

Research Article

Dim and Small Target Detection Based on Local Energy Aggregation Degree of Sequence Images

Fan Xiangsuo¹ and Xu Zhiyong ²

¹*School of Electrical and Information Engineering, Guangxi University of Science and Technology, Liuzhou 545006, China*

²*Institute of Optics and Electronics Chinese Academy of Sciences, Chengdu 610209, China*

Correspondence should be addressed to Xu Zhiyong; xzy158@163.com

Received 19 August 2019; Accepted 18 November 2019; Published 13 December 2019

Academic Editor: Paramasivam Senthilkumaran

Copyright © 2019 Fan Xiangsuo and Xu Zhiyong. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to improve the detection ability of dim and small targets in dynamic scenes, this paper first proposes an anisotropic gradient background modeling method combined with spatial and temporal information and then uses the multidirectional gradient maximum of neighborhood blocks to segment the difference maps. On the basis of previous background modeling and segmentation extraction candidate targets, a dim small target detection algorithm for local energy aggregation degree of sequence images is proposed. Experiments show that compared with the traditional algorithm, this method can eliminate the interference of noise to the target and improve the detection ability of the system effectively.

1. Introduction

Dim and small target detection algorithm is a critical technique for the imaging detection system. The system mainly combines the corresponding algorithms in the field of machine vision to automatically discriminate the targets in low signal-noise rate (SNR) scenes. The detection distance of the imaging detection system depends heavily on the quality of an algorithm [1]. Since the long-distance detection target is of a small size in the image, the energy of the target signal arriving at the detector's plane is weak. The target is easily submerged in the background clutter especially when it is affected by strong noise. As a result, it is difficult to accurately detect dim small target due to the lack of effective information such as shape and texture in the image.

In recent years, many researchers at home and abroad have carried out a lot of research work on dim small targets detection algorithms. Some algorithms have achieved good detection results in engineering applications. They can be summarized into two categories: one is based on single frame filtering detection; the other is based on multiframe association detection. Specifically, the detection algorithm based on single-frame filtering mainly refers to the

preprocessing of single-frame images by means of filtering and learning classification algorithms in some signal processing fields, and acquiring corresponding candidate target points, and then use some prior knowledge (such as pretrained model, template matching, etc.) to detect the real target points from numerous candidate targets. These algorithms primarily include detection algorithms based on background modeling and detection algorithms based on machine learning. Of which, detection algorithms based on background modeling mainly are composed of traditional detection algorithms based on time domain [2], space domain [3, 4] or frequency domain filtering [5] and background modeling methods based on statistical characteristics [6, 7], etc. Detection algorithms based on machine learning mainly include detection algorithms based on visual saliency [8, 9], detection algorithms based on dictionary learning and sparse representation [10, 11], detection algorithms based on low-rank background [12, 13], and detection algorithms based on based on CNN (Convolutional Neural Network) [14–17], etc. When the target is in a low SNR scene, using merely the single-frame information to detect dim-small targets in the above single-frame filtering detection algorithm may result in a high

false alarm rate in the detection results. In this way, targets cannot be discriminated easily. So as to effectively improve the detection capability of the imaging system, most detection tracking algorithms today have adopted multi-frame detection. Detection based on multi-frame association selects real target points from masses of candidate targets making use of the association of multi-frame motion trajectories, such as Moyer et al. proposed projection transformation method [18], the method first transforms the image from the three-dimensional space into the two-dimensional space, and then uses the Hough transform to correlate the multi-frame images in the two-dimensional space to extract the suspicious target trajectory, finally energy accumulation is conducted on each suspicious target trajectory. When the target energy of a trajectory is larger than the preset threshold, the trajectory will be extracted and judged as a final target. Dong et al. proposed a pipeline filter dim small targets detection algorithm based on motion direction estimation [19]. This method estimates the moving speed and direction of the target according to the motion state of the target in multi-frame images to further eliminate the interference of noise on the target, and can effectively detect dim small targets in strong fluctuating background. Xiu put forward an modified high-order correlation method [20]. Based on the continuity of target motion in adjacent multiframe, the method first calculates the high-order correlation of many suspicious target points in the neighborhood region of the multi-frame images, and then extracts target points using the correlation degree and threshold of the target in the multi-frame images.

