structure element with respect to x and y .

Then, the open operation $(f \circ b)$ and the closed operation $(f \cdot b)$ can be expressed as [Equations 3, 4](#):

$$(f \circ b)(x, y) = (f \ominus b) \oplus b \quad (3)$$

$$(f \cdot b)(x, y) = (f \oplus b) \ominus b \quad (4)$$

By using the open operation, the background can be obtained after extracting the fore-ground. Subsequently, the background is



subtracted from the original image to high-light the target. The above process is also described as top-hat transform and is defined as Equations 5, 6:

$$T_{hat}(f(x,y)) = f - f(f \circ b) \quad (5)$$

$$B_{hat}(f(x,y)) = (f \cdot b) - f \quad (6)$$

where $T_{hat}(f(x,y))$ and $B_{hat}(f(x,y))$ are referred to as the open top-hat and closed top-hat operations, respectively.

As illustrated in Figure 1, after preprocessing with the top-hat transformation operation, the SNR of infrared images under different backgrounds is improved, and the targets are significantly enhanced. Although some clutter remains, the background clutter in local areas tends to spread in a certain direction, exhibiting local directional consistency in grayscale values. The areas containing targets show drastic changes in grayscale values, with little correlation to the surrounding neighborhood.

2.2 Calculation of the third-order center moment

By employing a sliding window of size $M \times N$, the image is traversed from top to bottom and left to right after preprocessing. The third-order central moment can reflect the intensity of pixel value changes within a certain spatial range, as well as whether the pixel value changes conform to a Gaussian distribution [35]. For image blocks with small grayscale changes or conforming to a Gaussian distribution, the third-order central moment is approximately zero. For an image patch with a size of $M \times N$, its

third-order central moment, J_3 , can be defined as Equations 7, 8:

$$J_3 = \frac{\sum_{x=1}^M \sum_{y=1}^N (f(x,y) - \bar{f})^3}{M \times N} \quad (7)$$

$$\bar{f} = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N f(x,y) \quad (8)$$

where, $f(x,y)$ represents the grayscale value of the pixel at (x,y) and $\bar{f}$ denotes the average grayscale value of the pixel in the image patch.

The third-order central moment is calculated for each window area, and this value is assigned to the central pixel of the local area. This process generates a saliency map based on the third-order central moments, as shown in Figure 2. In this figure, the brighter pixels indicate higher grayscale values, corresponding to larger values of the third-order central moments [36]. It is evident from the figure that the target areas have significantly higher third-order central moment values compared to the surrounding neighborhood pixels, creating a stark contrast. The residual background clutter in the preprocessed image has lower third-order central moment values, resulting in low contrast with the surrounding areas. Therefore, it is considered to further suppress the residual clutter by applying the contrast of local third-order central moments in the image post top-hat preprocessing.

Initially, a nested sliding window structure is constructed, as illustrated in the enlarged area of the figure. This model includes a central block T and eight surrounding neighborhood blocks $B_i (i = 1, \dots, 8)$, all of equal size [37]. The third-order central moments for the central block T and the eight surrounding blocks are calculated separately. The local third-order central moment contrast value for the area is determined and assigned to the central pixel within the sliding window.

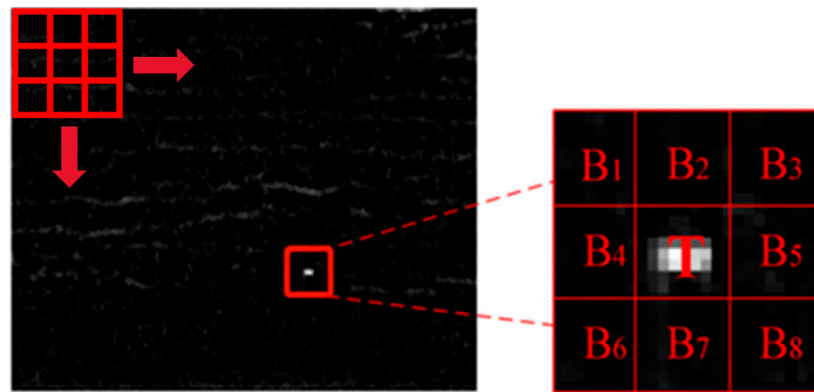


FIGURE 2
Third-order central moment saliency map and sliding window example.

The specific calculation process is as Equation 9:

$$C_i = \frac{J_T}{J_{B_i}} (i = 1, \dots, 8) \quad (9)$$

where J_T and J_{B_i} are the third-order central moments for the T area and B_i areas, respectively. The minimum value of C_i is selected as the contrast gain coefficient C for the local area, calculated as Equation 10:

$$C = \min C_i (i = 1, \dots, 8) \quad (10)$$

Then, the pixel value C_n in the third-order central moment contrast saliency map, defined as Equation 11:

$$C_n = m_T \times C \quad (11)$$

where m_T represents the mean value of the 3×3 region in the central of region T. The value range of the gain coefficient C is influenced by the position of the window within the image and the information contained within the image patches at various locations [38]. The behavior of the gain coefficient C can be described in the following scenarios:

1. When the window is positioned in a stable background area, the background and surrounding pixels display structural similarity. Consequently, both the central and nearest neighbor blocks exhibit high third-order central moment values. Nevertheless, the discrepancy between these values is minimal, resulting in a gain coefficient C that is approximately 1.
2. If the window is located in a target area and the central block encompasses complete target information, the difference in third-order central moment values between the central block and the neighboring blocks is substantial. This results in a gain coefficient C that is greater than 1, effectively enhancing the target.
3. In cases where the window includes partial target information and the central block represents a background area, the third-order central moment values between the central block and the neighboring background blocks are comparable. However, they significantly differ from those of the neighboring target

blocks, leading to a gain coefficient C that is less than or equal to 1. This characteristic effectively suppresses the background while enhancing the target.

The image regions where the central block is the target have large contrast gain coefficients C , while the contrast gain coefficients C of the gentle and undulating background regions are small. Although the gray value of the undulating background region changes greatly, it is strongly correlated with the gray value of the surrounding neighbourhood, and there is a similar structure in the local region, and the third-order central moments of the small blocks in the background region do not vary much, so the contrast gain coefficient of the value is much smaller compared to the target. The dynamic adjustment of the gain coefficient C based on local image characteristics allows for the selective enhancement of targets and the suppression of background noise, improving the overall detection performance. In summary, the local third-order central moment contrast-based detection algorithm has good target enhancement and background suppression capabilities, which is conducive to target detection.

