International Journal of Remote Sensing

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tres20

Research on infrared dim and small target detection algorithm based on frequency domain difference

Leihong Zhang, Haiqing Miao, Jian Chen, Weihong Lin, Runchu Xu & Dawei Zhang

INTERNATIONAL JOURNAL OF REMOTE SENSING
2022, VOL. 43, NO. 8, 2942–2964
https://doi.org/10.1080/01431161.2022.2075245

Research on infrared dim and small target detection algorithm based on frequency domain difference

Leihong Zhang[a], Haiqing Miao[a], Jian Chen[b], Weihong Lin[a], Runchu Xu[c] and Dawei Zhang[c,d]

[a]College of Communication and Art design, University of Shanghai for Science and Technology, Shanghai, China; [b]Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, Jilin, China; [c]and Computer Engineering, University of Shanghai for Science and TechnologySchool of Optical-Electrical, Shanghai, China; [d]Shanghai Institute of Intelligent Science and Technology, Tong ji University, Shanghai, China

ABSTRACT

The high accuracy of infrared dim and small target detection in complex backgrounds is of great relevance for infrared identification and tracking systems. Traditional infrared dim and small target detection methods suit scenes with a single and homogeneous continuous background. However, human vision system methods suffer from an undetectable or high false alarm rate in complex scenes with dim small targets. To address this shortcoming, this paper proposes an infrared dim and small target detection algorithm based on frequency domain differencing (FDD). The proposed algorithm consists of a spectrum residual module and a Gaussian greyscale difference module. Firstly, the target enhancement image is constructed by using the spectrum residual module to highlight small targets and suppress background noise. Secondly, the local contrast of the image is enhanced by the Gaussian grayscale difference module, which accurately depicts the edge information of small targets and locates them. Finally, the target enhancement image and Gaussian grayscale difference image are fused to detect infrared dim and small targets. Experimental results show that the proposed algorithm has higher accuracy under the evaluation metrics of local signal-to-noise ratio gain (LSNRG), signal-to-clutter ratio gain (SCRG) and background suppression factor (BSF). At the same time, compared with other algorithms, the detection rate of the proposed algorithm is higher for infrared dim and small targets in complex scenes. Code is available at https://github.com/m156879/FDD-module.

ARTICLE HISTORY
Received 29 December 2021
Accepted 30 April 2022

KEYWORDS
Spectrum residual; Gaussian grayscale difference; human visual system; infrared dim and small target detection

1. Introduction

Infrared imaging systems are now used in various applications such as fault diagnosis, target detection, and video surveillance (Goodall, Bovik, and Paulter 2016; Jakubowicz, Lefebvre, and Moulines 2012; Deng et al. 2017). As one of the key technologies in infrared imaging systems, infrared small target detection is still a difficult area of research (Zhong,

CONTACT Haiqing Miao m15679@163.com College of Communication and Art design, University of Shanghai for Science and Technology, Shanghai 200093, China

Li, and Miao 2014). The signal-to-noise ratio of infrared images is low, small targets generally appear in point shape, and there is no obvious shape and texture information, which brings trouble to infrared small target detection (Zhang, Cong, and Wang 2003).

The existing infrared dim small target detection methods fall into two main categories: sequence detection methods and single-frame detection methods (Pang et al. 2020). Due to the rapid relative movement between the small target and the imaging system, the static background hypothesis is less likely to be established (Pan et al. 2014). The performance of the sequence detection method is reduced. Therefore, researchers tend to study the method of single-frame detection. Single frame detection methods are roughly divided into filtering method, low-rank sparse matrix restoration method (Lu, Lin, and Yan 2015), and human vision system methods. Researchers have proposed a series of filtering methods. For example, Max-mean filter (Deshpande et al. 1999), Max-median filter (Deshpande et al. 1999) and Top-hat filter (Bai and Zhou 2010). This type of method uses filters to predict the background image to suppress background clutter. It is suitable for infrared image scenes with a single background and a small target size. When the target size changes within a larger range, it is usually impossible to detect accurately. In the low-rank sparse matrix restoration methods, the infrared patch-image (IPI) (Gao et al. 2013) model is first proposed. According to the sparseness of the target and the low rank of the background in the infrared image, the IPI model separates the background and the target in the infrared image. Subsequently, a large number of improved algorithms based on the IPI model appeared, such as weight infrared patch-image (WIPI) (Dai, Wu, and Song 2016) and reweighted infrared patch-tensor (RIPT) (Dai and Wu 2017). The low-rank sparse matrix restoration methods are almost suitable for all kinds of complex and rapidly changing backgrounds. Although the detected target position is correct, the target size and shape deviate. And with the higher computational complexity, it can not meet the requirements of real-time applications. Methods based on the human visual system include two types of methods based on spectral residuals and based on local contrast. The algorithm based on spectral residual is simple and easy to implement, but it can not suppress background clutter well, such as an adaptive Butterworth high-pass filter based on frequency domain (Yang, Yang, and Yang 2004). It suppresses the low-frequency components of the image, highlights and enhances the high-frequency information, thereby enhancing the target containing high-frequency components. The method based on local contrast can not lose the features of small targets, but it is not suitable for detecting dark targets. For example local contrast measure (LCM) (Chen et al. 2013), multiscale patch-based contrast measure (MPCM) (Wei, You, and Li 2016), and weighted local difference measure (WLDM) (Deng et al. 2016).

The above-mentioned methods based on the human visual system have poor detection performance or high false alarm rate in a complex scene with a small dark target and cannot well suppress background clutter. Aiming at this problem, this paper proposes an infrared dim and small target detection method based on frequency domain difference (FDD). It mainly includes the following research steps:

(1) The SR model first analyzes the logarithmic spectrum of the input image, then obtains the remaining spectrum of the input image in the spatial domain, and finally constructs a saliency map (Hou et al. 2007). This process highlights the small target information in the image and suppresses the background information, which enhances the contrast between the target and the background.

(2) Combining the characteristics of infrared image direction gray value, this paper constructs a Gaussian grayscale difference map. The Gaussian grayscale difference map accurately depicts the edge information of the small target and locates the small target.
(3) The advantages of the saliency map and the grayscale difference map are selected and merged into the response diagram. The response diagram is segmented by an adaptive threshold to obtain the detection result.

Experimental results show that the proposed method achieves a high detection rate of 0.981 in various scenarios. And it makes up for the shortcomings of poor detection or high false alarm rate based on the human visual system method in the complex scene of dim and small targets. At the same time, the experimental results in different scenes show that the proposed method is simple in operation and more robust to complex scenes, and can effectively detect small infrared targets.

2. Related principles

2.1. Spectrum residual module

The SR model is independent of object characteristics, categories or other forms of prior knowledge. By analysing the log spectrum of the input image, the residual spectrum of the input image can be obtained. Then the saliency map of the input image can be constructed quickly.

From the perspective of information theory (Olshausen and Field 1996), actual coding divides image information into two parts:

$$\mathbf{H}(\mathbf{Image}) = \mathbf{H}(\mathbf{Innovation}) + \mathbf{H}(\mathbf{PriorKnowledge}) \tag{1}$$

Where $\mathbf{H}(\mathbf{Innovation})$ indicates the 'salient' part of the image, and $\mathbf{H}(\mathbf{PriorKnowledge})$ represents the redundant information that needs to be compressed by the encoding system. In image statistics, redundant information is equivalent to the statistical invariant features of the environment. It is now generally accepted that natural images are not random; they follow a highly predictive distribution.

In the SR model, the remaining spectrum contains the 'salient' information of the image. This is similar to scene compression. Through the inverse Fourier transform, the saliency map of the input image can be constructed. Since the saliency map mainly contains the salient part of the scene, and the residual spectrum can also be understood as the salient part of the image, the estimation error can be expressed by squaring the value of each point of the saliency map. For a better visual effect, a Gaussian filter ($s = 2.5$) to smooth the image.

Given the input image $I(x)$:

$$\mathbf{A}(f) = R(F[\mathbf{I}(\mathrm{x})]) \tag{2}$$

$$\mathbf{P}(f) = S(F[\mathbf{I}(\mathrm{x})]) \tag{3}$$

$$\mathbf{L}(f) = \log(\mathbf{A}(f)) \tag{4}$$

$$\mathbf{R}(f) = \mathbf{L}(f) - h_n(f) \times \mathbf{L}(f) \tag{5}$$

$$\mathbf{S}(x) = g(x) \times F^{-1}[\exp(\mathbf{R}(f) + i \times \mathbf{P}(f))]^2 \tag{6}$$

Where F,$F^{-1}$and h_n are the Fourier transform, the inverse Fourier transform and mean filtering function, respectively. $S(F[\mathbf{I}(x)])$ refers to the phase value of the image after Fourier transform. $\mathbf{P}(f)$ represents the phase spectrum of the image. $R(F[\mathbf{I}(x)])$ is to calculate the amplitude of the image after Fourier transform. $\mathbf{A}(f)$ represents the amplitude spectrum of the image. $\mathbf{S}(\mathbf{x})$ is the SR saliency map obtained.