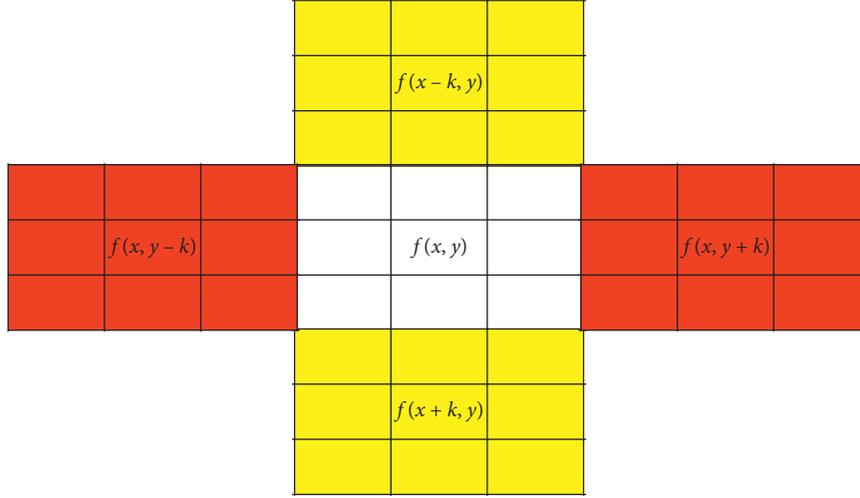
In most of the above algorithms, whether single or multiframe detection, some filtering algorithms are used to remove most of the background interference to the target in advance and then combined with the corresponding detection methods to extract small and weak targets, which provides a good research idea for this paper. Nevertheless, research deficiencies still exist in the above algorithms in the low SNR ($\text{SNR} \leq 3 \text{ dB}$) scenes with dynamic changes. For instance, some need gradually changed backgrounds or high signal-to-noise ratio and some information of the known target to achieve better detection results, some need to make a priori assumption of the model, some need to get the features of the target from the image, and then use these features to train the model in advance. It should be noted that it is practically impossible to obtain the prior information of the target, which will definitely restrict the application range of these algorithms. Furthermore, it is challenging to construct a training model due to lack of effective shapes and textures in dim-small targets. In order to enhance the capability of detecting dim-small targets under dynamic scenes, an anisotropic gradient background modeling algorithm combined with spatial and temporal information is proposed in this paper, the method compares the gradient relationship of neighborhood blocks in four directions within the spatial region of a pixel point of the t -th frame image and the block in the middle is filled by the mean of pixels in the corresponding position of two neighborhood blocks with the least

gradient difference among blocks in four direction to be the background image of the image in the t -th frame image, so that the acquired background image contains the information of the pixel point in the spatial domain. Next, the image background can be constructed by calculating the gray mean of the multiframe image of the pixel point in the continuous-time domain. At this point, the background image containing the spatial and temporal information is obtained. After that, the difference image is segmented with the multidirectional gradient maximum of neighborhood blocks. Finally, combined with the coherence of the motion of the target in the sequence image and the randomness of the noise, that is, the degree of aggregation of the target energy in the sequence image is higher than the noise, a dim small target detection algorithm based on local energy aggregation degree of sequence images is proposed, which effectively eliminates the interference of noise on the target and obtains the real target point.

2. Anisotropic Gradient Background Modeling Algorithm

The anisotropic gradient background modeling method mainly combines the imaging mechanism of target signal which can be described by point spread function, that is, the energy of a target block is spread outward from the central pixel, so that the energy in the region where the background locates is lower than that in the target area. In this case, the background can be estimated by comparing the gradient relationships between background blocks in four directions within the neighborhood range of the target block, as shown in Figure 1.