2.3 Target extraction

After the above operation, the target and background are well enhanced and suppressed. Subsequently, the adaptive threshold segmentation operation is employed to extract targets, and the threshold calculation formula is defined as Equation 12:

$$Th = \mu + k\sigma \quad (12)$$

where μ and σ denote the mean and standard deviation of the processed image, respectively, and k is an adjustable parameter that allows for threshold adjustment in different scenarios, and experimental results show that a setting of 5 is quite appropriate in our work. The framework of the proposed method in this paper is summarized in Algorithm 1.

Input: Infrared image $f(x,y)$, size of local window v , size of large window u , contrast enhancement exponent C , structural element for top-hat transformation $b(m,n)$, and parameter k

Output: Detection result

```

1: Calculate the top-hat transformation map
 $f(x,y)$  using Equation 5;
2: for 1 to  $i$  do
3: Calculate the local mean  $\bar{f}$  using Equation 8;
4: Calculate the local contrast
 $J_3$  using Equation 7;
5: Calculate the absolute mean contrast;
6: end for
7: Calculate the contrast metrics of the
sub-blocks and normalize them;
8: Obtain the saliency map based on the
third-order central moment;
9: Obtain the final detection
result using Equation 12.
```

Algorithm 1. Specific Target Detection Steps of The Proposed Method.

3 Experiment and analysis

Here, we first introduce the datasets used in the experiments and then analyses the performance of the proposed method from both qualitative and quantitative perspectives. Qualitatively, the performance of the method is described through the detection result images and their three-dimensional distributions. Quantitatively, the analysis is conducted based on several metrics: Signal-to-Noise Ratio (SNR), Signal-to-Noise Ratio Gain (SNRG), Background Suppression Factor (BSF), Receiver Operating Characteristic (ROC) curves, Area Under the Curve (AUC), and the average runtime of the algorithm.

3.1 Datasets and baseline methods

To further validate the effectiveness and robustness of the proposed algorithm, six infrared image sequences and an infrared dataset containing mixed frames are select-ed as the experimental dataset. Among these, the background of scenes 1 and 6 contain strong cloud clutter, and the background of scene 2 includes high-brightness buildings. The single-frame mixed dataset contains high-brightness point-like noise similar to re-al targets, mountain-forest environments, high-brightness interference objects, and multi-target scenarios, significantly increasing the detection difficulty. Table 1 provides a detailed feature description of the dataset.

To illustrate the efficacy and robustness of the proposed approach, six established methods were chosen as baseline comparisons. The IPI [10] model was selected to represent non-local information-based techniques. Local information-based methods comprised MPCM [21], LoG [12], AAGD [9] and FKRW [13]. Furthermore, the RIPT [33] method, which combines local and non-local *a priori* information, was selected as a comparison method. The

parameters of different methods are adjusted to the best through experiments.

3.2 Evaluation indicators

The evaluation of infrared small target detection algorithms can be conducted from both qualitative and quantitative perspectives. Qualitative analysis involves subjective assessment based on the detection result images and their corresponding three-dimensional distributions, such as whether targets are detected, the number of false alarms, and the degree of background clutter suppression compared to the original image. Due to the influence of human subjective factors, it is essential to perform a quantitative analysis to objectively evaluate the experimental results. Common evaluation metrics include Signal-to-Noise Ratio (SNR), Signal-to-Noise Ratio Gain (SNRG), Background Suppression Factor (BSF), Probability of Detection (PD), False Alarm Rate (FA), Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC). SNR is expressed as Equation 13:

$$SNR = \frac{I_{max} - I_{mean}}{\sigma} \quad (13)$$

where, I_{max} and I_{mean} represent the maximum grey value and the average grey value of the image respectively, and σ signifies the standard deviation of the image. The SNRG is defined as Equation 14:

$$SNRG = 20 \times \log_{10} \frac{SNR_{out}}{SNR_{in}} \quad (14)$$

where, SNR_{out} and SNR_{in} denote the SNR of the processed image and the original image, respectively. In general, the larger SNRG is, the better the target enhancement performance of the method.

BSF can be used to describe the background suppression ability of the corresponding method as Equation 15:

$$BSF = \frac{\sigma_{in}}{\sigma_{out}} \quad (15)$$

where, σ_{out} and σ_{in} indicate the standard deviation of the processed image and the standard deviation of the original image, respectively. Normally, the higher the BSF value, the better the method suppresses the infrared background.

The ROC curve is a critical metric for assessing target detection performance, composed of PD and FA rates, defined as Equation 16:

$$PD = \frac{TD}{AT}, FA = \frac{FD}{NP} \quad (16)$$

where, TD represents the number of correctly detected targets, AT denotes the actual number of true targets, FD and NP respectively represent the number of pixels in false alarm regions and the total number of pixels in the test data.

When there is an overlap between the detected target pixels and the real target pixels and the central distance between them is less than 5 pixels, the detected target is considered to be a real target, and *vice versa*, it is considered to be a false target.

TABLE 1 Presentation of specific details for sequences 1–6.

	Frames number	Image size	Specific characteristics of target and background
Seq1	30	256 × 200	Gloomy sky background, strong edge cloud clutter The target intensity is weak and overlaps with the cloud
Seq2	40	300 × 300	Gloomy sky background, high brightness tree and building clutter Extremely weak target intensity and low contrast
Seq3	400	330 × 230	Dim sky background, strong edge cloud clutter Extremely weak target intensity and low contrast
Seq4	400	198 × 200	Ground - sky background, railing Extremely weak target intensity and low contrast
Seq5	100	506 × 404	Sky background, buildings Extremely weak target intensity and low contrast
Seq6	200	254 × 200	Sky background, strong edge cloud clutter Target and cloud overlap

Additionally, the AUC can be calculated as a supplementary quantitative evaluation metric for the ROC curve. Generally, a larger AUC signifies better detection performance represented by the ROC curve. The area under the ROC curve, AUC, is calculated using a non-parametric method. The area under the curve is calculated as Equation 17:

$$AUC = \frac{1}{2} \sum_{i=1}^n (x_i - x_{i-1})(y_i + y_{i-1})$$

(17)

where x_i and y_i represent the false alarm rate and the detection rate, and n means the total number of operating points on the ROC curve.

3.3 Analysis and discussion of experimental results

This section employs a series of evaluation metrics to assess the detection performance of different methods, thereby validating the effectiveness and robustness of the proposed approach. The experiments utilized six sequence datasets and one single-frame mixed dataset. The proposed method was compared with baseline methods to verify its effectiveness and robustness.