2.2. Gaussian greyscale difference module

Firstly, the average grey value of the pixels in the neighbourhood of the input image is calculated to obtain the image $\mathbf{I}_{ave}$. Secondly, the image $\mathbf{I}_R$ is obtained by convolving the input image using Gaussian kernel ($\mathbf{R}_g$). Finally, subtract $\mathbf{I}_{ave}$ and $\mathbf{I}_R$ obtain a Gaussian greyscale difference map. In order to effectively remove noise and reduce the amount of calculation, the Gaussian kernel ($\mathbf{R}_\mathrm{g}$) we choose the size 5×5:

$$\mathbf{R}_g = \frac{1}{256}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \tag{7}$$

$\mathbf{H}(x, y)$ is the gaussian greyscale difference image of a given image:

$$\mathbf{H}(x, y) = |\mathrm{I}_{ave}(x, y) - \mathbf{I}_R(x, y)| \tag{8}$$

3. Proposed method

Aiming at the deficiency of the human vision system method in detecting small dark targets in complex scenes, an infrared small target detection method based on frequency domain difference is proposed in this paper. It makes up for the shortcomings of low detection performance or high false alarm rate based on the human visual system method in the complex scene of dim and small targets. Figure 1a shows the detection infrared dim and small targets using the method proposed in this paper. The specific steps are as follows:

(1) The generation of SR saliency map and Gaussian greyscale difference map.

In the SR model, this paper first analyzes the logarithmic spectrum of the input image, obtains the remaining spectrum of the input image, and finally generates the saliency map of the input image. The SR model suppresses the background information in the original image and highlights the target information. Figure 2 shows the result of the SR model. At the same time, in the Gaussian greyscale difference model, the average value of the neighbourhood pixels of the original image and the Gaussian convolution result of the original image is calculated, and the Gaussian grayscale difference image is generated. Gaussian grayscale difference image enhances the contrast between small target and local background, and obtains a more accurate position of a small target. Figure 3 shows the processing results of the Gaussian grayscale difference model.

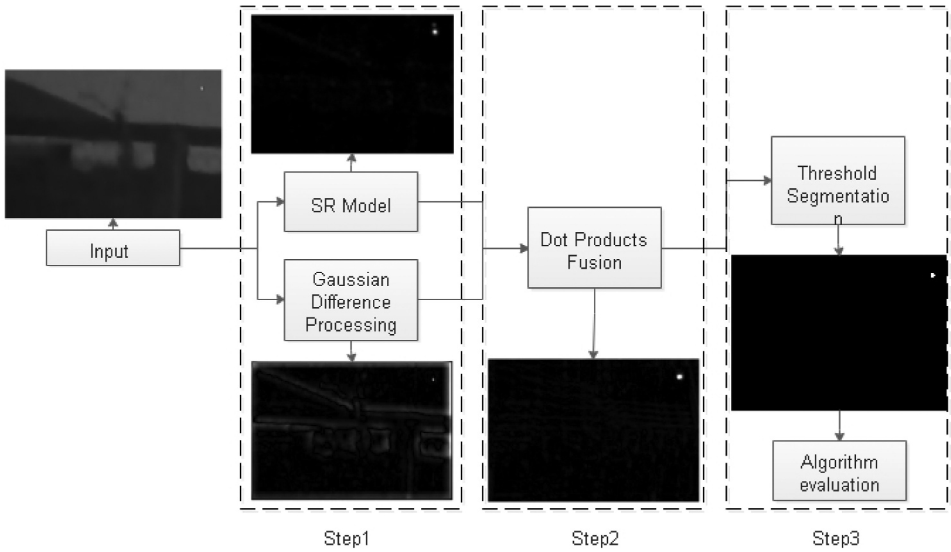


Figure 1a.

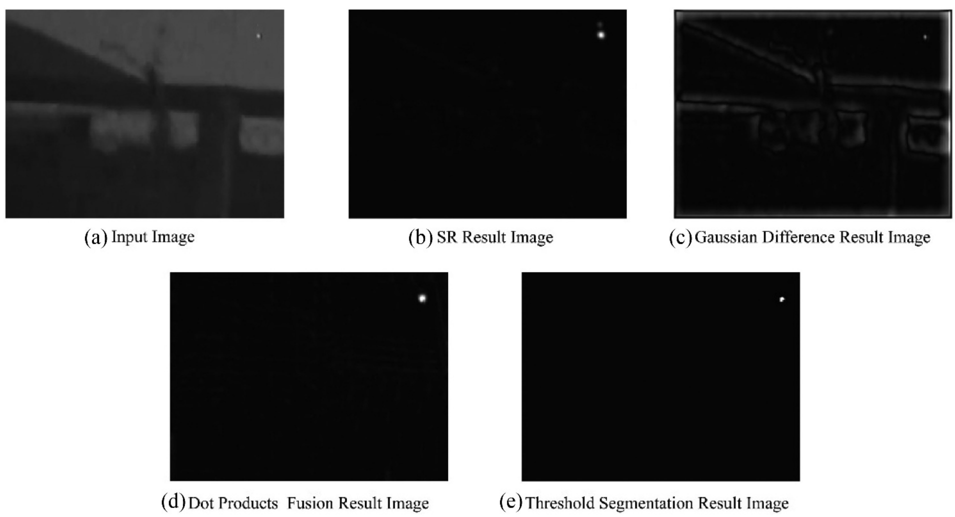


Figure 1b. FDD algorithm flowchart.

(2) The fusion of SR saliency map and Gaussian greyscale difference map.

Through the Otsu method (Otsu 2007), the thresholds of the SR saliency map and the Gaussian greyscale difference map are obtained. Binarize both according to the obtained threshold. The binary images of the SR model and Gaussian grayscale difference model are fused by matrix point multiplication to obtain the response map. The non-small target areas in both binary images are filtered out by the dot product operation. The small targets are thus highlighted.

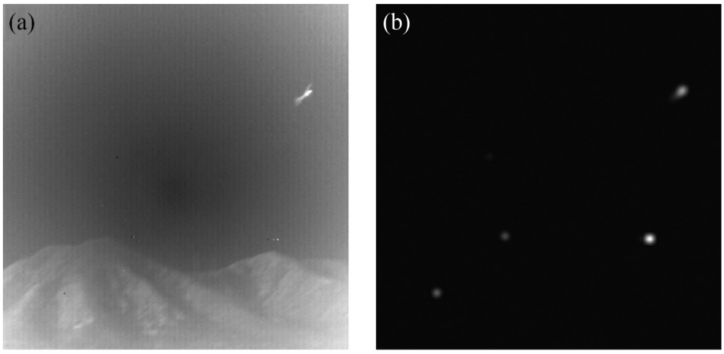


Figure 2. Schematic image of SR model processing results. (a) original image; (b) SR model processing result image.

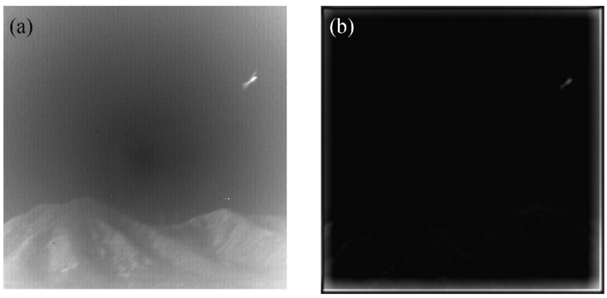


Figure 3. Schematic image of Gaussian greyscale difference module processing results. (a) original image; (b) Gaussian grayscale difference module processing result image.

(3) Target detection and evaluation index.

The fused image is segmented by adaptive threshold segmentation. The excess clutter is filtered out, and the accurate position of the small target is obtained. Adaptive thresholds were derived from the mean and standard deviation of the response map. In order to comprehensively evaluate the detection results, this paper uses local signal-to-noise ratio gain (LSNRG), signal-to-clutter ratio gain (SCRG) and background suppression factor (BSF), detection rate (Pd), and false detection rate (Fa) to compare the background suppression performance and detection performance of each algorithm. LSNRG describes the intensity of target enhancement in its neighbourhood before and after the algorithm processing, which is defined as:

$$\mathrm{LSNRG} = \frac{LSNR_{out}}{LSNR_{in}}, \mathrm{LSNR} = \frac{P_T}{P_B} \tag{9}$$

where $LSNR_{\mathrm{in}}$ and $LSNR_{out}$ represent the LSNR values before and after background suppression, respectively. P_T and P_B are the maximum grey values of the target and its neighbourhood. Neighborhood width value is $d = 20$. BSF describes the degree of background suppression before and after the algorithm is processed and is specifically defined as:

$$\mathrm{BSF} = \frac{\sigma_{in}}{\sigma_{out}} \tag{10}$$

where σ_{in} and σ_{out} indicates the standard deviation of the background neighbourhood before and after background suppression. SCRG is the most used evaluation index. It describes the degree of enhancement of the target relative to the background before and after the algorithm and is also used to describe the difficulty of detecting small targets. It is defined as:

$$\mathrm{SCRG} = \frac{SCR_{out}}{SCR_{in}}, \mathrm{SCR} = \frac{|\mu_t - \mu_b|}{\sigma_b} \tag{11}$$

where μ_t is the average grey value of the target, $\mu_b$andσ_b represents the average gray value and standard deviation of the neighbourhood area. For the above three indicators, the higher their values, the better the background suppression performance of the algorithm. Pd and Fa are the key indicators to evaluate the test results. The larger the value of Pd and the lower the value of Fa, the better the detection performance of the algorithm.

4. Experiments and results

In this section, we first introduce the dataset SIRST (Dai et al. 2020) used in this article. The dataset contains 427 pictures. Each picture is selected from hundreds of infrared image sequences of different scenes. Then analyse the operation results of each module in the proposed algorithm in detail. The detection results of the proposed algorithm are compared and analysed with other eight detection algorithms. The proposed algorithm reaches the optimal value under various evaluation indicators. At the same time, after adding different proportions of Gaussian noise and salt and pepper noise to the original data set to simulate the noise interference in the imaging process, the detection results are compared again.

4.1. Dataset

We chose the SIRST data set for algorithm verification. Figure 4 shows the infrared dim and small target images with representative backgrounds such as heavy sky clouds, sea level scenes, and obvious building obscurations as the detection images. In Figure 5, we have selected three images representing the sea level background, sky background, and ground background to draw the three-dimensional diagram and the three-dimensional diagram of the target neighbourhood. It can be seen that the target is not obvious in the whole image, but it is prominent in its neighborhood.