The anisotropic gradient background modeling algorithm combining spatial and temporal information is presented as follows: first, move the pixel point (x, y) of the image in the t -th frame upwards, downwards, towards left, and towards right at a certain step length k to obtain neighborhood blocks in four directions (the upper and lower blocks are presented in the yellow areas, while the left and right blocks are in the red areas in Figure 1). Second, the mean of four neighborhood blocks is calculated separately. And, the block in the middle is filled by the mean of pixels in the corresponding position of two neighborhood blocks with the least gradient difference among the gradient differences of the upper and lower blocks and the left and right blocks, which is regarded as the background image of the image in the t -th frame. In this way, the information of the pixel point in the spatial domain is contained in the background image obtained. At last, taking the accumulated frame length as M , the background for the image in the cur-th frame is estimated by calculating the gray mean of the image pixel point (x, y) in the temporal domain. The background can be estimated making full use of the information of the pixel point in its spatial and temporal domain, while the target signal is effectively reserved in the difference image. Formulas are presented as follows:

FIGURE 1: The anisotropic gradient image of the t -th frame.

$$\left\{ \begin{array}{l} F_U(x, y - k, t) = \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} f(x + m, y - k + n, t), \\ F_D(x, y + k, t) = \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} f(x + m, y + k + n, t), \\ F_L(x - k, y, t) = \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} f(x - k + m, y + n, t), \\ F_R(x + k, y, t) = \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} f(x + k + m, y + n, t), \end{array} \right. \quad (1)$$

$$F(t) = F(x + m, y + n, t) = \begin{cases} \frac{[F_U(x + m, y - k + n, t) + F_D(x + m, y + k + n, t)]}{2}, & \text{if } |F_U - F_D| \leq |F_L - F_R|, \\ \frac{[F_L(x - k + m, y + n, t) + F_R(x + k + m, y + n, t)]}{2}, & \text{else } |F_U - F_D| > |F_L - F_R|, \end{cases} \quad (2)$$

$$\left\{ \begin{array}{l} \bar{F}(\text{cur}) = \frac{1}{M} \sum_{t=\text{cur}-M}^{\text{cur}} F(t), \\ \bar{f}(\text{cur}) = |f(\text{cur}) - \bar{F}(\text{cur})|, \end{array} \right. \quad (3)$$

where F_U , F_D , F_L , and F_R represent the mean of the upper, lower, left, and right neighborhood blocks, respectively; $F(t)$ is the background image of the t -th frame image, which is determined by the mean of the corresponding coordinates of the upper and lower neighborhood blocks or the left and right neighborhood blocks. When the gradient difference of the left and right neighborhood blocks is greater than or equal to that of the upper and lower neighborhood blocks, the background block in the middle will be filled by the mean of the corresponding coordinates of the upper and lower neighborhood blocks.

Otherwise, it will be filled by the mean of the corresponding coordinates of the left and right neighborhood blocks. By doing so, the background image at the moment t can be obtained. $f(\text{cur})$ is the original image of the current frame cur; $\bar{F}(\text{cur})$ is the background estimation of the cur-th frame image; $\bar{f}(\text{cur})$ is the differential image; $L * L$ is the range size of neighborhood with the value taking from $3 * 3, 5 * 5, 7 * 7, \dots$; m and n are the coordinate positions of the neighborhood $L * L$ with the value range of $m, n = (-\text{fix}(L/2), \dots, 0, \dots, -\text{fix}(L/2))$; and $\text{fix}()$ is a rounding function.

3. Multidirectional Gradient Maximum Segmentation Method of Neighborhood Blocks

According to the investigation, dim-small targets are segmented by searching the local maximum of the pixel point in [21]. This method achieves certain segmentation effect, but it only compares the gradient relationship between the central pixel and the surrounding pixel. There is insufficient utilization of local information and does not use the neighborhood block information of the pixel to construct the gradient difference. Therefore, a multidirectional gradient maxima-segmentation method based on neighborhood block is proposed in this paper as follows:

- (1) First of all, construct five direction blocks according to Figure 1, including a central block with the pixel point $f(x, y)$ as the center at the width and height of $L \times L$, an upper block with the pixel point $f(x - k, y)$ as the center at the width and height of $L \times L$, a lower block with the pixel point $f(x + k, y)$ as the center at the width and height of $L \times L$, a left block with the pixel point $f(x, y - k)$ as the center at the width and height of $L \times L$, and a right block with the pixel point $f(x, y + k)$ as the center at the width and height of $L \times L$.
- (2) Secondly, construct a multidirectional gradient of four directional blocks according to the following formula (4):

