

Research on infrared dim and small target detection algorithm based on local contrast and gradient

Weihong Lin, Leihong Zhang, Zimin Shen, Dawei Zhang, Jian Chen, Jie Zhou, Wei Peng & Fengshou Wu

To cite this article: Weihong Lin, Leihong Zhang, Zimin Shen, Dawei Zhang, Jian Chen, Jie Zhou, Wei Peng & Fengshou Wu (2022): Research on infrared dim and small target detection algorithm based on local contrast and gradient, Journal of Spatial Science, DOI: [10.1080/14498596.2022.2140714](https://doi.org/10.1080/14498596.2022.2140714)

To link to this article: <https://doi.org/10.1080/14498596.2022.2140714>

 View supplementary material [↗](#)

 Published online: 08 Nov 2022.

 Submit your article to this journal [↗](#)

 Article views: 53

 View related articles [↗](#)

 View Crossmark data [↗](#)



Research on infrared dim and small target detection algorithm based on local contrast and gradient

Weihong Lin ^a, Leihong Zhang^a, Zimin Shen ^a, Dawei Zhang^a, Jian Chen^b, Jie Zhou^c, Wei Peng^d and Fengshou Wu^d

^aCollege of Communication and Art design, University of Shanghai for Science and Technology, Shanghai, China; ^bChangchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, China; ^cLibrary, Shanghai University of Engineering Science, Shanghai, China; ^dCSSC Survey, Design and Research Institute Co. Ltd, Shanghai, China

ABSTRACT

The rapid development of infrared technology makes it widely applicable in such fields as military, medical, testing and communications. This paper proposes an algorithm based on local contrast and gradient (LCG). Specifically, the algorithm uses the difference of gaussian and threshold segmentation to preprocess to obtain the position of possible target points. Then, local contrast processing is conducted for the possible target points. Finally, the global is processed using an improved directional gradient. According to the experimental results, the proposed algorithm outperforms the existing algorithms in terms of detection probability and false detection probability.

KEYWORDS

Infrared dim and small target detection; local contrast; gradient

1. Introduction

With the constant advancement of infrared imaging technology, infrared imaging systems have been extensively adopted in both military and civilian fields, such as precision guidance, target detection, video surveillance and tracking. However, it is quite difficult to detect the infrared dim and small targets due to low ratios, fewer pixels, and the lack of texture features.

According to the different characteristics of small targets used by the algorithm, the existing detection algorithms can be divided into spatial filtering methods and frequency domain filtering methods based on filtering methods. As for the former, it includes spatial high-pass filtering (Yang *et al.* 2004), median filtering (Deshpande *et al.* 1999) and morphological filtering (Bai and Zhou 2010). Concerning the latter, it includes wavelet transform (Deng *et al.* 2016) and phase spectrum (Guo and Zhang 2009). The above methods are applicable to make simple predictions of the image background. After the high-frequency components are smoothed by using the filtering method, the detection of dim and small target is performed through the differences between the infrared image and the smoothed image. Featuring a fast running speed and simple operation, the above method is applicable for small target detection under simple backgrounds. For those

CONTACT Weihong Lin  422170545@qq.com

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/14498596.2022.2140714>

© 2022 Mapping Sciences Institute, Australia and Surveying and Spatial Sciences Institute

images in complex scenes, small targets are often blurred and frequency characteristics are not obvious. Consequently, the finally-obtained differential image contains more background clutter, which increases the likelihood of false detection. Ultimately, there is a high false alarm rate, which poses challenge to ensuring the validity of detection in complex scenes.

Based on the human visual attention mechanism, it mainly refers to the processing mechanism of the field of view information that is simulated by the human eyes and scans the image to find the region of interest (ROI) (Wang *et al.* 2012). In recent years, many algorithms have been proposed based on the human eye attention mechanism, mainly including Local Contrast Measure (LCM) (Chen *et al.* 2014), Relative Local Contrast Measure (RLCM) (Han *et al.* 2018), Multiscale Patch-Based Contrast Measure (MPCM) (Wei *et al.* 2016) and Weighted Strengthened Local Contrast Measure (WSLCM) (Han *et al.* 2020), etc. These methods can simulate the attention mechanism and contrast mechanism of the human visual system, quickly locate and extract the region of interest, and suppress the background while increasing the saliency of the target. Still, the algorithm is not robust and has a high false alarm rate.

Based on the method of low-rank sparse matrix recovery, sparse representation was mainly used in classification tasks, and target detection can be realised by using frequency feature differences. At present, the algorithms based on low-rank sparse recovery mainly include Infrared Patch-Image Model (IPI) (Gao *et al.* 2013), Weighted Infrared Patch Image (WIPI) (Dai *et al.* 2016), Reweighted Infrared Patch-Tensor Model (RIPI) (Dai and Wu 2017), Non-Convex Rank Approximation Minimisation Joint $l_{2,1}$ Norm (NARM) (Zhang *et al.* 2018), Non-Convex Optimisation with L_p -Norm Constraint (NLOC) (Zhang *et al.* 2019) and Self-Regularised Weighted Sparse Model (SRWS) (Tz *et al.* 2021), etc. These algorithms comprehensively consider the characteristics of the target and the background, can adapt to various scenes, and greatly improve the detection probability of infrared dim and small targets. However, the algorithm based on low-rank sparse recovery is prone to false detection of strong edges and high computational complexity.

In recent years, deep learning methods have been widely applied in various fields. Among them, convolutional neural network has improved significantly in terms of target detection, tracking and segmentation. In respect of infrared dim and small target detection, for example, (Li *et al.* 2022) proposed to achieve progressive feature interaction and adaptive feature enhancement through a tri-directional dense nested interactive module (DNIM) with a cascaded channel and spatial attention module (CSAM). (Chen *et al.* 2022b) proposed a novel end-to-end framework intended for infrared small target detection and segmentation. (Chen *et al.* 2022a) developed a hierarchical overlapped small patch transformer (HOSPT) to encode multi-scale features from the single-frame image, rather than relying on the CNN. However, due to the lack of obvious texture and shape information in infrared dim and small target, it is difficult to learn practical feature information for the algorithm based on depth learning. In addition, it is limited in the training model due to the lack of training datasets.