4.2. Analysis of the detection results of the proposed algorithm

4.2.1. Analysis of spectrum residual module and Gaussian grayscale difference module

By observing the existing infrared small target dataset, we find that the grey value of the small target is higher than that of the neighbourhood background. Therefore, we use the SR saliency detection model to suppress the background information in the

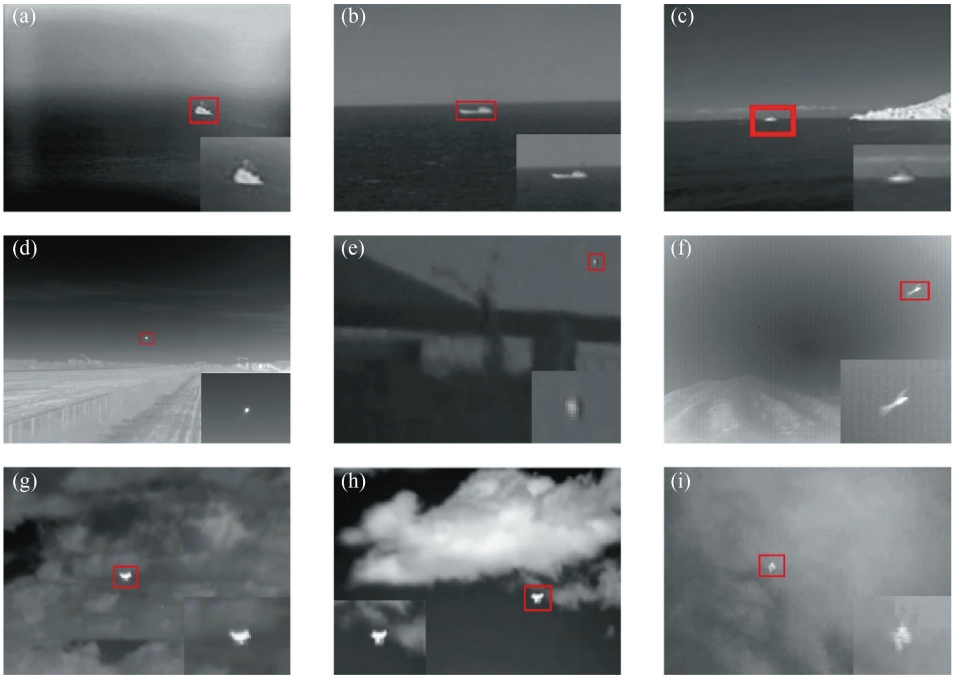


Figure 4. Representative images in the SIRST dataset. (a), (b) and (c) are all sea-level scenes. (a) contains slight noise. (b) has obvious sea horizon. (c) the target is small and has a lot of clutter. (d), (e) and (f) contain the information of the ground scene. The target brightness in (e) is low. (g), (h) and (i) are the sky background. The cloud cover in (g) and (i) is large, and the brightness of the clouds is dark. In (h), the area of the cloud is small, but the brightness is larger, which is not very different from the small bright target.

image, generate a saliency map of the original image, and highlight the target information. The generated saliency map is shown in Figure 6. It can be seen from Figure 6: (1) The convex part in Figure 6(b) is exactly the position corresponding to the target in the image, which proves that the target is more prominent relative to the neighborhood background. (2) The saliency map obtained after processing by the SR module suppresses the background information, makes the target more prominent, which verifies the effectiveness of the SR module, and lays the foundation for further target detection.

The small target area in the saliency map in Figure 6(c) is larger than the real small target area, and false targets may appear. Therefore, according to the original image, a Gaussian greyscale difference image is constructed to obtain a more accurate pixel-level target area.

Figure 7 shows the Gaussian grey difference map of two images of different scenes. The target position is marked by a rectangular frame in the original image. Figure 7 shows: (1) For point and rectangular targets, the Gaussian gray-scale difference map can effectively smooth the background and suppress isolated noise to obtain accurate target positions; (2) The edge position of the Gaussian greyscale difference map is also obtained larger response value.

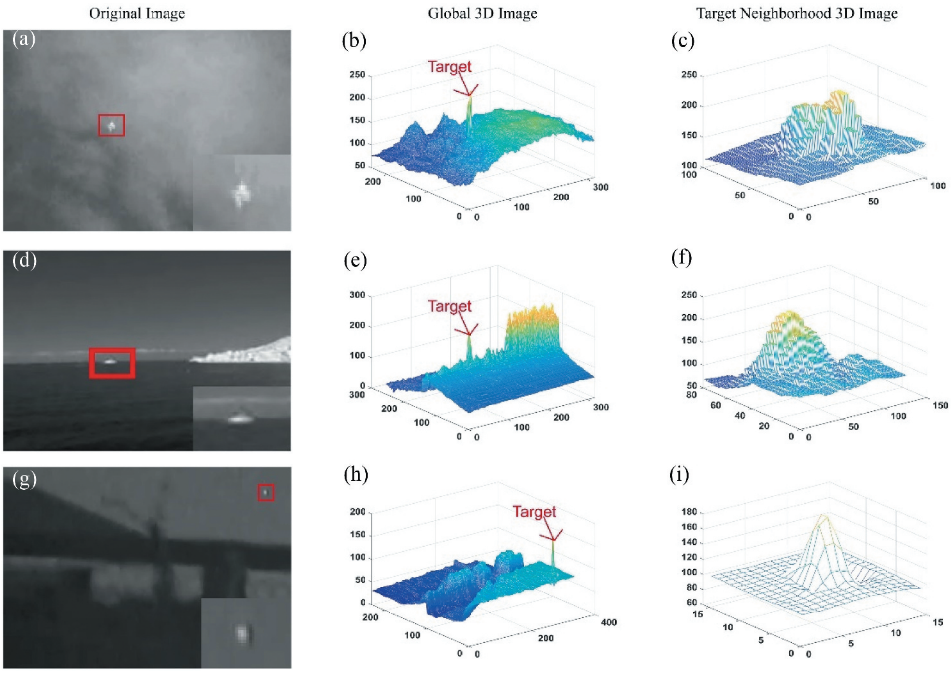


Figure 5. Global 3D image and target 3D image. (a), (d) and (g) are original images selected from the data set with representative scenes; (b), (e) and (h) are global 3D images of (a), (d) and (g) respectively. (c), (f) and (i) are the target 3D images of (a), (d) and (g) respectively.

4.2.2. Image fusion

Although the saliency map highlights the small target and suppresses the background noise, the obtained small target shape is larger than the real target shape. The Gaussian greyscale difference map obtains an accurate target area. It enhances the contrast between the target and the background, but a larger response value appears at the edge of the image. Therefore, the two can complement each other. The threshold value of the saliency map and the Gaussian grayscale difference map is calculated by the Otsu method. According to the threshold, a binary image of the saliency map and the Gaussian grayscale difference map is generated. The two binary images are fused by matrix dot multiplication to obtain the response map. Only small target information is left in the response map, so that the target can be accurately located.

4.2.3. Target detection

The grey value of the target area in the response map is close to 255. The gray value of the neighbourhood background area, image edges, and isolated noise is minimal and approaches zero. The peak area in the three-dimensional diagram of the response map is the dim and small target position. Therefore, an adaptive threshold segmentation method is adopted to segment the response map to obtain the target detection result. Adaptive threshold segmentation is $T = \bar{I} + k\sigma$, where T is an adaptive threshold, $\bar{I}$ is mean of response map, σ is the standard deviation of response map, k is constant. According to experience, k set to 7. As shown in Figure 8, the fusion map and the

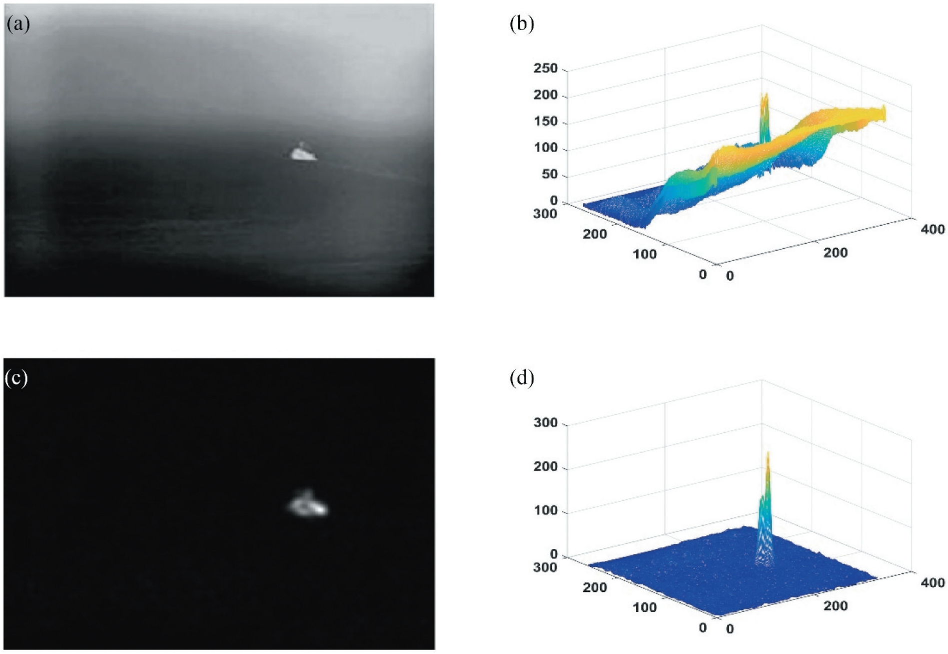


Figure 6. SR model result image. (a) Original image; (b) Three-dimensional diagram of (a); (c) Salient map of (a); (d) Three-dimensional diagram of (c).