3.3.1 Visual comparison

The detection results are presented in Figures 3, 4. Figure 3 displays the out-comes for the sequence datasets, while Figure 4 shows the results for the single-frame mixed dataset. In Sequence 1, despite the significant cloud clutter in the background, the targets are distinct and brightly illuminated. The baseline methods succeed in detecting the targets, but the heightened background clutter leads to an elevated false alarm rate. In Sequence 2, where the targets are less bright and the background interference is more pronounced, the performance of the LoG and AAGD methods exhibit a marked decline in detection accuracy, failing to identify the actual targets. The remaining baseline methods manage to

detect the actual targets but also flag numerous suspicious targets. For Seq 3 and 6, which feature more uniform scenes, all seven methods yield favorable detection results. However, Seq 4 and 5 present challenges due to the high-brightness interference clutter in the backgrounds, leading to suboptimal performance from the baseline methods. The MPCM method detects real targets in both sequences but also registers numerous background noises and suspicious targets. The AAGD, FKRW, IPI, and RIPT methods fail to detect real targets in Sequence 4 and retain false targets. In Sequence 5, although AAGD, IPI, and RIPT detect real targets, they also identify suspicious targets, and FKRW fails to detect real targets.

As observed in Figure 4, the proposed method performs excellently in various scenes, demonstrating strong target response capabilities and background suppression abilities, and is capable of detecting multiple targets.

3.3.2 Quantitative comparison

SNRG and BSF are used to evaluate the target enhancement and background suppression capabilities of the proposed method compared to baseline methods. Higher values of SNRG and BSF indicate superior performance of the respective methods. Tables 2, 3 display the SNRG and BSF values for different methods under six different backgrounds, where bold numbers represent the highest values of SNRG and BSF in each sequence, and underlined numbers indicate the second highest values. The proposed method achieved the highest SNRG and BSF in Sequences 2, 4, and 5. Notably, high SNRG and BSF values are primarily found in the RIPT, FKRW, and the proposed methods. Among these, the RIPT model showed the highest SNRG and BSF values in Scenes 1 and 3, although background clutter still existed in the detection results of Sequence 1.

As shown in Figure 5, the ROC curves indicate the relationship between Pd and Pf in Seq 1–6 for all methods. Also, in order to visually evaluate the detection performance of different algorithms, we calculated the area under the curve (AUC) values of different methods, as shown in Table 4. Bold and underline indicate the

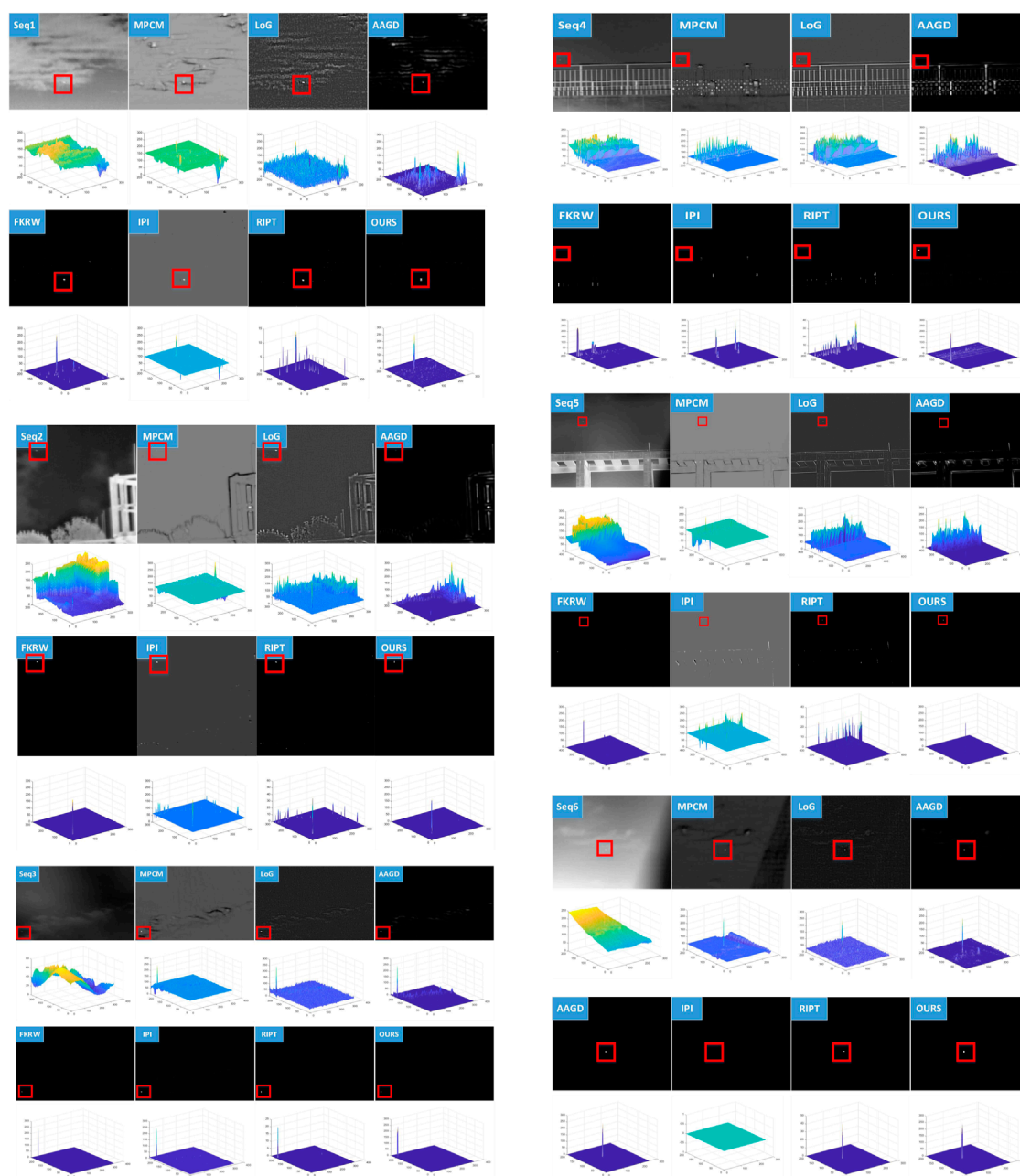


FIGURE 3
Visual comparison results of different approaches on sequences 1–6. For better resolution, the red rectangle are used to mark the target, respectively.

maximum and second largest values, respectively. In general, the closer the ROC curve is to the upper left corner and the larger the AUC value is, the better the detection performance is. For seq 1, the target detection result of LoG and MPCM methods show the presence of obvious interfering targets, which fail to be detected accurately. Therefore, a lower Pd value corresponds to it when the pf is the same, causing a low AUC value. In seq 2 and 5, MPCM obtained the closest detection accuracy to the proposed method. The AUC values of the corresponding ROC

curves were also second only to our method. In seq 3 and 4, the AUC values of LoG and RIPT were low. In sequence 6, AAGD accurately detects the target, closest to the detection accuracy of the proposed method. The AUC values of the corresponding ROC curves are also second only to our method. The ROC curves of the proposed method in this paper consistently maintain the optimal Pd value and the maximum AUC value in all six scenarios, indicating that our method has excellent target detection capability in all six scenarios.