In summary, although many infrared small target detection methods have been proposed, these methods still have problems such as poor robustness, insufficient stability, high algorithm complexity, long computing time and insufficient real-time performance. Therefore, the precise detection of dim and small targets in infrared images is still

a problem to be solved urgently. This paper has thoroughly studied the extreme value, contrast, gradient and other characteristics of dim and small targets in infrared images and proposed a detection model of dim and small targets in infrared images based on local contrast and gradient, which successfully improved the saliency of the target from multiple aspects. The validity of the model is verified on different sets of actual infrared image data. This model has better background suppression and target detection performance than existing methods.

The remainder of this paper is organised as follows. Section 2 describes the relevant principles, including the Gaussian difference calculation, extreme point detection, local contrast calculation and gradient calculation of dim and small targets in infrared images. Section 3 describes the flow chart of the detection algorithm and the effect diagram of each step is given. Section 4 analyzes the overall performance of the method and the comparison of results with other weak and small infrared target detection methods. Conclusions are given in section 5.

2. Related principles

2.1 Multi-scale spatial structure

The multi-scale space construction is generated by convolving the input image with Gaussian kernels of different scales. The multi-scale space adopts a multi-layer Gaussian pyramid image structure (Lowe 2004). The scale-space of the image is expressed as a function $M(x, y, s)$, which is filtered by Gaussian kernels of different variances:

$$M(x, y, s) = G(x, y, s) * I(x, y) \quad (1)$$

Where $G(x, y, s)$ is the Gaussian function $G(x, y, s) = \frac{1}{2\pi s^2} e^{-\frac{x^2+y^2}{2s^2}}$ with different variances, $I(x, y)$ represents the infrared dim and small image, s represents the variance. The smoothed images of different scale Gaussian filters corresponding to the original image are obtained by selecting different variances.

2.1.1 Difference of Gaussian

To detect interest points in multi-scale space, this paper uses the difference of gaussian (DoG) to subtract the Gaussian filtering results of two images with different parameters k to obtain a DoG image. The specific formula is as follows:

$$\begin{aligned} D(x, y, s) &= (G(x, y, ks) - G(x, y, s)) * I(x, y) \\ &= F(x, y, ks) - F(x, y, s) \end{aligned} \quad (2)$$

2.1.2 Eliminate edge response

The difference of gaussian has a solid response to the edge, and the edge of the image after the difference of gaussian may interfere with the extreme value detection. This paper uses adaptive threshold segmentation to remove unstable edge response points, enhancing matching stability and improving anti-noise ability.

Adaptive threshold segmentation obtains a binary image that calculates the threshold T on the input image and distinguishes the target from the background. For the dim and

small targets in infrared image $I(x, y)$, the image size is $L \times H$. Assume that the threshold value is T , and the number of pixels in the image whose pixel greyscale is less than the threshold value T is denoted as N_0 , whose greyscale value is set to 0. The number of pixels in the image whose greyscale is greater than the threshold T is denoted as N_1 , and its greyscale value is set to 255. Thus, $N_0 + N_1 = L \times H$. The ratio of the number of pixels belonging to the target to the entire image is denoted as $\omega_0 = N_0/L \times H$, and its average grey level is denoted as μ_0 . The ratio of background pixels in the entire image is denoted as $\omega_1 = N_1/L \times H$, and its average grey level is denoted as μ_1 . The total average grey level of the image is denoted as $\mu = \omega_0 \times \mu_0 + \omega_1 \times \mu_1$, and the variance between classes is denoted as g . Thus,

$$\omega_0 + \omega_1 = 1 \quad (3)$$

$$\mu = \omega_0 \times \mu_0 + \omega_1 \times \mu_1 \quad (4)$$

Establish the objective function $g(T)$:

$$g(T) = \omega_0(\mu_0 - \mu)^2 + \omega_1(\mu_1 - \mu)^2 \quad (5)$$

The threshold value that maximises the variance c between classes is obtained by traversing 0 to 255, which is the required value.

2.2 Detection and location of extreme point

To find the extreme point of scale-spaces, the extreme value of scale space is detected for each layer of the DoG pyramid. Take each extreme point as a candidate point, determine the position and scale of the extreme point through a fit function at each candidate position, and remove low-contrast candidate points simultaneously. The specific formula is as follows:

Taylor expansion of spatial scale function:

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (6)$$

Where $X = (x, y, s)^T$, By deriving the above formula and making the equation equal to zero, the offset of the extreme point can be obtained as follows:

$$X = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X} \quad (7)$$

Formula (7) removes low-contrast points. Then take the first two items of formula (6) to get:

$$D(X) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} X \quad (8)$$

Where $\hat{X} = (x, y, s)^T$ represents the offset relative to the interpolation centre. Experiments show that the precise position of the extreme point can be obtained when $D \geq 0.03$.

2.3 Local contrast

Due to the complex background of small infrared targets, background noise and edge response may interfere with extreme value detection. Therefore, this paper uses the difference between the area where the small target is located and the neighbouring area, and proposes an effective local comparison measure. It can suppress background clutter while enhancing the target and significantly improve the signal-to-noise ratio of the image. Use a 3×3 image patch slider to slide the input image (Table 1), where 0 represents the area where the target may appear, and calculate the grey average value in 0–8 cells:

$$M_i = \frac{1}{N_a} \sum_{j=i}^{N_a} C_j^i \tag{9}$$

Where N_a represents the number of pixels in the i th cell, and C_j^i represents the grey value of the j th pixel in the i th cell. The ratio of the centre cell to the surrounding cells is $h_i^n = \frac{L_n}{M_i}$, and L_n represents the maximum grey value of the centre cell of the n th image patch.

The grey value of a small target is generally the minimum value of its field, so the local contrast is defined as:

$$H_n = \min_i L_n \times h_i^n = \min_i L_n \times \frac{L_n}{m_i} = \min_i \frac{L_n^2}{m_i} \tag{10}$$

The larger H_n is, the more likely it is to be the target.