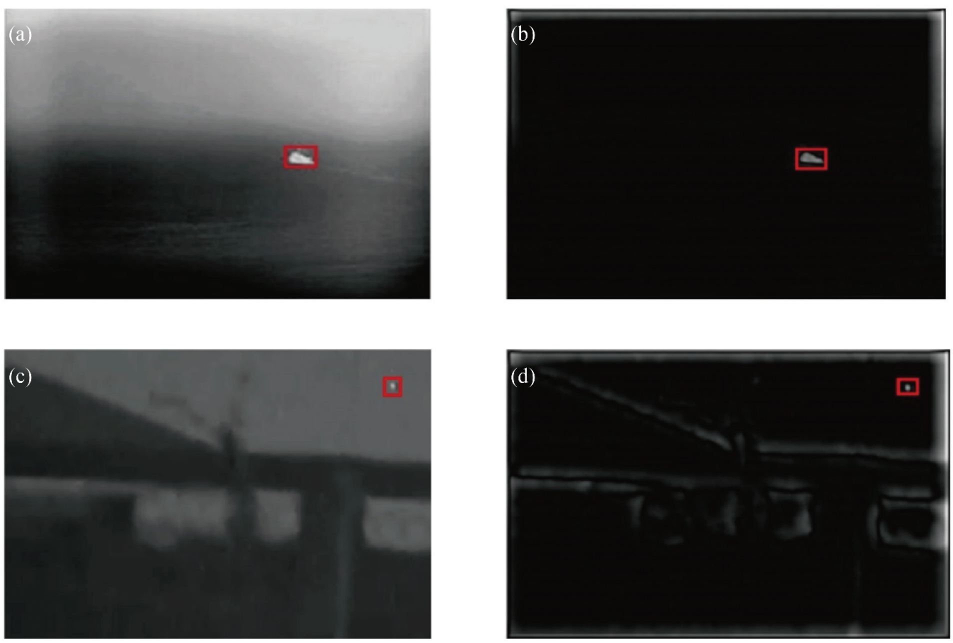


Figure 7. The result image of the Gaussian grey difference model. (a) Original image; (b) Gaussian greyscale difference image of (a); (c) Original image; (d) Gaussian grayscale difference image of (c).

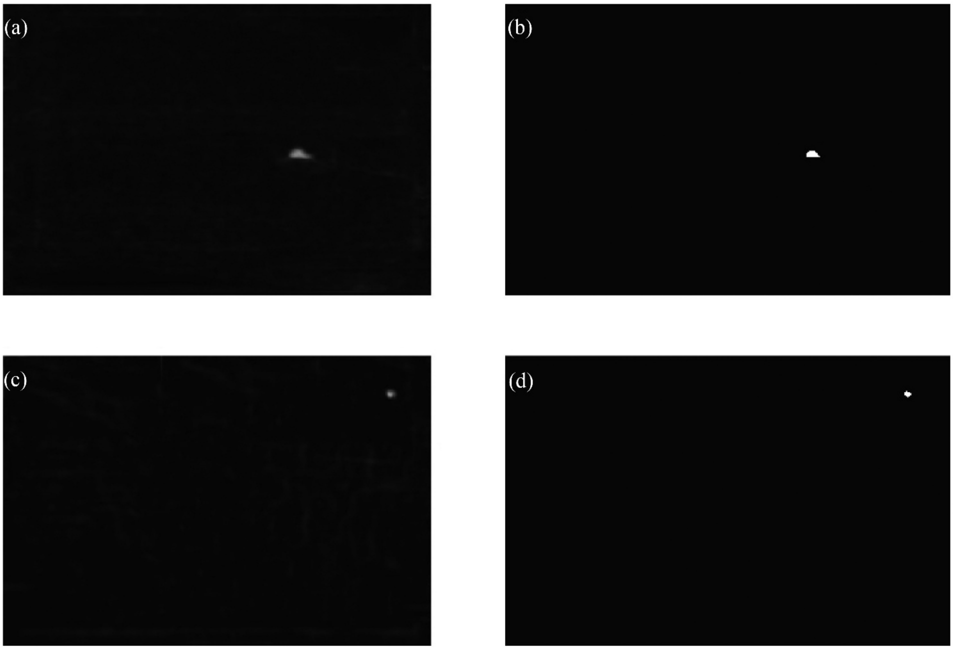


Figure 8. Image fusion and target detection example images. (a) the fusion map of the SR module and Gaussian greyscale difference module in Figure 7(a); (b) the target detection result map of (a); (c) the fusion map of the SR module and Gaussian grayscale difference module in Figure 7(c); (d) the target detection result map of (c).

detection result map of Figure 7(a,c) are displayed. The fusion map only retains the small target information, and the background information is completely filtered out, which shows a good background suppression ability.

4.3. Analysis of simulation experiment results

In order to evaluate the feasibility and effectiveness of the proposed algorithm, we use infrared small target images of various scenes for experiments. At the same time, our experimental results are compared with classical traditional detection algorithms (Top-Hat, Max-median), detection algorithms based on the human visual system (RLCM [Han et al. 2018], AMWLCM [Liu et al. 2018]), contrast measurement detection algorithm based on multi-scale patch (MPCM), detection algorithm based on low rank sparse matrix recovery (RIPT), detection algorithm based on the cumulative directional derivative weighted absolute mean difference (AAGD) (Aghaziyarati, Moradi, and Talebi 2019) and detection algorithm based on absolute direction average difference (ADMD) (Moradi, Moallem, and Sabahi 2018). All algorithms are implemented on MATLAB 2020b.

It can be seen from Figure 9 that the backgrounds of dim and small targets are complex and diverse and both bright small targets and small dark targets exist. After fusing the saliency map and the Gaussian greyscale difference map, the background clutter is suppressed and the small target is enhanced. At the same time, Figure 9(i,m) show that our method can be used for both bright small targets and small dark targets.

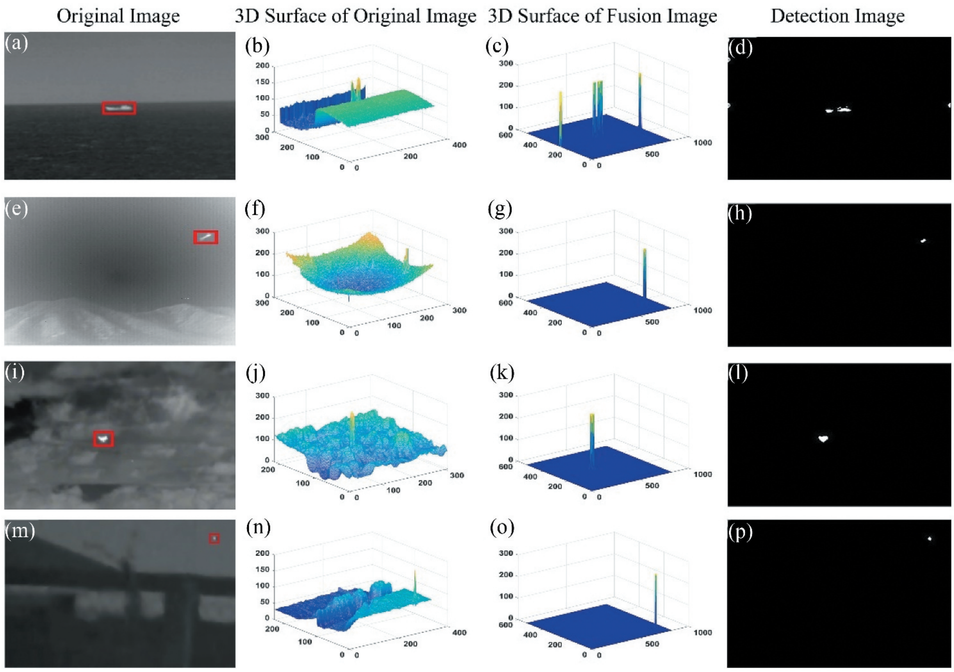
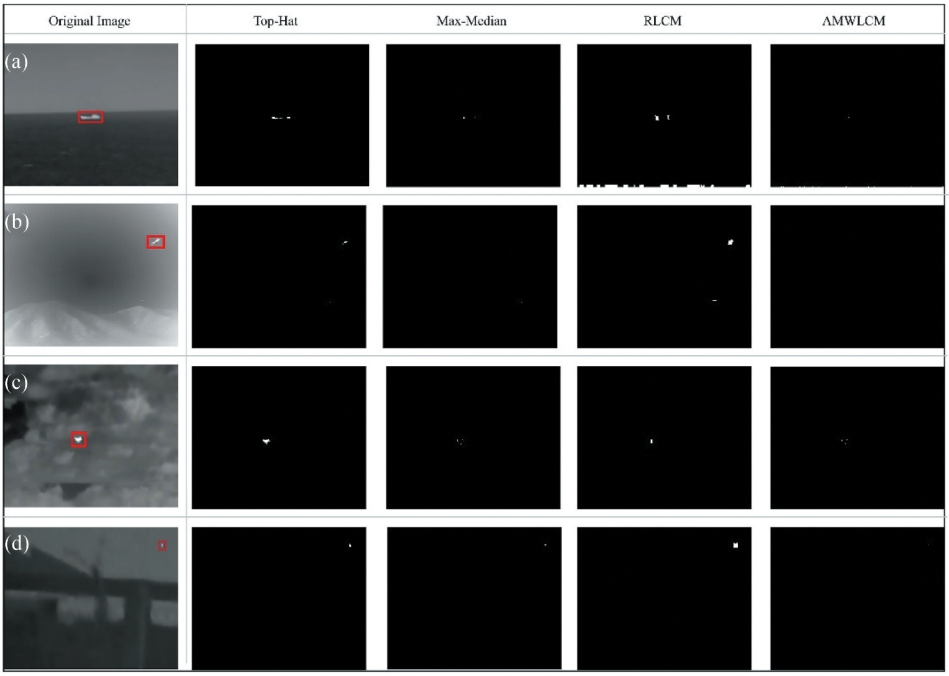


Figure 9. FDD model test results. (a), (e), (i) and (m) are the original images in the SIRST dataset. (b), (f), (j) and (n) are three-dimensional diagrams of (a), (e), (i) and (m). (c), (g), (k) and (o) are the three-dimensional diagram of the fusion map of (a), (e), (i) and (m).(d), (h), (l) and (p) are the detection results of (a), (e), (i) and (m).