$$\left\{ \begin{array}{l} \Delta F_U = \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} \bar{f}(x+m, y+n) - \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} \bar{f}(x+m, y-k+n), \\ \Delta F_D = \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} \bar{f}(x+m, y+n) - \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} \bar{f}(x+m, y+k+n), \\ \Delta F_L = \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} \bar{f}(x+m, y+n) - \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} \bar{f}(x-k+m, y+n), \\ \Delta F_R = \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} \bar{f}(x+m, y+n) - \frac{1}{L * L} \sum_{m=-\text{fix}(L/2)}^{\text{fix}(L/2)} \sum_{n=-\text{fix}(L/2)}^{\text{fix}(L/2)} \bar{f}(x+k+m, y+n), \end{array} \right. \quad (4)$$

where $\bar{f}(\cdot)$ is the differential image; k is the step length moved by the pixel; other parameters are defined as above; ΔF_U , ΔF_D , ΔF_L , and ΔF_R represent the gradient differences between the central block and the upper, lower, left, and right neighborhood blocks.

- (3) Obtain the maximum gradient by transforming the step length based on the gradient difference obtained in step 2. If there are more than two directions are greater than a preset threshold Th in the gradient difference of four directions in the same step length, the center block will be segmented and extracted.
- (4) Make a similar judgment after transforming the step length. The step length should be transformed for three times in total. If more than half of the three step lengths sizes satisfy the judging criterion in step 3, the central block will be deemed as a candidate target point, obtaining a binary image, record as $F'(\cdot)$.

4. The Proposed Detection Algorithm

The local energy aggregation degree of sequence images is defined primarily by the motion coherence of the target in the sequence image, that is, the aggregation of the energy of

the target in the sequence image is high due to the motion coherence of the target. The aggregation of the target is measured by the mean gray scale of the target in the continuous multiframe local neighborhood, as follows:

$$\left\{ \begin{array}{l} f'(x, y, t) = F'(x, y, t) \times f(x, y, t), \\ T = \sum_{t=1}^N \text{Num}(F'(x, y, t)), \\ \text{Area}_t = \bigcup_{(x, y, t) \in \epsilon(o_x, o_y, t)} F'(o_x, o_y, t), \\ E_{\text{Gray}} = \frac{\sum_{t=1}^N \sum_{i=-r}^r \sum_{j=-r}^r f'(x+i, y+j, t)}{\sum_{t=1}^N \text{Area}_t}, \end{array} \right. \quad (5)$$

where $(2r+1) \times (2r+1)$ is the local neighborhood; (x, y) is the coordinate of the pixel point; $f(\cdot)$ is the original image; $F'(\cdot)$ is the binary image, it is obtained by multidirectional gradient maximum segmentation algorithm; $f'(\cdot)$ is the gray image of the candidate target point, which is obtained by multiplying the original image by the binary image; t is the

image number; N is the number of frames in the continuous time domain; $\text{Num}(\cdot)$ means the number of times a candidate target point moves within its neighborhood of $(2r + 1, 2r + 1)$; T is the total number of a candidate point moving on the continuous N frame; (o_x, o_y) is the coordinate of the central point of a candidate block in the binary image; $N(o_x, o_y)$ is the neighborhood of a candidate block; $(x, y) \in N(o_x, o_y)$ is the set of coordinates of all candidate points belonging to the neighborhood $N(o_x, o_y)$; the area of the candidate block is Area_t ; E_{Gray} is the mean gray value of the candidate target points (that is the aggregation), which is obtained with a ratio of the sum of the gray values of the target in the continuous N frame and the total area.

The mean gray value of each candidate target of the continuous N frame in the neighborhood of $(2r + 1) \times (2r + 1)$ is calculated in this paper. When the area, the mean gray value, and the total number of moving of a candidate block meet the following formula, it can be judged as a real target point, as follows:

$$R(i, j) = \begin{cases} T \geq \text{Th}_1, \\ 1, & E_{\text{Gray}} \geq \text{Th}_2, \\ \text{Th}_3 < \text{Area} \leq \text{Th}_4, \\ 0, & \text{other}, \end{cases} \quad (6)$$

where Th_i ($i = 1, 2, 3, 4$) is the preset threshold.

Pseudocodes of the dim-small target detection based on local energy aggregation of the sequence image are shown in Table 1.