2.4 Direction gradient

An Infrared image can be divided into three parts: background, target and noises. There are two types of backgrounds in the direction gradient: the gradient direction distribution of the flat background is irregular, and the gradient direction of the background with the strong edge is usually consistent. The target is a salient area with local intensity characteristics, and the target pixels have gradient characteristics in each direction. Noises also have multi-directional gradient characteristics but have been suppressed after difference of gaussian. The gradient direction of the background with strong edges is usually the same. The three-part gradient directions are shown in Figure 1(b–d). Figure 1(b) is the red box in Figure 1(a,c) is the yellow box in Figure 1(a,d) is the blue box in Figure 1(a). Therefore, gradient information is an important feature of infrared dim and small targets (Hong et al. 2017).

This paper uses an image slider of size $n \times n$ to traverse the image. A coordinate system is established with the centre of the image block as the origin, and four-square areas

Table 1. 3×3 image patch slider.

1	2	3
4	0	5
6	7	8

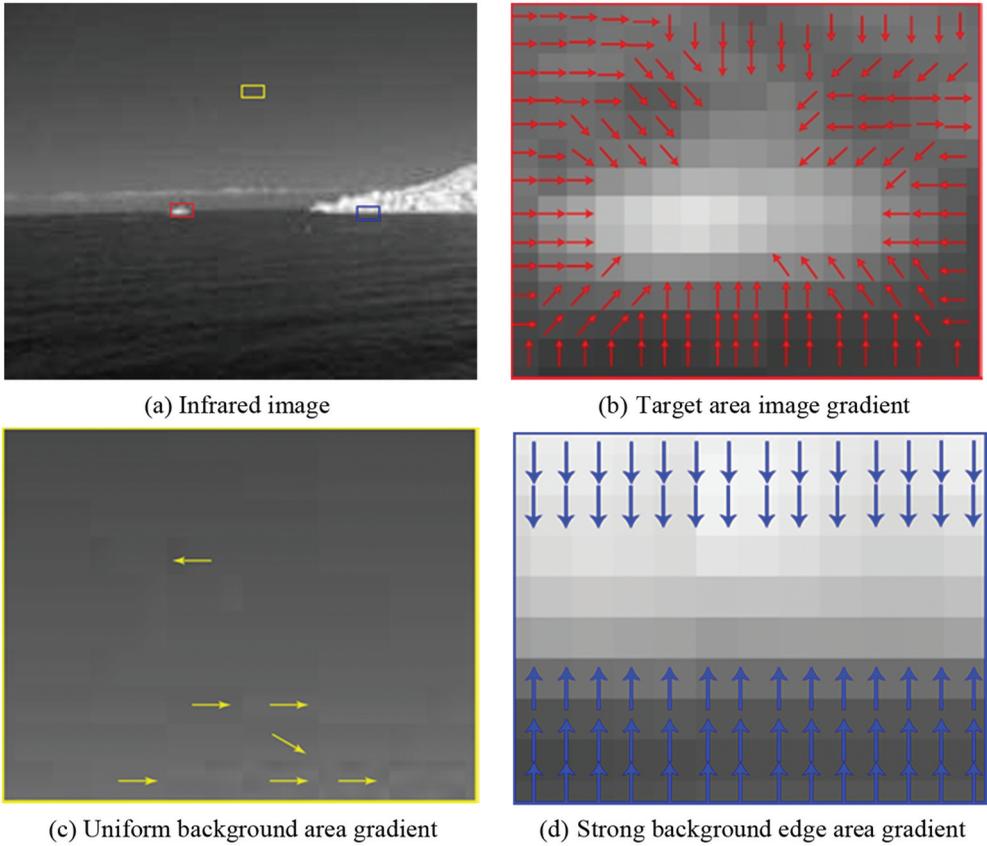


Figure 1. Gradient direction information. (a) Infrared image, (b) Target area image gradient, (c) Uniform background area gradient, (d) Strong background edge area gradient.

R_1, R_2, R_3 and R_4 are formed respectively. The range of each area block is 90° . The i th square area can be expressed as R_i :

$$R_i = \left\{ (\gamma, \theta) \mid \frac{\pi(i-1)}{2} < \theta \leq \frac{\pi i}{2} \right\} \quad (11)$$

Where (γ, θ) is the coordinate of the image block. The gradient direction does not strictly point to the central area because of the complexity of the infrared weak and small target images. The constraints on the gradient vector of infrared weak and small targets in this paper are:

$$\psi_{R_i} = \left\{ g_{v_i}(\alpha, \beta, \gamma, \theta_i) \mid \frac{\pi(i-1)}{2} + \pi < \frac{\pi i}{2} + \pi, (\gamma, \theta_i) \in R_i \right\} \quad (12)$$

Where ψ_{R_i} represents the set of gradients roughly pointing to the centre in the R_i area; $g_{v_i}(\alpha, \beta, \gamma, \theta_i)$ is the gradient at point (γ, θ) in the R_i region, where α and β represent the magnitude and direction of the gradient g_{v_i} respectively. In the i th region, calculate the mean square of the magnitude of the gradient g_{v_i} that roughly points to the centre:

$$G_i = \frac{1}{N_i} \sum_{j=1}^{N_i} ||g^j R_i||^2 \quad (13)$$

Where N_i is the number of g_{V_i} in the set ψ_{R_i} . After obtaining all G_i , the maximum value $G_{\max} = \max_{1 \leq i \leq 4} G_i$ and minimum value $G_{\min} = \min_{1 \leq i \leq 4} G_i$ can be obtained. Therefore, all G_i in an image block can be accumulated as:

$$G = \begin{cases} \sum_{i=1}^4 G_i, & \frac{G_{\min}}{G_{\max}} > k \\ 0, & \frac{G_{\min}}{G_{\max}} \leq k \end{cases} \quad (14)$$

Where G represents the local gradient distribution value in an image block; The threshold k obtained through experiments can determine whether the local gradient distribution is uniform. Therefore, the ratio of G_{\min} to G_{\max} can well suppress background clutter with local directivity.

3. Proposed method

Aiming at the characteristics of infrared dim and small targets, this paper proposes an infrared dim and small target detection model based on local contrast and gradient. The LCG model makes up for the inability to detect or high false alarm rate in complex scenes of small and weak targets. At the same time, the LCG is robust in various backgrounds and achieves better detection results. Figure 2 shows the whole flow chart of the method proposed in this paper. The specific steps are as follows: (1) Construct Gaussian difference pyramid; (2) Eliminate edge response; (3) Extremum point detection; (4) Local contrast calculation; (5) Gradient direction calculation.