Figure 10 is the detection results obtained by the other eight algorithms for the four types of infrared images in Figure 9, respectively. Figure 11 is three-dimensional diagram of the detection results of 8 comparison algorithms. According to Figures 10 and 11, we can find that: (1) The classic filtering method enhances the target. But for different targets, the enhancement effect is different, especially when the target is too large, and the detected result will easily lose the original shape of the target. There are many infrared dim and small target images that cannot be detected. (2) For detection algorithms based on the human visual system, small targets detected by RLCM have errors in shape and size, and there are false targets. In the process of suppressing the background, AMWLCM suppresses the small target and causes the detection to fail. The shape of the target after the detection of the multi-scale patch-based contrast measurement detection algorithm (MPCM) is smaller than the original shape of the target. (3) The detection algorithm based on low-rank sparse matrix recovery (RIPT) can accurately recover the target matrix, but the number of false targets is more than the other eight methods. Based on the above phenomena, it can be concluded that among the nine test methods, the proposed FDD model has achieved the most satisfactory test results.

In detecting infrared dim and small targets, the interference of complex background is the biggest problem in the detection. Severe background clutter will increase the false alarm rate of detection and cover up weak and small targets. Therefore, the background suppression factor is used to evaluate the background clutter suppression ability of the



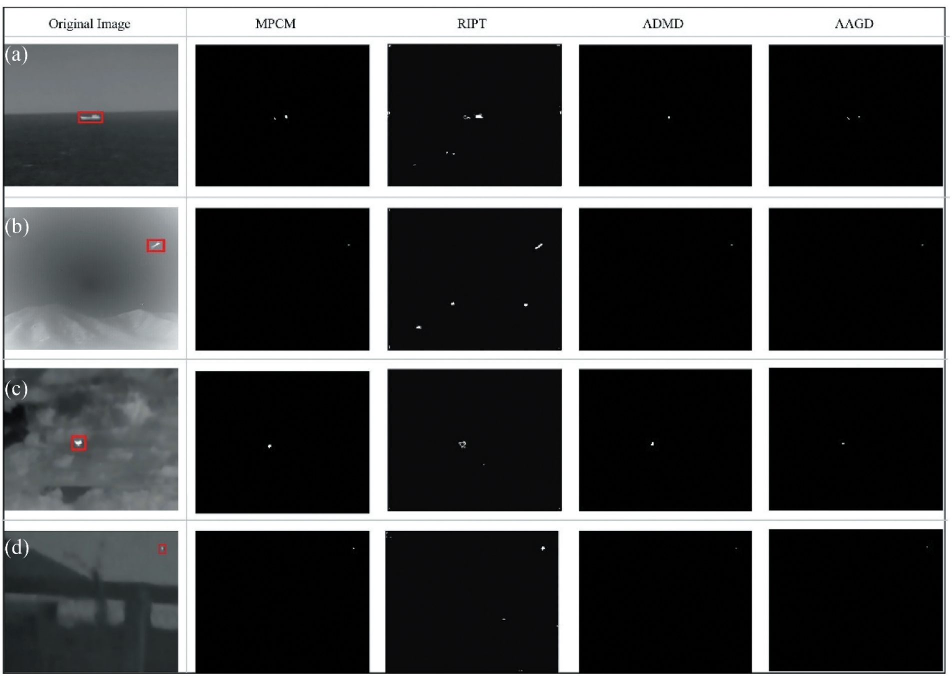


Figure 10. Comparison results of eight detection algorithms. (a), (b), (c), and (d) are (a), (e), (i), and (m) in Figure 9, and each column thereafter is the detection results corresponding to an algorithm.

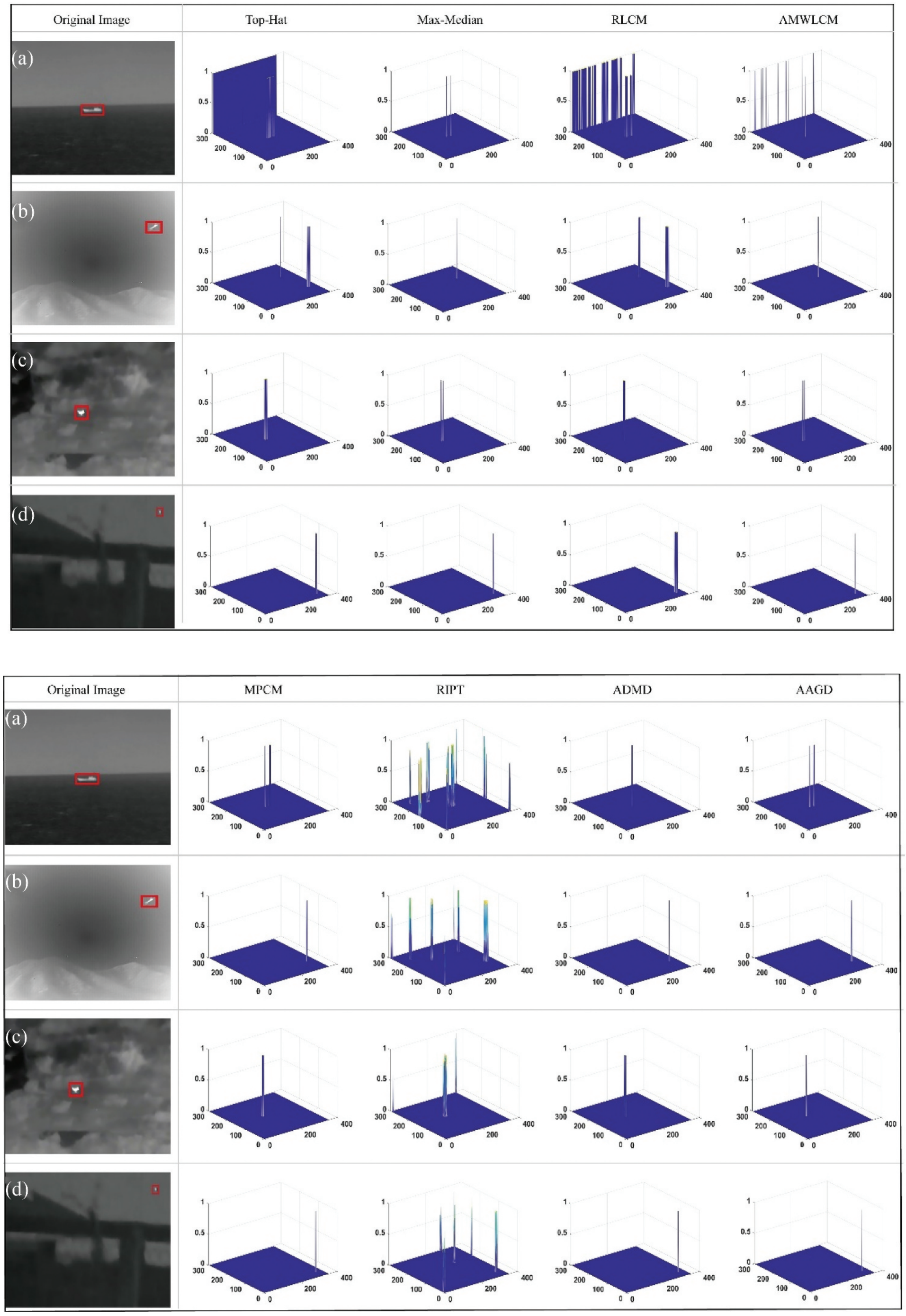


Figure 11. Three-Dimensional diagrams of the detection results of eight comparison algorithms. (a), (b), (c) and (d) are (a), (e), (i) and (m) in Figure 9. Each column is a three-dimensional diagram of the detection result corresponding to an algorithm.

Table 1. Evaluation table of the proposed algorithm and other eight algorithm detection results.

	Sea level background infrared images			Ground background infrared images			Cloud background infrared images			Dark and small target infrared images		
Method	LSNRG	SCRG	BSF	LSNRG	SCRG	BSF	LSNRG	SCRG	BSF	LSNRG	SCRG	BSF
Top-Hat	1.93	13.20	5.62	3.65	8.55	5.54	1.86	5.41	2.97	2.46	1.72	0.32
Max-Median	0.88	1.85	2.21	0.98	1.56	2.49	0.99	1.15	1.24	0.75	0.63	0.32
RLCM	8.94	98.64	49.60	1.04	2.45	1.29	11.27	15.96	9.87	0.99	0.57	0.26
AMWLCM	1.64	5.15	2.53	1.27	2.66	3.27	2.29	4.96	3.03	2.47	1.06	0.60
MPCM	4.40	49.22	21.88	16.03	75.39	42.05	75.36	202.59	111.31	10.43	6.79	3.77
RIPT	1.95	21.29	16.32	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**
ADMD	6.62	76.11	34.07	6.92	25.56	13.29	66.18	199.35	126.49	12.14	6.27	3.72
AAGD	94.29	1000.86	411.53	31.31	174.47	100.18	**Inf**	**Inf**	**Inf**	37.93	48.74	24.76
OURS	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**	**Inf**

Table 2. Detection rate and false detection rate of different algorithms.