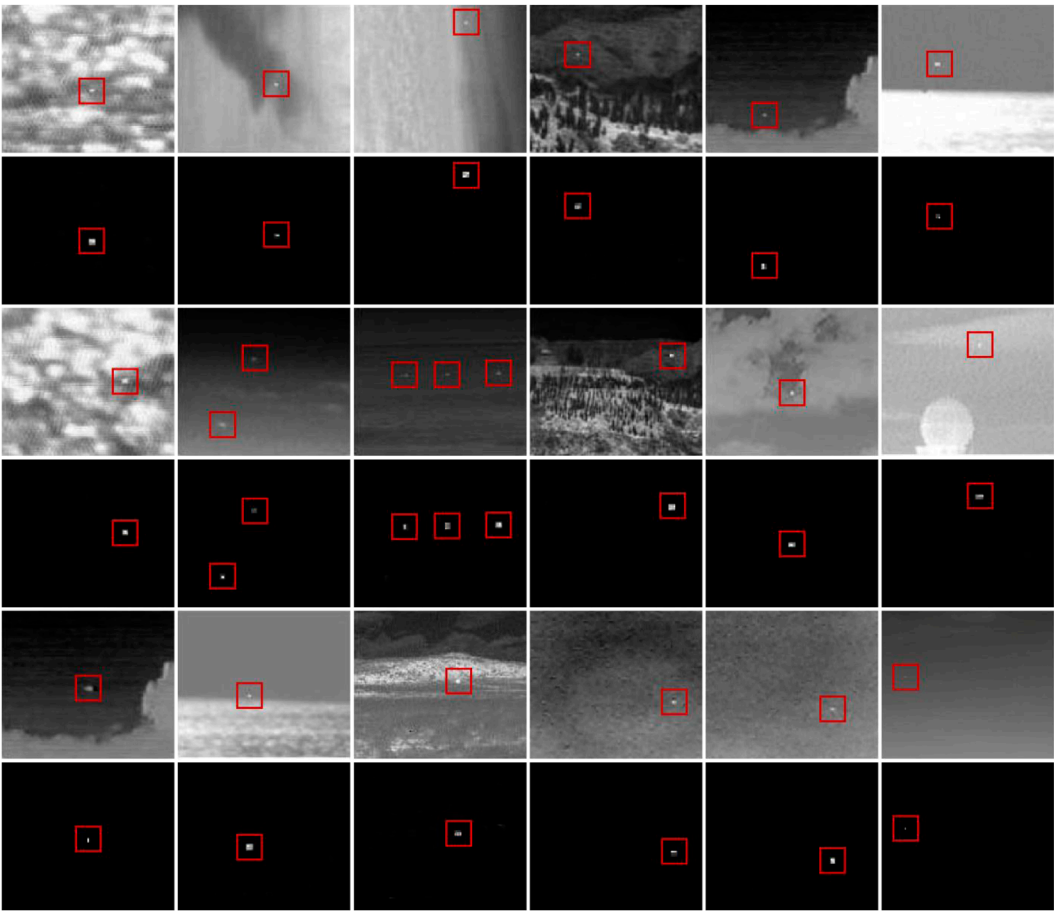


FIGURE 4
Visual comparison of single-frame mixed datasets, the red rectangle are used to mark the target, respectively.

TABLE 2 Average SNRG values of different methods in six real scenes. The bold value is the maximum value.

	LoG [12]	AAGD [9]	MPCM [21]	FKRW [13]	IPI [10]	RIPT [33]	Proposed
Seq 1	42.9126	37.2999	18.1030	46.5820	45.7398	47.1608	46.5872
Seq 2	12.6842	22.7342	−3.3249	33.7752	25.8948	31.9745	35.1651
Seq 3	38.6286	39.0160	25.0280	40.0749	39.6803	41.1559	38.9639
Seq 4	3.1224	12.4164	11.8490	24.5923	25.5748	22.5606	26.3360
Seq 5	15.2126	20.7905	19.9816	37.1702	16.5926	29.0504	41.4241
Seq 6	42.5041	43.1355	44.4799	42.7732	NaN	42.6279	41.4010

4 Discussion

4.1 Discussion of detection performance

The detection of infrared small targets is rendered challenging by the complexity and variability of the infrared image environment, which frequently includes substantial background clutter and interference noise. Conventional detection techniques,

such as LoG, MPCM, and AAGD, are characterized by high computational complexity and noise sensitivity, rendering them prone to image noise and resulting in false or missed detections. Methods like IPI and FKRW are highly responsive to variations in target gray intensity, leading to unstable detection outcomes when the target and background exhibit minimal grayscale differences or when influenced by factors such as illumination. The RIPT detection method, which integrates

TABLE 3 Average BSF values of different methods in six real scenes. The bold value is the maximum value.

	LoG [12]	AAGD [9]	MPCM [21]	FKRW [13]	IPI [10]	RIPT [33]	Proposed
Seq 1	17.4953	9.1209	5.9284	26.2413	23.2279	27.5569	26.3840
Seq 2	3.7184	11.4748	2.4414	40.9638	17.4044	33.9158	48.8688
Seq 3	49.0199	52.9625	27.3042	59.2640	55.6245	66.0983	51.7224
Seq 4	1.3675	3.8351	3.5887	16.6863	17.6726	12.2383	19.9535
Seq 5	4.5858	8.6565	2.1983	58.7325	5.3834	22.4054	92.9656
Seq 6	42.1676	45.3576	53.0407	43.6129	Inf	42.7872	37.6230