3.1 Constructing difference of gaussian pyramid and eliminating edge response

The first step of constructing the Difference of Gaussian pyramid to extract the points of interest of the image is to filter the image using Gaussian filters with different variances s . The second step is to establish four octaves images of different scales, each group contains five intervals of different Gaussian blur scale images (Figure 3). The first octave of scales is the size of the original image, and each subsequent octave is the result of the previous octave of subsampled. The ratio of the scale factor of the Gaussian blur between each layer in a group of images is b , and the Gaussian blur scale of the adjacent group is two times the relationship. Then the optimal threshold T is calculated through the adaptive threshold, which distinguishes the background of the infrared picture from possible targets. Figure 4(a) is a threshold segmentation image, Figure 4(b) is a three-dimensional image segmented by adaptive threshold. After threshold segmentation, the edge response is suppressed and the target is enhanced.

3.2 Extremum point detection

To find the extreme point of the scale space, the middle detection point needs to be compared with its 8 neighbouring points of the same scale and 9×2 points

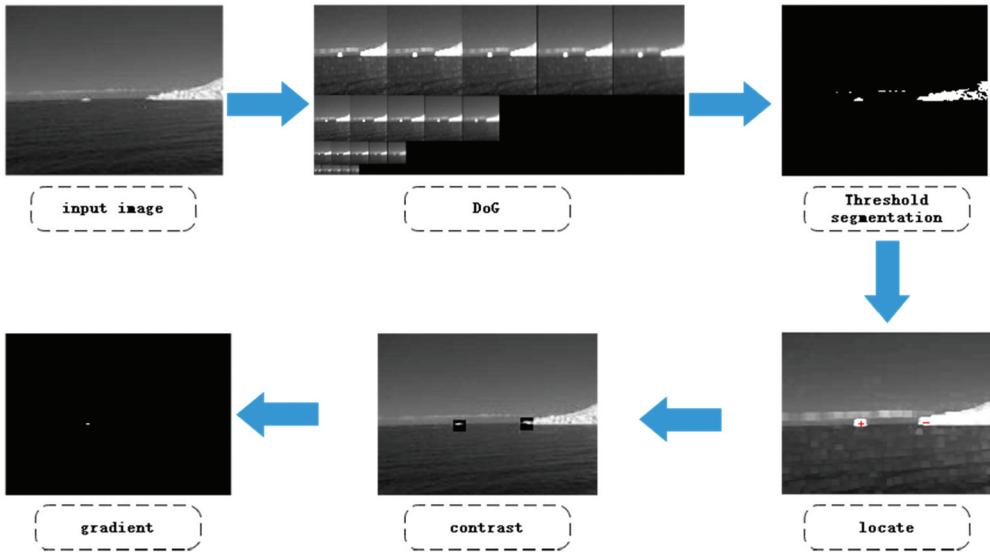


Figure 2. LCG algorithm flowchart.

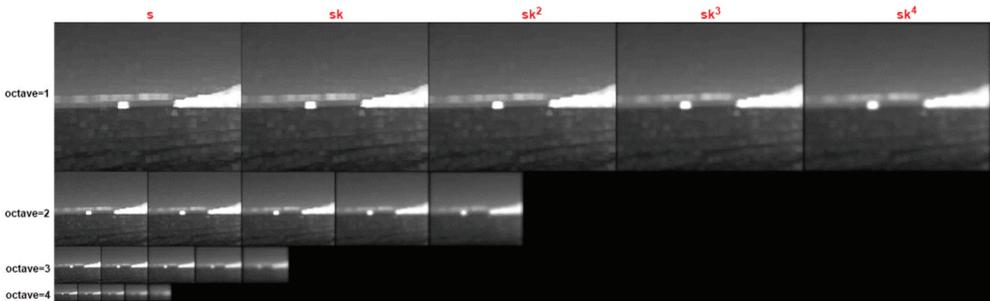


Figure 3. Difference of Gaussian pyramid.

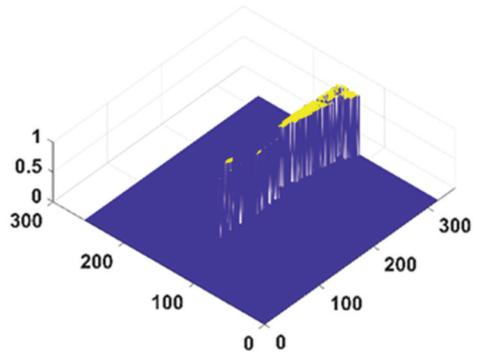
corresponding to the upper and lower neighbouring scales. This ensures that extreme points are detected in both the scale space and the two-dimensional image space (Figure 5(a)). If a point is the maximum value in the 26 domains of this layer and the upper and lower layers of the DoG scale space, the point is considered to be a feature point of the image at that scale (Figure 5(b)).

3.3 Calculation of local contrast and gradient direction

From the definition, it can be seen that local contrast can enhance the object's saliency and suppress the background. After the local contrast calculation process, there is still some interference background, which affects the detection probability of the target (Figure 6(a)). Therefore, the algorithm in this paper adds gradient calculation, which can accurately eliminate the interference background and improve the accuracy. The processing effect is shown in Figure 6(b):

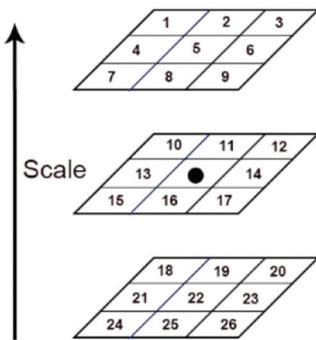


(a) Threshold segmentation image



(b) Three-dimensional image

Figure 4. Elimination of edge response.(a) Threshold segmentation image;(b).Three-dimensional image.



(a) Find extreme points



(b) Locate extreme points

Figure 5. Extreme point detection. (a) Find extreme points; (b) Locate extreme points.