	Top-Hat	Max-Median	RLCM	AMWLCM	MPCM	RIPT	ADMD	AAD	OURS
Pd	0.973	0.973	0.928	0.955	0.964	0.955	0.964	0.946	**0.981**
Fa	1.386	1.980	**0.069**	3.079	0.297	2.376	0.178	0.148	<u>0.128</u>

infrared small target detection method. Table 1 shows the local signal-to-noise ratio gain (LSNRG), signal-to-noise ratio gain (SCRG), and background suppression factor (BSF) values of the eight test methods in Figure 10 and the method proposed in this article. The test results of each method have been normalized to ensure the accuracy of the experimental data. It can be seen from Table 1: (1) The proposed algorithm achieves optimal values on all evaluation indicators when detecting images of different scenes; (2) Although some images are processed by RIPT, the target neighbourhood is completely reduced to zero. The RIPT detection result reaches Inf (infinity) in the evaluation index, but the RIPT processing results are insufficient, the shape and contour information of the detected small targets are lost, and there are too many false targets. Therefore, the small target detected by the proposed algorithm contains the shape and contour information of the small target and achieves excellent results under various indicators.

In order to further show the advantages of the proposed algorithm, we give the detection rate and false detection rate of each detection algorithm in Table 2. It can be seen that the proposed algorithm has the highest detection rate and the second-lowest false detection rate. This shows that our method is superior to other methods in detection performance.

4.4. Noise attack analysis

Noise is also a key factor affecting the detection results of infrared dim and small targets. An infrared imaging system mainly includes three parts: optical-mechanical structure, infrared detector, and electronic system. However, the final infrared image may contain noise generated in these various parts. Therefore, we simulate the possible noise of each part by adding different percentages of salt and pepper noise and Gaussian noise. Then evaluate the detection performance of the proposed algorithm. Figure 12 shows the infrared image with different percentages of Gaussian noise. Figure 13 shows the infrared images with different percentages of salt and pepper noise.

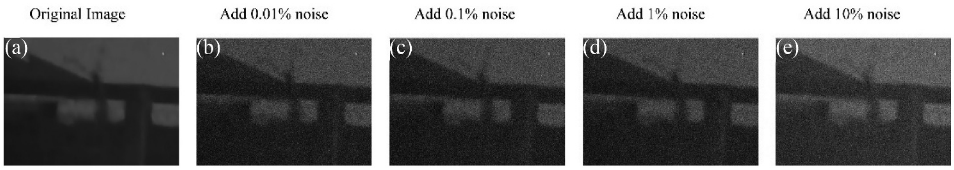


Figure 12. Infrared image after adding different proportions of Gaussian noise. (a)original image. (b) add 0.01% Gaussian noise. (c) Add 0.1% Gaussian noise. (d) Add 1% Gaussian noise. (e) Add 10% Gaussian noise.

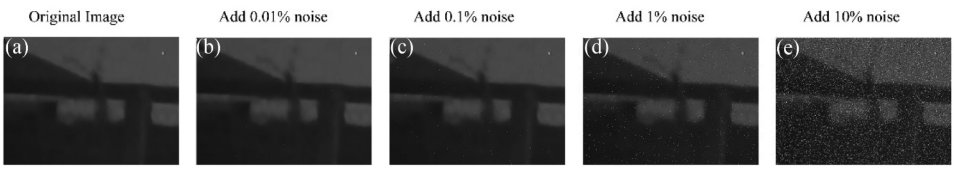


Figure 13. Infrared image after adding different proportions of salt and pepper noise. (a) Original image. (b) Add 0.01% salt and pepper noise. (c) Add 0.1% salt and pepper noise. (d) Add 1% salt and pepper noise. (e) Add 10% salt and pepper noise.

Figures 14–17 show the detection rate and false detection rate of the proposed algorithm and other algorithms with different percentages of salt and pepper noise and Gaussian noise. We add Gaussian noise and salt and pepper noise to the image at 10%, 1%, 0.1%, and 0.01%, respectively. The Gaussian grey-level difference module in the proposed algorithm calculates the average value of the neighbouring pixels and the Gaussian convolution value and subtracts the two to obtain the Gaussian greyscale difference map. It can better suppress the interference of noise in the detection of small targets. After adding salt and pepper noise or Gaussian noise, we find that: (1) The detection rate of traditional filtering-based methods has been low, and the false detection rate is high. In other methods, when different proportions of noise are added, the fluctuations of the detection rate and the false detection rate are also great. (2) The detection rate of the proposed algorithm is the highest among all algorithms after adding different proportions of different noises, and the false detection rate is also the lowest, which is close to zero. Therefore, the proposed algorithm still has a higher detection rate than other algorithms in the presence of strong noise.

4.5. Experimental data analysis

In this paper, a tracking dedicated injection infrared dynamic scene simulation evaluation system (Chen, Li, and Cao 2021) is used to simulate multiple infrared dim and small target images for the verification of the proposed algorithm. The steps of generating the entire infrared simulation image are as follows:

(1) Select the desired target type from the air target model library. And add the selected target to the scene.

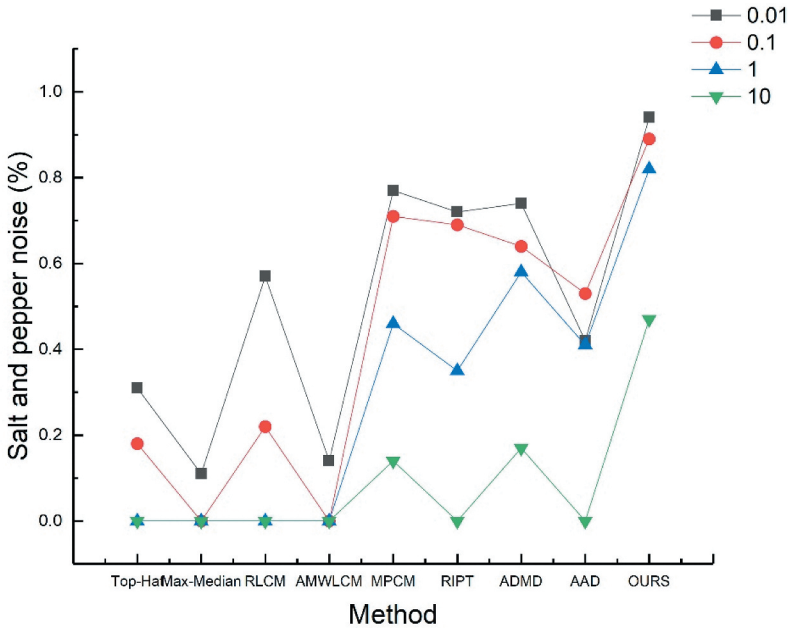


Figure 14. The detection rate of each algorithm after adding different proportions of salt and pepper noise.

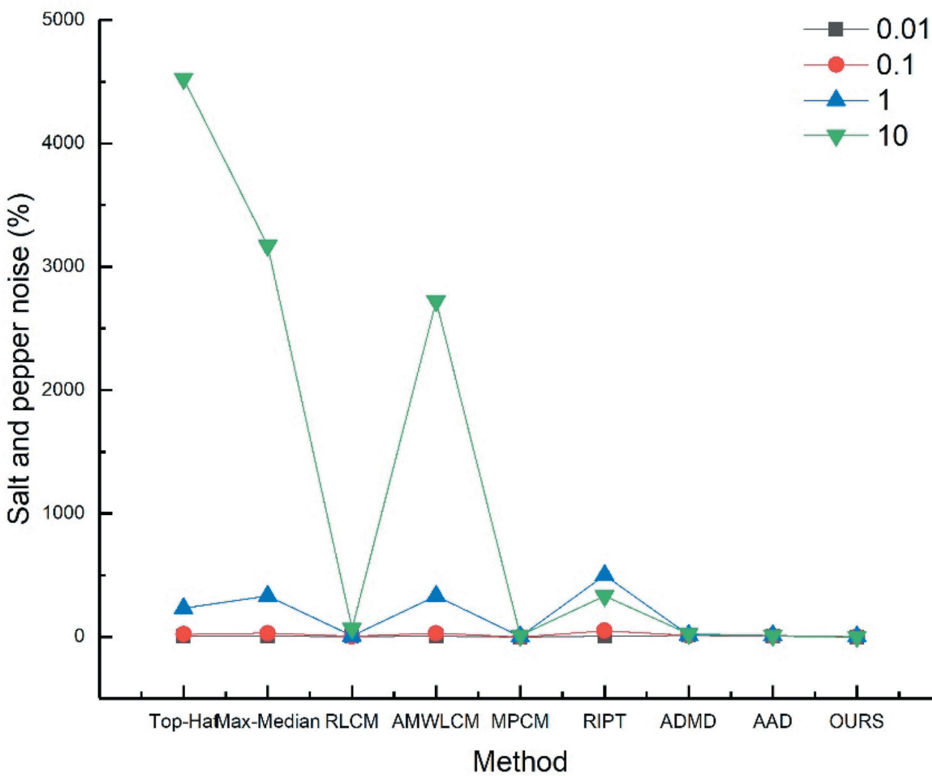


Figure 15. The false detection rate of each algorithm after adding different proportions of salt and pepper noise.