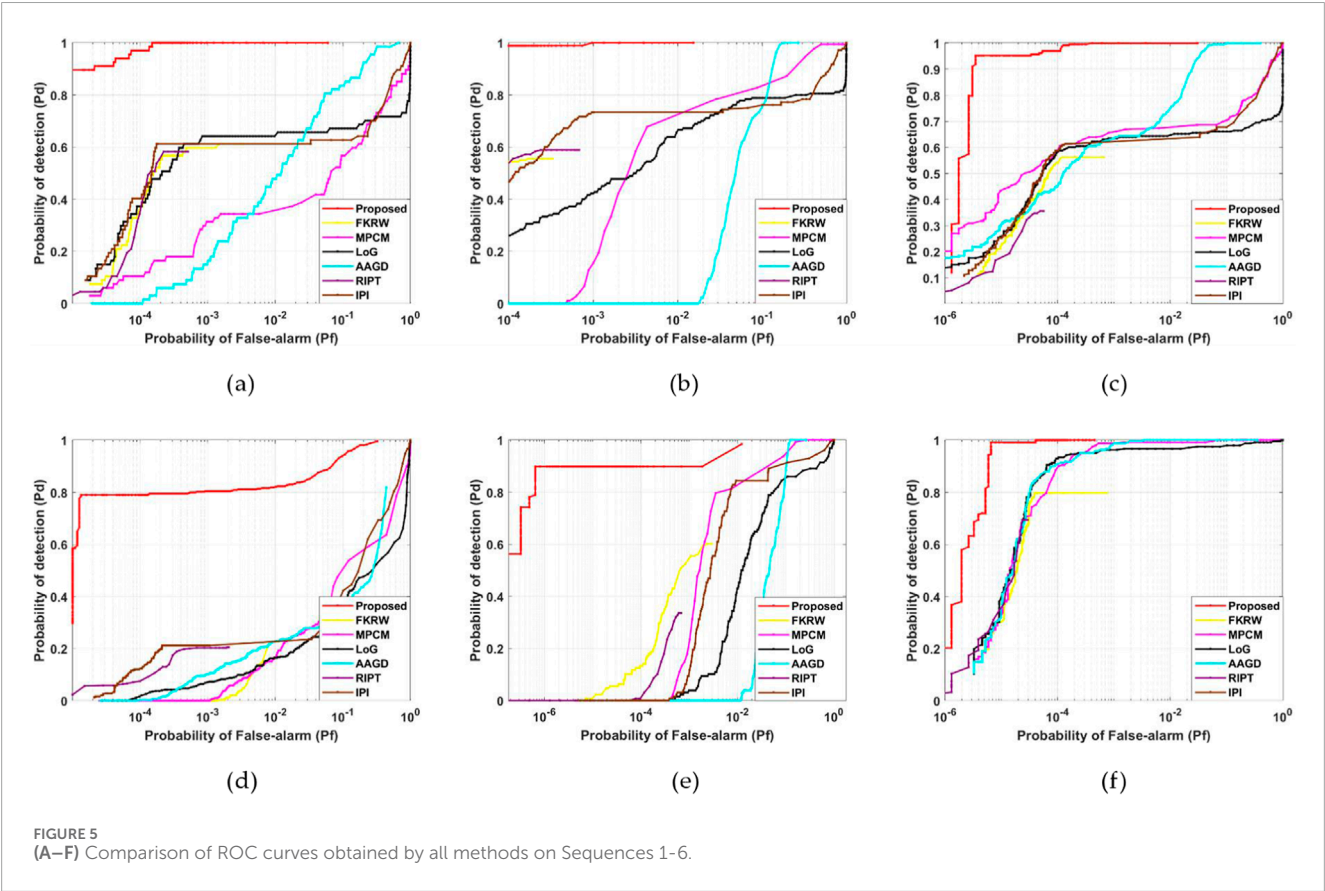


TABLE 4 AUC values corresponding to ROC curves of all approaches on sequences 1.6 ($\times 10^{-3}$). The bold value is the maximum value.

	MPCM [21]	LoG [12]	AAGD [9]	FKRW [13]	IPI [10]	RIPT [33]	Proposed
Seq 1	758.7167	717.5151	947.0283	805.2301	806.2288	790.8851	999.9911
Seq 2	944.0244	801.0041	935.2437	777.6925	859.6449	794.2826	999.9886
Seq 3	857.5838	702.3206	993.0174	780.9553	852.5991	678.0886	999.9926
Seq 4	693.3668	581.772	718.3679	607.5147	726.6379	601.7396	981.3734
Seq 5	984.0468	895.36	945.2263	799.9512	946.2618	667.6413	991.48
Seq 6	999.4926	990.079	999.9209	898.6716	500	847.0958	999.9969

TABLE 5 Parameter settings for all approaches.

No.	Abbreviations	Parameter settings
1	MPCM [21]	Local window size: $N = 3, 5, 7, 9$. Mean filter size: 3×3
2	LoG [12]	$\sigma = 1.5$, scale size: $n = 5$
3	AAGD [9]	$l_{max} = 19, 19, 19, 19$; $l_{min} = 3, 5, 7, 9$
4	FKRW [13]	Local window size: 11×11
5	IPI [10]	Patch size: 50×50 , sliding step: 10, $\lambda = 1/\sqrt{\min(m, n)}$, $\varepsilon = 10^{-7}$
6	RIPT [33]	Patch size: 50×50 , sliding step: 10, $\lambda = \frac{L}{\sqrt{\min(I, P)}}$, $L = 1$, $h = 10$, $\varepsilon = 0.01$, $\varepsilon = 10^{-7}$
7	Proposed	Local window size: $v = 3, 5, 7, 9$ $u = 3 \times v$

TABLE 6 Presentation of specific details for sequences 1–4.

	Frames number	Image size	Specific characteristics of target and background
Seq1	30	256×200	Gloomy sky background, strong edge cloud clutter The target intensity is weak and overlaps with the cloud
Seq2	30	300×300	Gloomy sky background, high brightness building clutter Extremely weak target intensity and low contrast
Seq3	30	330×230	Dim sky background, strong edge cloud clutter Extremely weak target intensity and low contrast
Seq4	30	198×200	Ground - sky background, railing Extremely weak target intensity and low contrast

local and non-local in-formation, is effective in suppressing interference factors and enhancing target detection accuracy in specific scenarios, nonetheless, in complex scenes, it is more vulnerable to background interference, thereby decreasing detection accuracy.

This paper presents a novel methodology that commences with the preprocessing of infrared images using the top-hat operator. It proceeds to enhance target objects and suppress background clutter by calculating the third-order central moment contrast within the local region of each candidate target pixel. The extraction of targets is subsequently accomplished through adaptive threshold segmentation. The proposed method is evaluated against baseline techniques on six real sequence datasets and one single-frame hybrid dataset, demonstrating consistently superior performance across all datasets. Furthermore, the proposed method is marked by low computational complexity and is amenable to acceleration through GPU or Field-Programmable Gate Array (FPGA) technologies.