4. Experiments evaluation

In this section, we first introduce commonly used evaluation indicators and then introduce the data set SIRST (Dai *et al.* 2020) used in this article. The data set contains 427 pictures selected from hundreds of infrared image sequences of different scenes. The test results and experimental comparison are carried out under the environment of Intel Core i7-6700HQ CPU@2.60 GHz, and the simulation platform is MATLAB R2020b. Finally, it analyzes the running results of each module of the proposed model, gives the detection results of the model proposed in this paper and compares it with four other small target detection algorithms. The model proposed in this paper reaches the optimal value under various evaluation indicators.

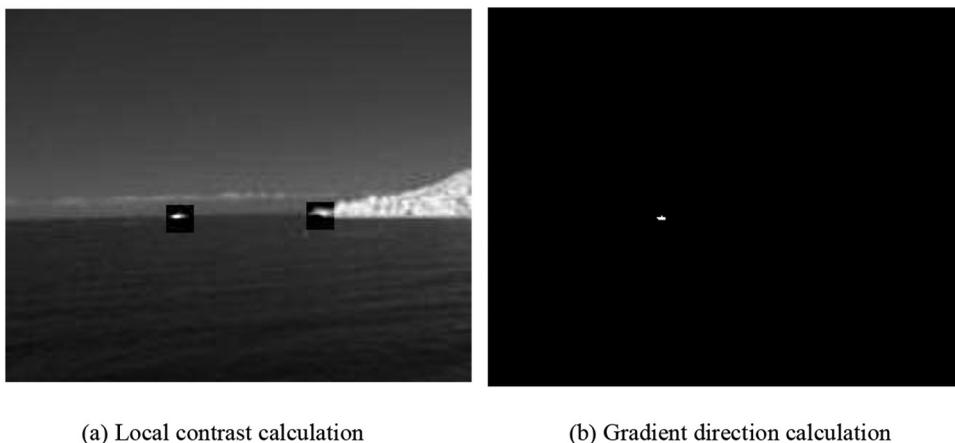


Figure 6. Local contrast and gradient direction calculation processing. (a) Local contrast calculation; (b) Gradient direction calculation.

4.1 Evaluation index

4.1.1 Signal-to-clutter ratio gain and background suppression factor

To quantitatively verify the effectiveness of the algorithm and further objectively evaluate the experimental results, this article uses two common evaluation indexes in target detection: Signal-to-Clutter Ratio Gain (SCRG) and Background Suppression Factor (BSF). Their calculation formulas are:

$$SCRG = 20 \times \lg \left(\frac{(m_T/\sigma_B)_{out}}{(m_T/\sigma_B)_{in}} \right) \quad (15)$$

$$BSF = 20 \times \lg \left(\frac{\sigma_{in}}{\sigma_{out}} \right) \quad (16)$$

Where m_T represents the mean value of the target area; σ_B represents the standard deviation of the background; σ_{in} represents the standard deviation of the input original image background; σ_{out} represents the standard deviation of the image background processed by the algorithm; $(m_T/\sigma_B)_{out}$ represents the parameters of the original input image; $(m_T/\sigma_B)_{in}$ represents the parameters of the image processed by the algorithm.

The SCRG value represents the gain effect of the image background before and after the algorithm is processed. The BSF value represents the suppression effect of the algorithm on the image background. For the same image, the higher the SCRG and BSF values obtained by the algorithm, the better the effect of enhancing the saliency of the target and suppressing background clutter. At the same time, the detection result of the algorithm is more accurate.

4.1.2 ROC curve

ROC (Receiver Operating Characteristic) curve is currently the most commonly used visual saliency model evaluation index. It is a two-dimensional curve that changes with the

change of detector decision threshold. ROC curve is composed of detection probability P_d and false alarm rate F_a , which is defined as follows:

$$P_d = \frac{\text{Number of real targets detected}}{\text{Actual target number}} \quad (17)$$

$$F_a = \frac{\text{Number of false targets detected}}{\text{Total number of targets detected}} \quad (18)$$

The number of detected targets is calculated by grey-scale clustering of the detection result image. When the detected target pixel overlaps the real target pixel or the distance between the centre pixels of the two is within five pixels, it means that the detected target is the real target. Otherwise, it is a false target.

4.2 Introduction to dataset

To intuitively evaluate the detection performance of the detection algorithm proposed in this paper for dim and small targets in infrared images, this section selects four real infrared images under different background conditions for experimental simulation and comparative analysis of the algorithm. The representative images of the four experimental atlases are shown in Figure 8. The small target in each image has been marked with a red box, and a three-dimensional image of each image is given.

The experimental atlas includes four kinds of infrared images with different backgrounds: The image in Figure 7(a) is a cloud background, and the cloud structure in the image is complex. Figure 7(e) is the three-dimensional image of Figure 7(a). Figure 7(b) is a background image of the sea and sky. There is strong interference at the junction of sea and sky in the image. Figure 7(f) is the three-dimensional image of Figure 7(b). The image in Figure 7(c) is a background image of complex buildings and clouds. The edges of strong buildings in the image have strong interference. Figure 7(g) is the three-dimensional

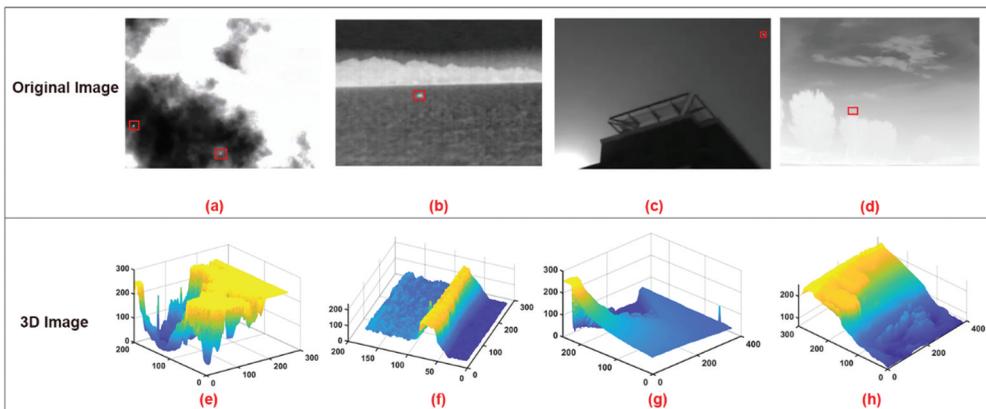


Figure 7. Infrared images and three-dimensional images of four different backgrounds (a) Cloud background image, (b) Sea-sky background image, (c) Building background image, (d) Ground background image, (e) 3D drawing of cloud layer background, (f) 3D drawing of the sea-sky background, (g) 3D drawing of the back of the building, (h) 3D drawing of the ground background.