5. Results and Analysis

Sequence images in real scenes are used for analyzing, which are discussed from the evaluation of background modeling effect, the analysis of segmentation effect, and the evaluation of detection effect.

5.1. Background Modeling Results and Analysis. The algorithm proposed in this paper is compared with other algorithms, and contrast gain, signal-to-noise ratio gain and background inhibitor factor are used to evaluate the background modeling effect of different algorithms. The specific indicators are as follows [23]:

$$\left\{ \begin{array}{l} C = \frac{|\mu_t - \mu_b|}{|\mu_t + \mu_b|}, \\ I_C = \frac{C_{\text{out}}}{C_{\text{in}}}, \\ \text{SNR} = \frac{|\mu_t - \mu_b|}{\sigma_b}, \\ I_{\text{SNR}} = \frac{\text{SNR}_{\text{out}}}{\text{SNR}_{\text{in}}}, \\ \text{BSF} = \frac{\sigma_{\text{in}}}{\sigma_{\text{out}}} \end{array} \right. \quad (7)$$

where μ_t and μ_b are the mean of the target area and the mean of the background area, respectively; C_{in} and C_{out} are the contrasts of the original image and the difference image, respectively; I_C is the contrast gain; σ_b is the standard deviation of the background area; SNR_{in} and SNR_{out} are the SNR of the original image and the difference image, respectively; I_{SNR} is the SNR gain; σ_{in} and σ_{out} are the standard deviation of the original image and the difference image, respectively; and BSF is the background inhibitor factor.

$L \times L = 3 \times 3$ is taken. The length of the cumulative frame $M = 3$. SNR images in different scenes are selected for testing. Moreover, it also compares with TDLMS [3], Top-Hat [24], improved bilateral filtering (IBF) [25], and anisotropic filtering (AF) [26]. Specific results are shown in Tables 2–4.

Indicators in Tables 2–4 indicate that the contrast gain, SNR gain, and background inhibitor factor obtained by the algorithm proposed in this section are greater than 18, 3, and 6, respectively, in different SNR images. The reason why it has achieved a better background modeling effect than the traditional algorithm is that the proposed algorithm in this paper combines the information of the spatial domain to construct the gradient difference of the different direction blocks and fills the image with the mean of the corresponding pixels in the two neighborhood blocks with the smallest gradient difference and then obtains the image mean value in the continuous time domain as the background image. This algorithm utilizing the spatial and temporal information of the pixel blocks to construct the background image can achieve a better background modeling effect that those algorithm relying merely on temporal domain or spatial domain.

Besides, in order to demonstrate the modeling effects of different algorithms, the above five algorithms are used to model the background of the same image. At the same time, background images, difference images, and 3D images obtained by different algorithms are plotted, as shown in Figure 2. Of which, (a) is the original image (the target position is marked in red); (b) is the background image, the difference image, and the 3D image obtained by TDLMS; (c) is the background image, the difference image, and the 3D image obtained by Top-hat; (d) is the background image, the difference image, and the 3D image obtained by IBF; (e) is the background image, the difference image, and the 3D image obtained by AF; (f) is the background image, the difference image, and the 3D image obtained by the algorithm proposed in this paper.

According to Figure 2, the lightness of the background image obtained by TDLMS is unbalanced. And only signals of targets are reserved in the difference image with dim energy, as shown in Figure 2(b). Since Top-Hat filtering is susceptible to structural elements, the background image it obtained is obscure with false targets presenting in blocks in the difference image, as shown in Figure 2(c). The improved bilateral filtering reserves a great number of edge noises in the difference image, although it can construct the background effectively using the space position of the pixel and the similarity in the gray level of pixel to obtain smooth

TABLE 1: Pseudocode of dim small target detection for local energy convergence degree of sequence images.

Initialization: let k be transformed three times, $k = 3, 4, 5$; $L \times L$ is the size of target block, $L \times L = 3 \times 3$; M is the cumulative frame number, $M = 3$; Th is the threshold of gradient difference, $\text{Th} = 5$; r is the size of local neighborhood, $r = 2$; N is the number of frames in the continuous time domain, $N = 5$; Th_1 is threshold of the total number of times the target moves, $\text{Th}_1 = 3$; Th_2 is the average grayscale threshold of candidate block, $\text{Th}_2 = 10$; Th_3 and Th_4 are the lower and upper bounds of candidate block area, respectively, $\text{Th}_3 = 4$, $\text{Th}_4 = 9$.