4.2 Discussion of the key parameter V

We briefly discuss in this section the selection of the key parameter v in the proposed method, and the value of v in Table 5 will directly determine the quality of the generated

target salient maps. To obtain the optimal value of v , we select different sequences of infrared images for simulation experiments. In the experiments, we set v as 3, 5, 7, and 9 respectively. The details of different sequences of infrared images are described in Table 6. In addition, we use ROC curves to evaluate the detection performance of the proposed method at different v values. Figure 6 shows the ROC curves for different scenes.

As shown in Figure 6, we find that the setting of v value significantly affects the final target detection results. v value is set to 3, the detection results are the worst in all the scenes. v value is set to 7, the optimal detection results are obtained in all the scenes. Therefore, it is recommended to set the v value to 7.

4.3 Robustness to noisy scene

Furthermore, the algorithm must also overcome the challenge of detecting objects in a noisy scene. In order to verify the robustness of the proposed method in noisy environments, we add Gaussian white noise with different standard variances to the original infrared image. As shown in Figure 7. Gaussian white noise with standard deviation of 5, 10, and 15 was added to the original infrared image in the first, third, and fifth rows, respectively. From the detection results in the figure, it can be

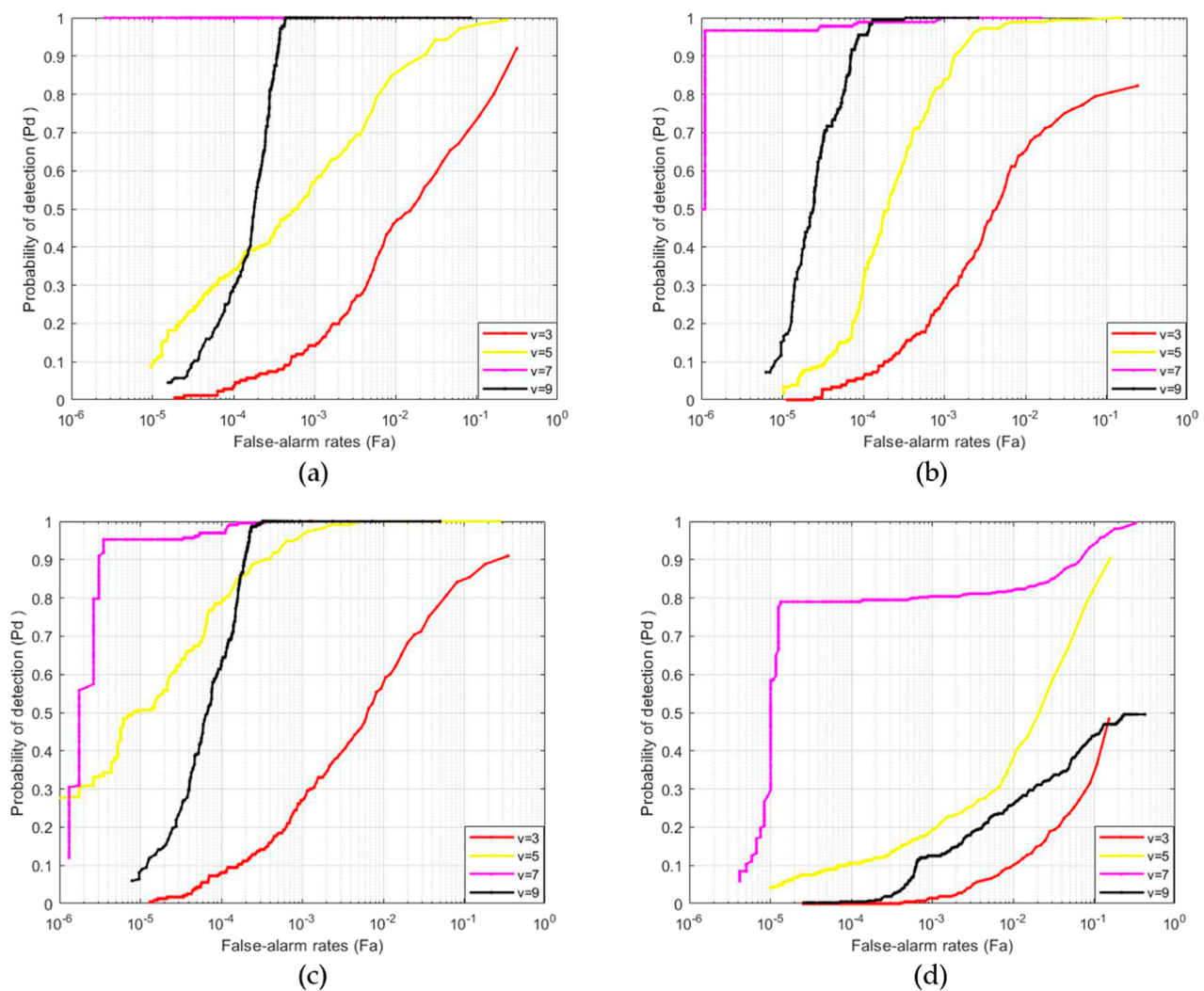


FIGURE 6
(A–D) Comparison of ROC curves obtained by different v value on Sequences 1–4.

seen that the proposed method can effectively detect the target in different degrees of noise scenarios. Thus, the ability of the proposed method in combating noise scenarios is verified. It shows that the proposed method possesses strong robustness in noisy scenes.

5 Conclusion

In this paper, a novel target detection method based on the contrast of local third-order central moments is proposed. Initially, the top-hat operator is used to process infrared images to suppress most of the background clutter and extract candidate target pixels. Subsequently, the contrast of the third-order central moments in the local areas of each candidate target pixel is calculated to enhance targets and suppress clutter. Finally, we conducted extensive experiments and compared six SOTA approaches on a dataset containing six real IR

videos with various scenes and a mixed dataset consisting of 18 single-frame infrared images with different backgrounds. Experimental results demonstrate that the proposed method can efficiently detect small infrared targets and shows significant advantages across a range of evaluation metrics. However, given the complexity and variability of real-world scenes, further in-depth research is warranted. This future work will primarily focus on the following aspects: (1) The small target detection method presented in this study is currently limited to single-frame infrared images and does not leverage the temporal information inherent in image sequences. Future investigations will integrate time-domain information to enhance target detection, considering the correlations across multiple frames. (2) The proposed method has been simulated in MATLAB, and we intend to explore its implementation on FPGA and other hardware platforms, in conjunction with existing laboratory equipment, to develop a real-time infrared small target detection system.