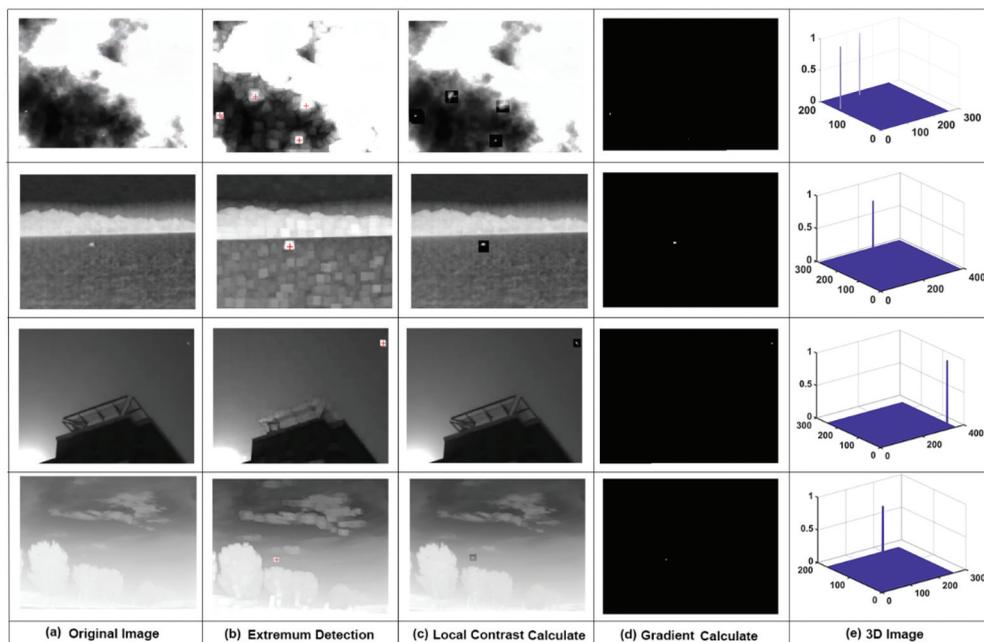


Figure 8. Detection results. (a) original image; (b) feature extraction diagram; (c) local contrast calculation diagram; (d) gradient calculation diagram; (e) 3D image of test results.

image of Figure 7(c). Figure 7(d) is a ground background image in which the ground background has strong interference. Figure 7(h) is the three-dimensional image of Figure 7(d).

4.3 Analysis of test results

We use four types of images of the dataset to verify the robustness and reliability of the detection algorithm proposed in this paper for detecting infrared targets in different backgrounds. The experimental results are shown in Figure 8. Figure 8 (column a) is the original image. Figure 8 (column b) is the image after extreme value detection. Figure 8 (column c) is the image after local contrast processing. Figure 8 (column d) is the image after gradient processing image. Figure 8 (column e) is a three-dimensional view of the detection.

It can be known from the simulation results that all the targets in the image have been detected. (1) For the a-type complex cloud background image, because the cloud is bright, small targets and cloud edges may be detected as local extrema. It can be known that after the local contrast and gradient calculation, the smooth background is almost all filtered out and the edge of the cloud is well suppressed. Infrared small and weak targets are accurately detected. (2) For c- type sea and sky background image, the contrast between the sea and sky is high, and the edge can easily cause false alarms. The target detection result is accurate after background suppression and edge response elimination. (3) For the b-type building background image, the grey distribution of the building is uneven, and the edge of the building can easily cause false alarms. After the

gradient calculation, the edge can be suppressed, and the target detection probability is improved. (4) For the d-type ground background image, the ground background is complex and the greyscale distribution is uneven, which is easy to detect errors in the extreme value detection. After local contrast and gradient calculation, false targets are eliminated and the target detection probability is improved. Through simulation experiments on infrared images with different backgrounds, the detection algorithm proposed in this paper can well suppress the background and accurately detect the target under various complex backgrounds.

4.4 Comparative analysis with other algorithms

4.4.1 Comprehensive evaluation of different algorithms

To further measure the robustness and reliability of the detection algorithm proposed in this paper for infrared dim and small target detection under different backgrounds, Tri-layer Template Local Difference Measure (TTLDM) (Jing and Weihua 2022), WSLCM, Nonconvex Approximate Tensor Fibred Rank (TFNN) (Kong *et al.* 2022) and SRWS algorithms and the algorithm LCG in this paper are compared and analysed in the data set. The experimental simulation comparison results are shown in Figure 9. The red mark indicates the target and the green mark indicates the target of the detection error.

The algorithm comparison analysis is as follows:

(1) The TTLDM and WSLCM algorithms are based on grey-scale local contrast. Because the LCM algorithm relies too much on the difference of grey levels, when the background is too complex, they easily cause serious false alarms. TTLDM uses the difference of grey distribution between three-layer templates to detect targets. Although it is better than LCM, the false alarm rate is higher in complex scenes. WSLCM is an improved local