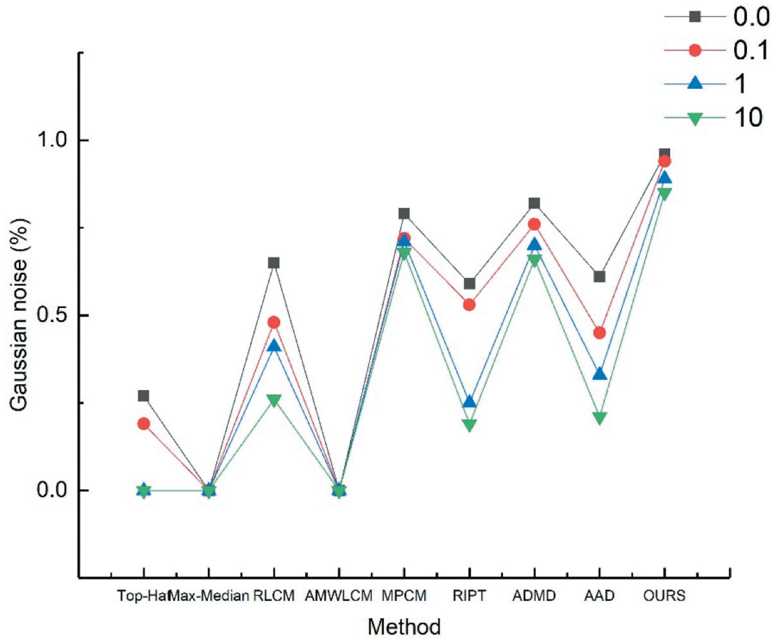


Figure 16. The detection rate of each algorithm after adding different proportions of Gaussian noise.

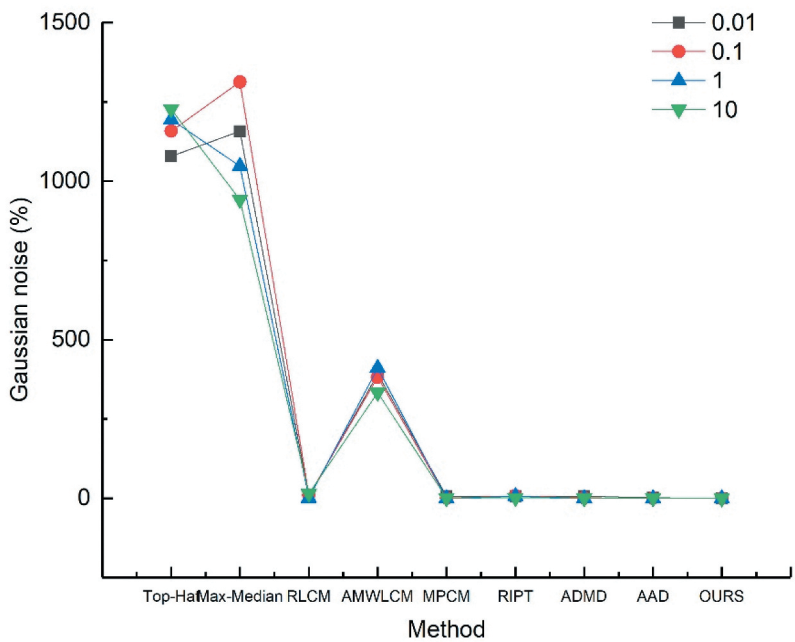


Figure 17. The false detection rate of each algorithm after adding different proportions of Gaussian noise.

(2) In the scene-setting module, load the background file. Add background data to the simulation scene.
(3) Load trajectory data to the target selected in step 1 and link the loaded trajectory to the target.
(4) Add an infrared camera to the scene and set the sight axis of the infrared camera to aim at the target selected in step 1.
(5) After all the parameters are set, the infrared dynamic scene image is generated and saved.

In the generated infrared scene images, images with clouds, cirrus clouds, thin clouds, and buildings as the background are selected for display. The detection result of the proposed algorithm is shown in Figure 18. Figure 19 show the detection results of the eight algorithms listed in Section 4.3.

According to the comparison of Figure 19, we find that: (1) The small target area in the simulated infrared small target image is clearer than the infrared small target image in other data sets. (2) When the background does not contain too complex building information, as shown in Figure 18(a,i,m), the nine algorithms, including the proposed algorithm, can accurately detect the small target location. However, the proposed algorithm better presents the shape of small targets and suppresses various clutter in the original image. (3) When the background contains complex building information, as shown in Figure 18(e), the eight comparison algorithms

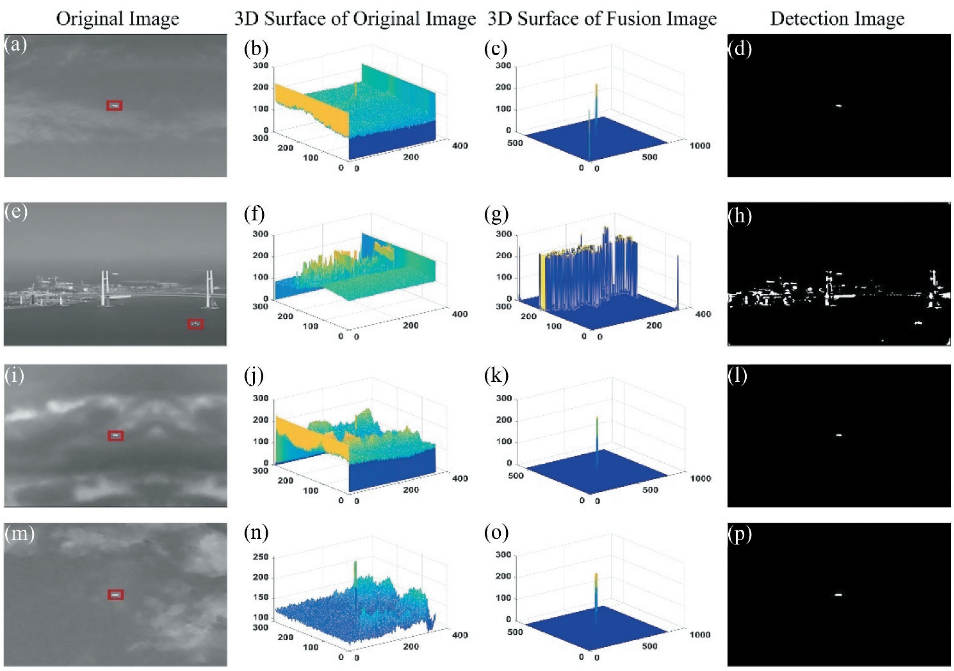


Figure 18. FDD model test results. (a), (e), (i) and (m) are the simulated infrared dim and small target image. (b), (f), (j) and (n) are three-dimensional diagrams of (a), (e), (i) and (m). (c), (g), (k) and (o) are the three-dimensional diagram of the fusion map of (a), (e), (i) and (m). (d), (h), (l) and (p) are the detection results of (a), (e), (i) and (m).

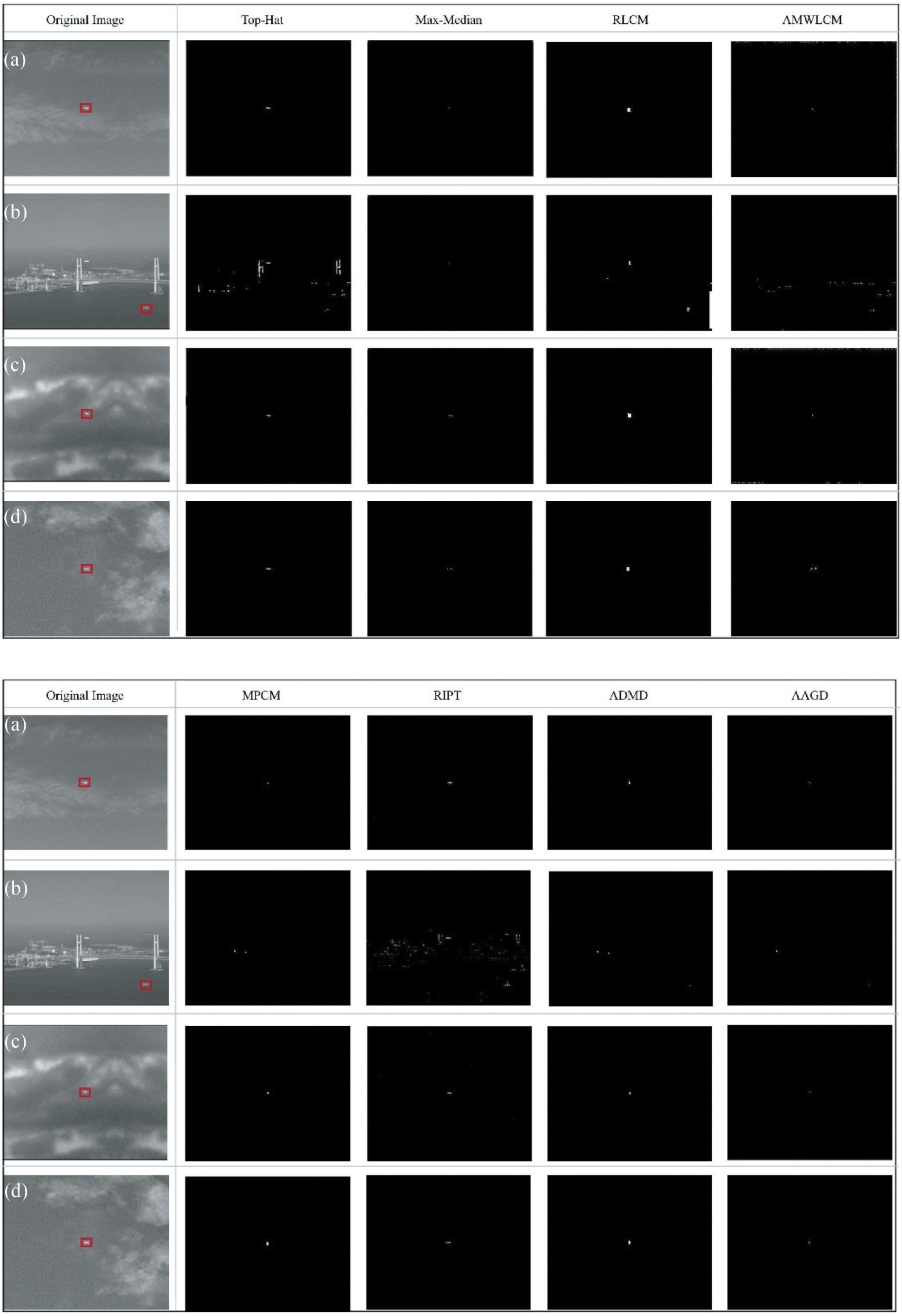


Figure 19. Comparison results of 8 detection algorithms. (a), (b), (c), and (d) are (a), (e), (i), and (m) in Figure 18, and each column thereafter is the detection results corresponding to an algorithm.