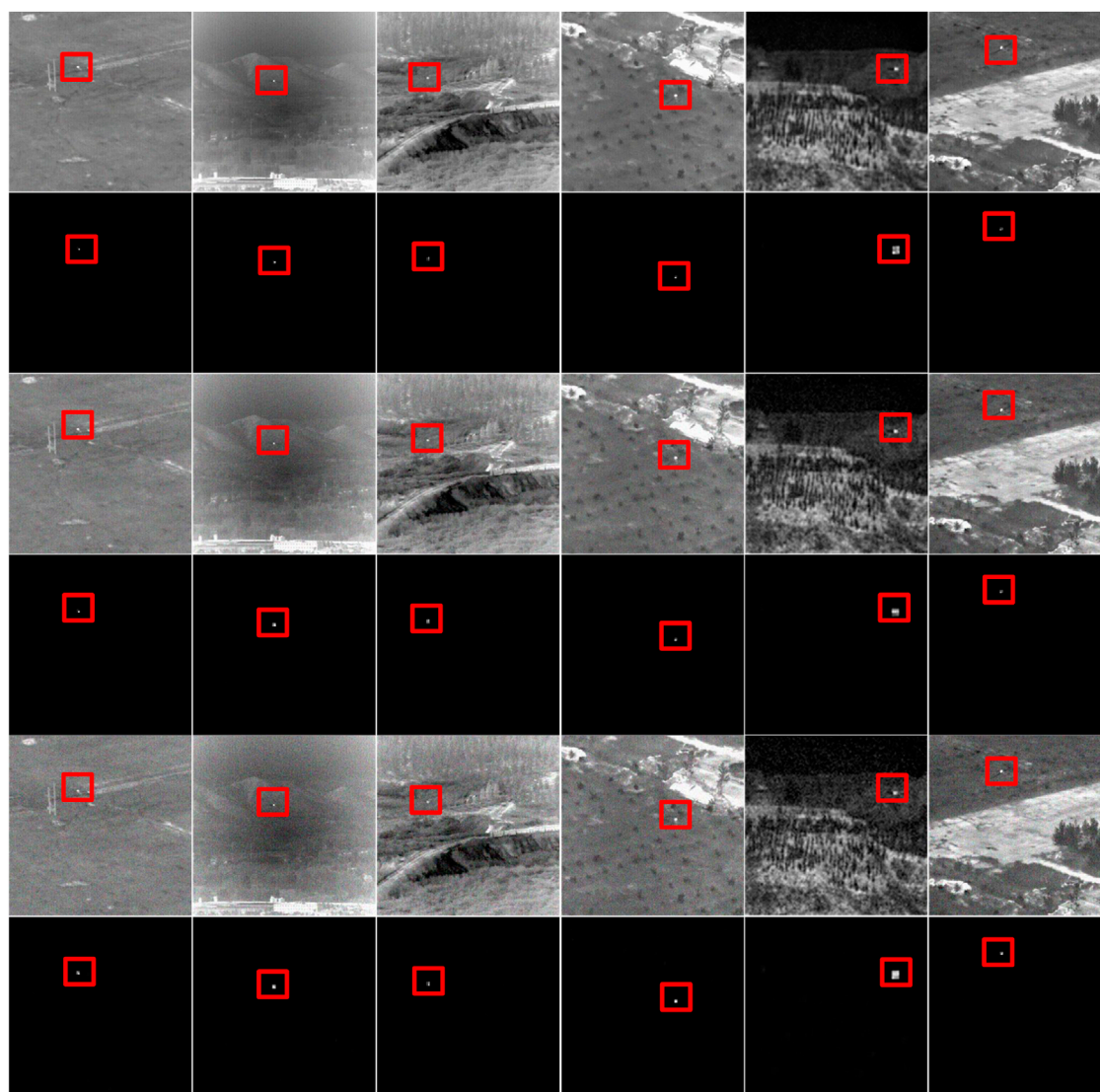


FIGURE 7
Detection results obtained in a noisy scene using the proposed method. The standard deviation of Gaussian white noise in the first, second and third rows are 5, 10, and 15, respectively.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

DX: Conceptualization, Data curation, Formal Analysis, Methodology, Writing—original draft, Writing—review and editing. **DP:** Conceptualization, Validation, Writing—original draft, Writing—review and editing. **SG:** Data curation, Writing—original draft, Writing—review and editing. **ZL:** Conceptualization, Writing—review and editing. **ZZ:** Conceptualization, Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was funded by National Natural Science Foundation of China [grant number 62301036]; The Open Fund Projects of Henan Key Laboratory of General Aviation Technology [grant number ZHKF-240201].