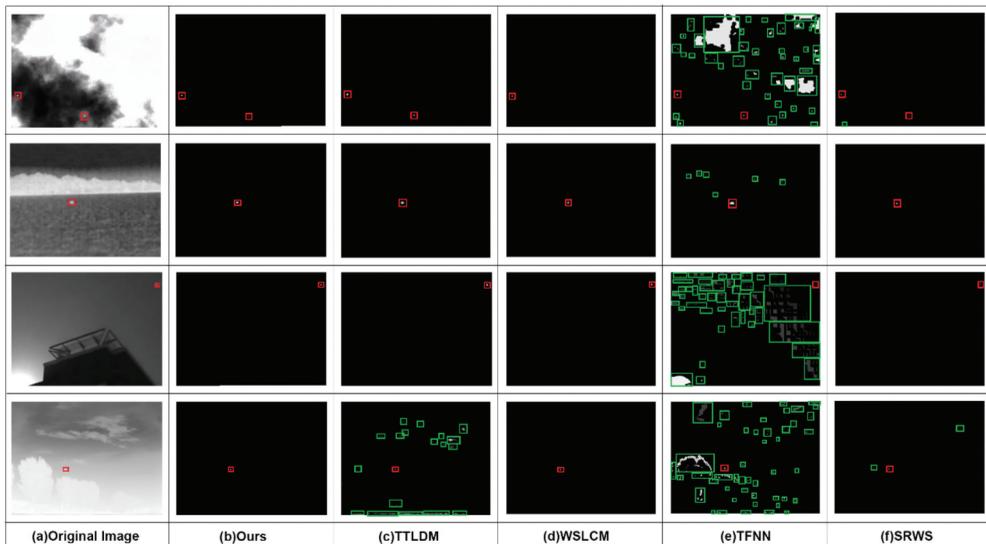


Figure 9. Comparison results with other eight algorithms. (a) Original image, (b) Ours detection images, (c) TTLDM detection images, (d) WSLCM detection images, (e) TFNN detection images, (f) SRWS detection images.

4.4.3 ROC curve analysis of different algorithms

This paper selects TTLDM, WSLCM, RLCM, NOLC and SRWS algorithms and the algorithm LCG of this paper to objectively measure the performance of the algorithm through the ROC curve. To measure the anti-noise ability of the algorithms, this paper adds Gaussian noise with a variance of 0.001 and 0.01 to the dim and small targets in infrared images. Figure 10 compares ROC curves of different algorithms under four types of data sets. The Area Under Curve (AUC) needs to be calculated for quantitative purposes. AUC can be used as a criterion for judging the pros and cons of an algorithm. The larger the AUC, the more accurate the algorithm detection. Table 3 shows the AUC value of the original images in Figure 10. Table 4 shows the AUC value of add 0.001 noise images in Figure 10. Table 5 shows the AUC value of add 0.01 noise images in Figure 10.

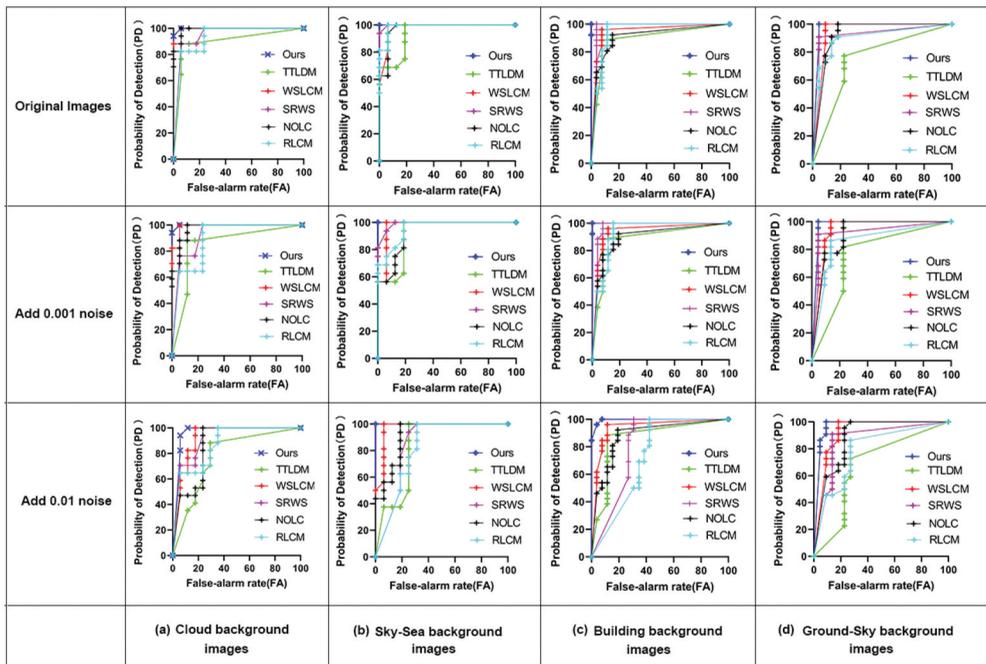


Figure 10. ROC curves of different algorithms under four types of data sets. (a) Cloud background images, (b) Sky-Sea background images, (c) Building background images, (d) Ground-sky background images.

Table 3. The AUC value of the original images in Figure 10.

Method	Cloud background images AUC	Sky-Sea background images AUC	Building background images AUC	Ground background images AUC
Ours	0.9983	1.0000	1.0000	0.9752
TTLDM	0.9014	0.9434	0.8994	0.7521
WSLCM	0.9931	0.9805	0.9497	0.9483
SRWS	0.9498	0.9980	0.9726	0.9236
NOLC	0.9896	0.9746	0.9127	0.9298
RLCM	0.9343	0.9883	0.9467	0.8998

Table 4. The AUC value of the add 0.001 noise images in [Figure 10](#).

Method	Cloud background images AUC	Sky-Sea background images AUC	Building background images AUC	Ground background images AUC
Ours	0.9983	1.0000	1.0000	0.9742
TTLDM	0.8547	0.9199	0.8942	0.7593
WSLCM	0.9896	0.9727	0.9423	0.9390
SRWS	0.9291	0.9902	0.9689	0.9236
NOLC	0.9723	0.9336	0.9024	0.9122
RLCM	0.8979	0.9570	0.9327	0.8595

Table 5. The AUC value of the add 0.01 noise images in [Figure 10](#).