Input: sequence original images

(1) Sequential difference images are obtained by using an anisotropic gradient background modeling method combined with spatial and temporal information, as shown in equations (1)–(3).

(2) The multidirectional gradient difference of the neighborhood block is constructed for the difference image. For details, see equation (4). According to the gradient difference, the gradient maximum value is obtained by transforming the step of k , and then the binary image is segmented and extracted.

(3) After obtaining the sequence binary images, the pipeline filtering algorithm of the literature [22] is used to obtain the centroid coordinate position (x, y) of the real target.

(4) Centering on the target centroid coordinates (x, y) , find the final target point in the local neighborhood $(2r + 1, 2r + 1)$ through equations (5) and (6).

Output: Sequence detection results

TABLE 2: Contrast gain results of different algorithms.

Method NO	SNR	TDLMS	Top-hat	IBF	AF	Proposed method
1	1.07	13.62	11.45	13.75	13.76	20.12
2	0.69	10.77	9.23	12.29	11.41	18.25
3	1.58	17.53	10.24	18.13	17.84	30.83
4	1.20	16.49	13.59	15.21	17.32	26.37
5	0.97	12.09	12.32	14.76	13.48	19.80

TABLE 3: SNR gain results of different algorithms.

Method NO	SNR	TDLMS	Top-hat	IBF	AF	Proposed method
1	1.07	2.95	1.49	3.27	3.12	5.17
2	0.69	1.37	1.16	2.45	1.78	3.25
3	1.58	3.16	2.08	4.88	3.57	10.23
4	1.20	2.51	1.78	3.79	3.46	6.87
5	0.97	2.45	1.31	2.87	2.51	4.69

TABLE 4: Background inhibitor factor results for different algorithms.

Method NO	SNR	TDLMS	Top-hat	IBF	AF	Proposed method
1	1.07	2.23	1.87	2.51	2.34	6.67
2	0.69	2.27	1.12	2.47	2.32	6.58
3	1.58	2.32	1.29	2.79	2.53	6.84
4	1.20	2.25	1.26	2.67	2.45	6.71
5	0.97	2.06	1.15	2.18	2.12	6.44

backgrounds, as shown in Figure 2(d). The anisotropic filtering algorithm only uses the mean of the four directions to be the difference filtering result and make it difficult to eliminate edge contour areas in the difference image, as shown in Figure 2(e). In terms of the proposed algorithm in this section, it constructs gradient differences of neighborhood blocks with a full utilization of spatial-temporal domain information of pixel blocks to obtain background images, which can effectively eliminate the edge contour areas in the difference image while well reserving target signals with a false rate in targets.

5.2. Segmentation Results and Analysis. Three scenes are adopted to verify the segmentation effect of different algorithms on images, which are also compared with OSTU and the algorithm proposed in [21]. Parameters in this section are set as follows: the motion $k = 3, 4, 5$; the neighborhood size $L * L = 3 \times 3$; and the preset threshold of gradient difference $\text{Th} = 5$. Simultaneously, the number of true frame (TR) and the number of failed frame (FR) in the same scene are counted, as follows:

$$\begin{cases} \text{TR} = \frac{T}{\text{TF}} \times 100\%, \\ \text{FR} = \frac{\text{TF} - T}{\text{TF}} \times 100\%, \end{cases} \quad (8)$$

where T and TF represent the number of true frame and the total number of frames of the image, respectively; TR is the rate of true segmentation frame; and FR is the rate of failed segmentation frame.

As can be observed from Table 5, a better segmenting effect can be achieved by the multidirectional gradient maximum segmentation algorithm based on neighborhood blocks in comparison to the other two algorithms. The obtained TR is greater than 95%, while the FR is less than 4.2%. The study found that transforming the step length to search the maximum value of the neighborhood block by constructing the gradient differences of neighborhood blocks in different directions can not only fully utilize the information of the neighborhood blocks but also effectively highlight the difference between the pixel blocks. Regarding the algorithm proposed in [21], as it depends merely on the differences between the pixel points and neighboring pixels, it has a less efficient segmenting effect in the unstable edge contour area with slight gradient differences, obtaining TR greater than 90% and FR less than 10%. OSTU has the worst segmentation effect. The main reason is that when faced with dynamic transformation and low signal-to-noise scenes, the algorithm is difficult to segment and extract targets because of the small gray difference between pixels.