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Zhou F, Wu Y, Dai Y, Wang P, Ni K. Graph-regularized laplace approximation for detecting small infrared target against complex backgrounds. *IEEE Access* (2019) 7:85354–71. doi:10.1109/access.2019.2925563
- Gregoris DJ, Simon KW, Tritchew S, Sevigny L. Wavelet transform-based filtering for the en-hancement of dim targets in FLIR images. *Wavelet Appl* (1994) 2242:573–83. doi:10.1117/12.170058
- Wu YQ, Luo ZJ, Wu WY. A method of small target detection in infrared image based on nonsubsampling contourlet transform. *J Image Graphics* (2009) 14(3):477–81.
- Kong X, Liu L, Qian Y, Cui M. Automatic detection of sea-sky horizon line and small targets in maritime infrared imagery. *Infrared Phys and Technology* (2016) 76:185–99. doi:10.1016/j.infrared.2016.01.016
- He W, Yunhong X. Infrared small target detection algorithm based on double-tree complex wavelet transform. *Laser and Infrared* (2020) 50(9):1145–52. doi:10.1109/aimsec.2011.6010000
- Sheng Z, Jian L, Wen T. Research of SVM-based infrared small object segmentation and clustering method. *Signal Process.* (2005) 21(5):515–9.
- Dong X, Huang X, Zheng Y, Bai S, Xu W. A novel infrared small moving target detection method based on tracking interest points under complicated background. *Infrared Phys and Technology* (2014) 65:36–42. doi:10.1016/j.infrared.2014.03.007
- Wu Y, Zhou Y, Long Y. Small target detection in hyperspectral remote sensing image based on adaptive parameter SVM. *光学学报* (2015) 35:0928001. doi:10.3788/aos201535.0928001
- Yilin J, Qi T, Rongbing Z, Haiyan W, Qingbo J. Adaptive gradient descent optimization algorithm for measurement matrix. *Application research of computers* (2017).
- Gao C, Meng D, Yang Y, Wang Y, Zhou X, Hauptmann AG. Infrared patch-image model for small target de-tECTION in a single image. *IEEE Trans Image Process* (2013) 22(12):4996–5009. doi:10.1109/tip.2013.2281420
- Zhang L, Peng Z. Infrared small target detection based on partial sum of the tensor nuclear norm. *Remote Sensing* (2019) 11(4):382. doi:10.3390/rs11040382
- Shao X, Fan H, Lu G, Xu J. An improved infrared dim and small target detection algorithm based on the contrast mechanism of human visual system. *Infrared Phys and Technology* (2012) 55(5):403–8. doi:10.1016/j.infrared.2012.06.001
- Qin Y, Bruzzone L, Gao C, Li B. Infrared small target detection based on facet kernel and random walker. *IEEE Trans Geosci Remote Sensing* (2019) 57(9):7104–18. doi:10.1109/tgrs.2019.2911513
- Chen CP, Li H, Wei Y, Xia T, Tang YY. A local contrast method for small infrared target detection. *IEEE Trans Geosci remote sensing* (2013) 52(1):574–81. doi:10.1109/tgrs.2013.2242477
- Han J, 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 Sensing Lett* (2014) 11(12):2168–72. doi:10.1109/LGRS.2014.2323236
- Qin Y, Li B. Effective infrared small target detection utilizing a novel local contrast method. *IEEE Geosci Remote Sensing Lett* (2016) 13(12):1890–4. doi:10.1109/lgrs.2016.2616416
- 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 Sensing Lett* (2018) 15(4):612–6. doi:10.1109/lgrs.2018.2790909
- 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 Sensing Lett* (2019) 17(10):1822–6. doi:10.1109/lgrs.2019.2954578
- Han J, Liu C, Liu Y, Luo Z, Zhang X, Niu Q. Infrared small target detection utilizing the enhanced clos-est-mean background estimation. *IEEE J Selected Top Appl Earth Observations Remote Sensing* (2020) 14:645–62. doi:10.1109/jstars.2020.3038442
- Han J, Moradi S, Faramarzi I, Zhang H, Zhao Q, Zhang X, et al. Infrared small target detection based on the weighted strengthened local contrast measure. *IEEE Geosci Remote Sensing Lett* (2020) 18(9):1670–4. doi:10.1109/lgrs.2020.3004978
- Wei Y, You X, Li H. Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recognition* (2016) 58:216–26. doi:10.1016/j.patcog.2016.04.002
- Fu J, Zhao J, Li F. Infrared sea-sky line detection utilizing self-adaptive Laplacian of Gaussian filter and vis-ual-saliency-based probabilistic Hough transform. *IEEE Geosci Remote Sensing Lett* (2021) 19:1–5. doi:10.1109/lgrs.2021.3111099
- Cui Z, Yang J, Jiang S, Li J. An infrared small target detection algorithm based on high-speed local contrast method. *Infrared Phys and Technol* (2016) 76:474–81. doi:10.1016/j.infrared.2016.03.023
- Cui Z, Yang J, Li J, Jiang S. An infrared small target detection framework based on local contrast method. *Measurement* (2016) 91:405–13. doi:10.1016/j.measurement.2016.05.071
- Deng H, Sun X, Liu M, Ye C, Zhou X. Small infrared target detection based on weighted local difference measure. *IEEE Trans Geosci Remote Sensing* (2016) 54(7):4204–14. doi:10.1109/tgrs.2016.2538295
- Deng H, Sun X, Liu M, Ye C, Zhou X. Infrared small-target detection using multiscale gray difference weighted image entropy. *IEEE Trans Aerospace Electron Syst* (2016) 52(1):60–72. doi:10.1109/taes.2015.140878
- Deng H, Sun X, Liu M, Ye C, Zhou X. Entropy-based window selection for detecting dim and small infrared targets. *Pattern Recognition* (2017) 61:66–77. doi:10.1016/j.patcog.2016.07.036
- Chen Y, Song B, Wang D, Guo L. An effective infrared small target detection method based on the human visual attention. *Infrared Phys and Technology* (2018) 95:128–35. doi:10.1016/j.infrared.2018.10.033
- Chen Y, Song B, Du X, Guizani M. Infrared small target detection through multiple feature analysis based on visual saliency. *IEEE Access* (2019) 7:38996–9004. doi:10.1109/access.2019.2906076
- Du P, Hamdulla A. Infrared small target detection using homogeneity-weighted local contrast measure. *IEEE Geosci remote sensing Lett* (2019) 17(3):514–8. doi:10.1109/lgrs.2019.2922347
- Xiong B, Huang X, Wang M. Local gradient field feature contrast measure for infrared small target detection. *IEEE Geosci Remote Sensing Lett* (2020) 18(3):553–7. doi:10.1109/LGRS.2020.2976208
- Han J, Liu S, Qin G, Zhao Q, Zhang H, Li N. A local contrast method combined with adaptive background estimation for infrared small target detection. *IEEE Geosci Remote Sensing Lett* (2019) 16(9):1442–6. doi:10.1109/lgrs.2019.2898893
- Dai Y, Wu Y. Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection. *IEEE J selected Top Appl earth observations remote sensing* (2017) 10(8):3752–67. doi:10.1109/jstars.2017.2700023
- Pang D, Shan T, Li W, Ma P, Liu S, Tao R. Infrared dim and small target detection based on greedy bilateral factorization in image sequences. *IEEE J Selected Top Appl Earth Observations Remote Sensing* (2020) 13:3394–408. doi:10.1109/jstars.2020.2998822
- Pang D, Ma P, Feng Y, Shan T, Tao R, Jin Q. Tensor spectral k-support norm minimization for detecting infrared dim and small target against urban backgrounds. *IEEE Trans Geosci Remote Sensing* (2023) 61:1–13. doi:10.1109/tgrs.2023.3277848
- Pang D, Shan T, Ma P, Li W, Liu S, Tao R. A novel spatiotemporal saliency method for low-altitude slow small infrared target detection. *IEEE Geosci remote sensing Lett* (2021) 19:1–5. doi:10.1109/lgrs.2020.3048199
- Pang D, Ma P, Shan T, Li W, Tao R, Ma Y, et al. STTM-SFR: spatial-temporal tensor modeling with saliency filter regularization for infrared small target detection. *IEEE Trans Geosci Remote Sensing* (2022) 60:1–18. doi:10.1109/tgrs.2022.3172745
- Pang D, Shan T, Li W, Ma P, Tao R, Ma Y. Facet derivative-based multidirectional edge awareness and spatial-temporal tensor model for infrared small target detection. *IEEE Trans Geosci Remote Sensing* (2022) 60:1–15. doi:10.1109/tgrs.2021.3098969