Method	Cloud background images AUC	Sky-Sea background images AUC	Building background images AUC	Ground background images AUC
Ours	0.9637	1.0000	0.9956	0.9669
TTLDM	0.7785	0.8359	0.8539	0.6798
WSLCM	0.9256	0.9707	0.9371	0.9236
SRWS	0.9187	0.8477	0.8047	0.8667
NOLC	0.8651	0.9180	0.8817	0.8853
RLCM	0.8120	0.8223	0.7300	0.7903

It can be known from [Figure 10](#), [Tables 3–5](#): (1) Compared with the other five algorithms, the algorithm in this paper obtains the highest detection probability in all cases, and the ROC curves in the four types of scenes are all located in the upper left corner. (2) The algorithm proposed in this paper achieves the maximum AUC value in the four types of scenarios. The closer the AUC is to 1.0, the higher the authenticity of the detection method. (3) The comparison by adding different noises shows that the effect of the six algorithms is not obvious when adding 0.001 noise, but when adding 0.01, TTLDM, SRWS, NOLC and RLCM show a significant decline and detect other noises. The algorithm in this paper and the WSLCM algorithm have better anti-noise ability.

5. Conclusion

Faced with the difficulty of detecting dim and small infrared targets in complex backgrounds, this paper proposes a detection algorithm for dim and small infrared targets based on local contrast and gradient. We improved both local contrast and directional gradient methods and combined the two methods. After pre-processing, local contrast is applied to possible target points to enhance the contrast of the target, and finally directional gradient is applied to the global. Compared with the traditional local contrast method or directional gradient method, the detection probability of the LCG is greatly improved. According to the simulation experiment results and comparative analysis of dim and small targets in infrared images, it proves that the algorithm in this paper has good robustness, background suppression effect, and detection effect for complex scenes. In the quantitative evaluation, the algorithm in this paper is superior to the compared algorithms in terms of SCRG, BSF, ROC and anti-noise.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China [No.61775140, No. 61875125, No. 62275153]; Natural Science Foundation of Shanghai [No. 18ZR1425800].

ORCID

Weihong Lin  <http://orcid.org/0000-0001-7349-1611>

Zimin Shen  <http://orcid.org/0000-0001-6699-1247>

References

- Bai, X. and Zhou, F., 2010. Analysis of new top-hat transformation and the application for infrared dim small target detection[J]. *Pattern Recognition*, 43 (6), 2145–2156. doi:10.1016/j.patcog.2009.12.023
- Chen, C.L.P, et al., 2014. A local contrast method for small infrared target detection [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 52 (1), 574–581. doi:10.1109/TGRS.2013.2242477
- Chen, Y., et al., 2022b. A multi-task framework for infrared small target detection and segmentation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–9.
- Chen, G., Wang, W., and Tan, S., 2022a. IRSTFormer: a hierarchical vision transformer for infrared small target detection[J]. *Remote Sensing*, 14 (14), 3258. doi:10.3390/rs14143258
- Dai, Y., et al., 2016. Infrared small target and background separation via column-wise weighted robust principal component analysis[J]. *Infrared Physics and Technology*, 77, 421–430. doi:10.1016/j.infrared.2016.06.021
- Dai, Y., et al., 2020. Asymmetric contextual modulation for infrared small target detection[J].
- Dai, Y. and Wu, Y., 2017. Reweighted infrared patch-tensor model with both non-local and local priors for single-frame small target detection[J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 10(8), 1–16.
- Suyog, D. D., M. Hwa Er, Ronda, V., and Philip, C. 1999. "Max-mean and max-median filters for detection of small target", *Proc. SPIE 3809, Signal and Data Processing of Small Targets 1999*. <https://doi.org/10.1117/12.364049>
- Deng, H., et al., 2016. Infrared small-target detection using multiscale gray difference weighted image entropy[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 52 (1), 60–72. doi:10.1109/TAES.2015.140878
- Gao, C., et al., 2013. Infrared patch-image model for small target detection in a single image[J]. *IEEE Transactions on Image Processing*, 22 (12), 4996–5009. doi:10.1109/TIP.2013.2281420
- Guo, C. and Zhang, L., 2009. A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression [J]. *IEEE Transactions on Image Processing*, 19 (1), 185–198.
- Han, J., et al., 2018. Infrared small target detection utilizing the multiscale relative local contrast measure[J]. *IEEE Geoscience and Remote Sensing Letters*, 15 (4), 1–5.
- Han, J., et al., 2020. Infrared small target detection based on the weighted strengthened local contrast measure. *IEEE Geoscience and Remote Sensing Letters*, 18(9), 1670–1674.
- Hong, Z., et al., 2017. Infrared small target detection based on local intensity and gradient properties[J]. *Infrared Physics & Technology*, 89, 88–96.
- Jing, M.U., et al., 2022. Infrared small target detection using tri-layer template local difference measure[J]. *Optics and Precision Engineering*, 30 (7), 869–882.

- Kong, X., et al., 2022. Infrared small target detection via nonconvex tensor fibered rank approximation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–21.
- Li, B., et al., 2022. Dense nested attention network for infrared small target detection[J]. *IEEE Transactions on Image Processing*, 99, 1. doi:[10.1109/TIP.2022.3199107](https://doi.org/10.1109/TIP.2022.3199107)
- Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 60 (2), 91–110. doi:[10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94)
- Tz, A., et al., 2021. Infrared small target detection via self-regularized weighted sparse model - ScienceDirect[J]. *Neurocomputing*, 420, 124–148. doi:[10.1016/j.neucom.2020.08.065](https://doi.org/10.1016/j.neucom.2020.08.065)
- Wang, X., Lv, G., and Xu, L., 2012. Infrared dim target detection based on visual attention[J]. *Infrared Physics & Technology*, 55 (6), 513–521. doi:[10.1016/j.infrared.2012.08.004](https://doi.org/10.1016/j.infrared.2012.08.004)
- Wei, Y., You, X., and Li, H., 2016. Multiscale patch-based contrast measure for small infrared target detection [J]. *Pattern Recognition*, 58, 216–226. doi:[10.1016/j.patcog.2016.04.002](https://doi.org/10.1016/j.patcog.2016.04.002)
- Yang, L., Yang, J., and Yang, K., 2004. Adaptive detection for infrared small target under sea-sky complex background[J]. *Electronics Letters- lee*, 40 (17), 1.
- Zhang, L., et al., 2018. Infrared small target detection via non-convex rank approximation minimization joint $l_{2,1}$ norm. *Remote Sensing*, 10(11), 1821.
- Zhang, T., et al., 2019. Infrared small target detection based on non-convex optimization with L_p -norm constraint. *Remote Sensing*, 11 (5), 559.