basically cannot detect the specific location of the small target, but the proposed algorithm accurately locates the location of the small target and detects the shape of a small target. However, the detection results are also inadequate, including too much background building information. Based on the above findings, it can be concluded that the 9 test methods have good detection results on the simulated infrared image data, and the FDD model proposed in this paper has the most satisfactory detection results on this data.

5. Conclusion

Aiming at the shortcomings of low target detection rate and high false alarm rate based on the human visual system method in complex scenes, this paper proposes an infrared dim and small target detection method based on frequency domain difference. The saliency map is constructed using the spectral residual module, and the greyscale difference map of the original image is calculated using the Gaussian difference bandpass filter. The two are merged into the final response map to detect dim and small targets. The detailed experimental results show that compared with the existing infrared small target detection methods, the proposed algorithm is robust to various scenes and can enhance the target and suppress background interference. The realization process is also simple, and the calculation speed is fast so that the infrared small target can be detected efficiently.

Disclosure statement

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

Funding

This research was funded by Natural Science Foundation of Shanghai [Grant No. 18ZR1425800], and the National Natural Science Foundation of China [Grant No. 61775140, 61875125].

ORCID

Haiqing Miao http://orcid.org/0000-0001-7627-687X

Data availability statement

Data underlying the results presented in this paper are available in Ref. [21].

References

Aghaziyarati, S., S. Moradi, and H. Talebi. 2019. "Small Infrared Target Detection Using Absolute Average Difference Weighted by Cumulative Directional Derivatives." *Infrared Physics & Technology* 101: 78–87. doi:10.1016/j.infrared.2019.06.003.

Bai, X. Z. and F. G. Zhou. 2010. "Analysis of New Top-Hat Transformation and the Application for Infrared Dim Small Target Detection." *Pattern Recognition* 43 (6): 2145–2156. doi:10.1016/j.patcog.2009.12.023.

Chen, C., H. Li, Y. Wei, T. Xia, and Y. Y. Tang. 2013. "A Local Contrast Method for Small Infrared Target Detection." *IEEE Transactions on Geoscience & Remote Sensing* 52 (1): 574–581. doi:10.1109/TGRS.2013.2242477.

Chen, J., Y. Li, and L. H. Cao. 2021. "Research on Region Selection Super Resolution Restoration Algorithm Based on Infrared Micro-Scanning Optical Imaging Model." *Scientific Reports* 11: 1. doi:10.1038/s41598-021-82119-1.

Dai, Y. M., Y. Q. Wu, and Y. Song. 2016. "Infrared Small Target and Background Separation via Column-Wise Weighted Robust Principal Component Analysis." *Infrared Physics and Technology* 77: 421–430. doi:10.1016/j.infrared.2016.06.021.

Dai, Y. M. and Y. Q. Wu. 2017. "Reweighted Infrared Patch-Tensor Model with Both Nonlocal and Local Priors for Single-Frame Small Target Detection." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10 (8): 3752–3767. doi:10.1109/JSTARS.2017.2700023.

Dai, Y. M., Y. Q. Wu, F. Zhou, and K. Barnard. 2020. "Asymmetric Contextual Modulation for Infrared Small Target Detection." Paper presented at the IEEE Winter Conference on Applications of Computer Vision. doi:10.48550/arXiv.2009.14530.

Deng, H., X. P. Sun, M. L. Liu, C. H. Ye, and X. Zhou. 2016. "Small Infrared Target Detection Based on Weighted Local Difference Measure." *IEEE Transactions on Geoscience & Remote Sensing* 54 (7): 4204–4214. doi:10.1109/TGRS.2016.2538295.

Deng, H., X. P. Sun, M. L. Liu, C. H. Ye, and X. Zhou. 2017. "Entropy-Based Window Selection for Detecting Dim and Small Infrared Targets." *Pattern Recognition: The Journal of the Pattern Recognition Society* 61: 66–77. doi:10.1016/j.patcog.2016.07.036.

Deshpande, S. D., H. E. Meng, V. Ronda, and P. Chan. 1999. "Max-Mean and Max-Median Filters for Detection of Small-Targets." In *Proceedings of SPIE-The International Society for Optical Engineering*, Vol. 3809, 74–83. doi:10.1117/12.364049.

Gao, C. Q., D. Y. Meng, Y. Yang, Y. T. Wang, X. F. Zhou, and A. G. Hauptmann. 2013. "Infrared Patch-Image Model for Small Target Detection in a Single Image." *IEEE Transactions on Image Processing* 22 (12): 4996–5009. doi:10.1109/TIP.2013.2281420.

Goodall, T. R., A. C. Bovik, and N. G. Paulter. 2016. "Tasking on Natural Statistics of Infrared Images." *IEEE Transactions on Image Processing* 25 (1): 65–79. doi:10.1109/TIP.2015.2496289.

Han, J. H., K. Liang, B. Zhou, X. Y. Zhu, J. Zhao, and L. L. Zhao. 2018. "Infrared Small Target Detection Utilizing the Multiscale Relative Local Contrast Measure." *IEEE Geoscience and Remote Sensing Letters* 15: 612–616. doi:10.1109/LGRS.2018.2790909.

Hou, X. D. and L. Q. Zhang. 2007. "Saliency Detection: A Spectral Residual Approach." Paper presented at the IEEE Conference on Computer Vision & Pattern Recognition. doi:10.1109/CVPR.2007.383267.

Jakubowicz, J., S. Lefebvre, and E. Moulines. 2012. "Detecting Aircraft with a Low-Resolution Infrared Sensor." *IEEE Transactions on Image Processing* 21 (6): 3034–3041. doi:10.1109/TIP.2012.2186307.

Liu, J., Z. Q. He, Z. L. Chen, and L. Shao. 2018. "Tiny and Dim Infrared Target Detection Based on Weighted Local Contrast." *IEEE Geoence & Remote Sensing Letters* 15: 1780–1784. doi:10.1109/LGRS.2018.2856762.

Lu, C. Y., Z. C. Lin, and S. C. Yan. 2015. "Smoothed Low Rank and Sparse Matrix Recovery by Iteratively Reweighted Least Squares Minimization." *IEEE Transactions on Image Processing a Publication of the IEEE Signal Processing Society* 24 (2): 646–654. doi:10.1109/TIP.2014.2380155.

Moradi, S., P. Moallem, and M. F. Sabahi. 2018. "Fast and Robust Small Infrared Target Detection Using Absolute Directional Mean Difference Algorithm-ScienceDirect." *Signal Processing* 177. doi:10.1016/j.sigpro.2020.107727.

Olshausen, B. A. and D. J. Field. 1996. "Natural Image Statistics and Efficient Coding." *Network Computation in Neural Systems* 7 (2): 333–339. doi:10.1088/0954-898X_7_2_014.

Otsu, N. 2007. "A Threshold Selection Method from Gray-Level Histograms." *IEEE Transactions on Systems Man & Cybernetics* 9 (1): 62–66. doi:10.1109/TSMC.1979.4310076.

Pan, H. B., G. H. Song, L. J. Xie, and Y. Zhao. 2014. "Detection Method for Small and Dim Targets from a Time Series of Images Observed by a Space-Based Optical Detection System." *Optical Review* 21 (3): 292–297. doi:10.1007/s10043-014-0044-7.

Pang, D. D., T. Shan, W. Li, P. G. Ma, S. H. Liu, and R. Tao. 2020. "Infrared Dim and Small Target Detection Based on Greedy Bilateral Factorization in Image Sequences." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13 (99): 3394–3408. doi:10.1109/JSTARS.2020.2998822.

Wei, Y. T., X. G. You, and H. Li. 2016. "Multiscale Patch-Based Contrast Measure for Small Infrared Target Detection." *Pattern Recognition: The Journal of the Pattern Recognition Society* 58: 216–226. doi:10.1016/j.patcog.2016.04.002.

Yang, L., J. Yang, and K. Yang. 2004. "Adaptive Detection for Infrared Small Target Under Sea-Sky Complex Background." *Electronics Letters* 40 (17): 1083–1085. doi:10.1049/el:20045204.

Zhang, W., M. Y. Cong, and L. P. Wang. 2003. "Algorithms for Optical Weak Small Targets Detection and Tracking: Review." Paper presented at the International Conference on Neural Networks and Signal Processing, 643–647. doi:10.1109/ICNNSP.2003.1279357.

Zhong, M. L., F. M. Li, and S. Miao. 2014. "A Survey on Infrared Weak Small Target Detection Method." *Advanced Materials Research* 945-949: 1558–1560. www.scientific.net/AMR.945-949.1558