Segmentation results of different algorithms in the three scenarios are presented in Figures 3–5. Specially, Figures 3(a)–3(d), 4(a)–4(d), and 5(a)–5(d) in order are the

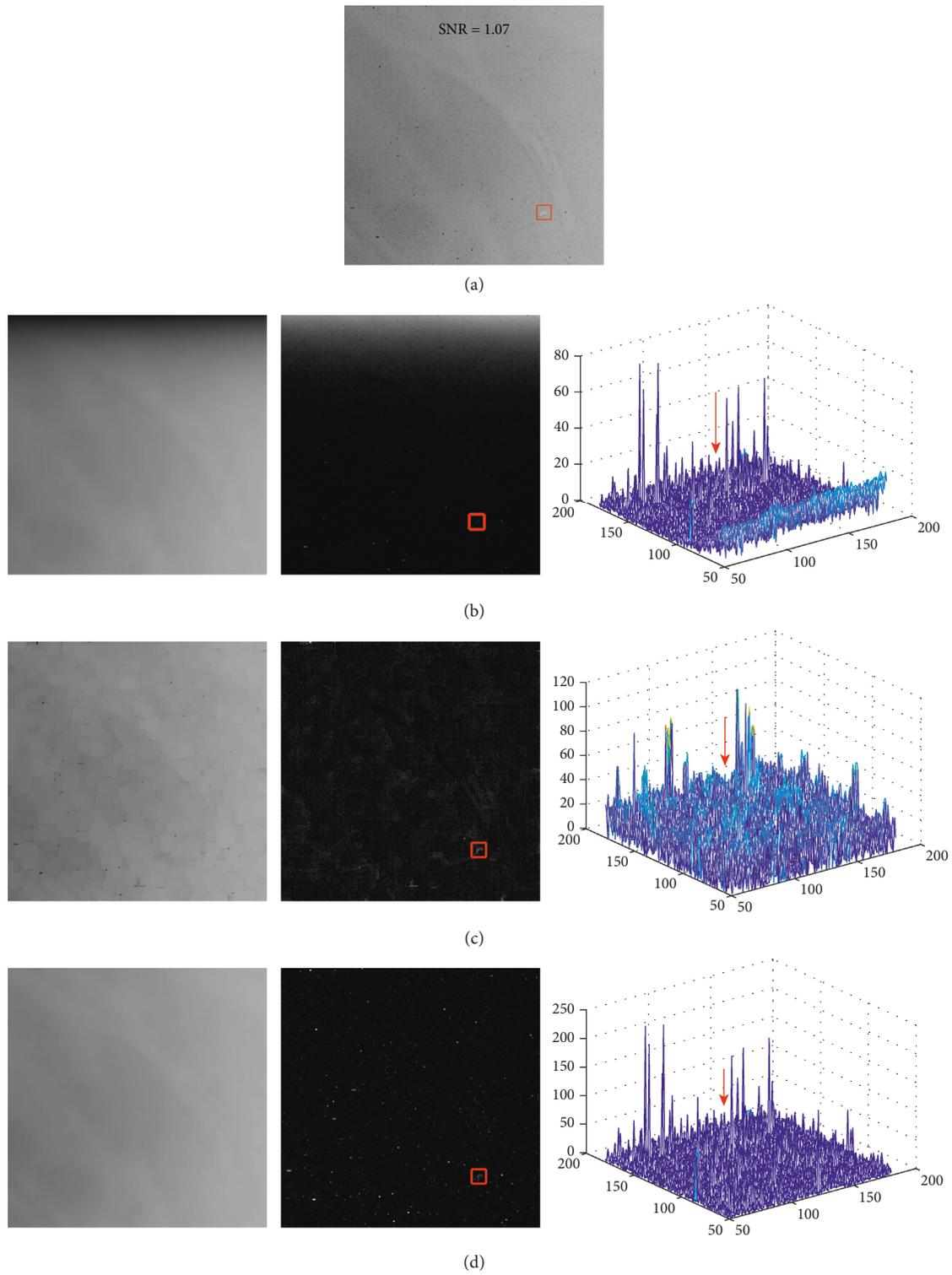


FIGURE 2: Continued.

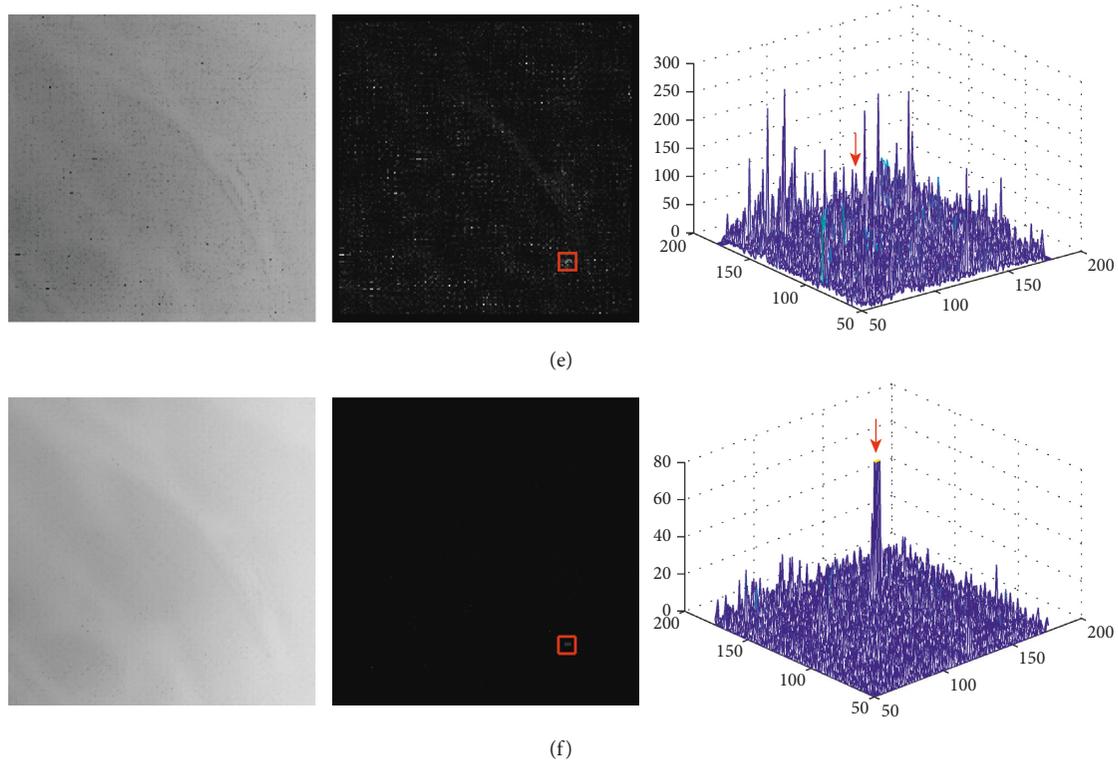


FIGURE 2: Background modeling results of different algorithms.

TABLE 5: Segmentation results of different algorithms.

Index	Method		OSTU (%)	Reference [21] (%)	Proposed method (%)
	NO	Frames			
TR	Scene 1	127	86.61	92.13	97.64
	Scene 2	120	85.00	90.00	95.83
	Scene 3	114	87.72	93.86	97.37
FR	Scene 1	127	13.39	7.87	2.36
	Scene 2	120	15.00	10.00	4.17
	Scene 3	114	12.28	6.14	2.63

original images, the OSTU segmentation results, the segmentation results of Reference [21], and the segmentation results of the algorithm proposed in this paper. It can be observed from the three scenes that there are a large number of false targets in the segmentation results obtained by the OSTU algorithm; at the same time, due to the small difference between the target gray and the background in the scene, some of the target signals are lost in the segmentation image; the algorithms obtained by [21] also lose pixels in the target area, although it obtains less false targets than the OSTU algorithm, while the algorithm proposed in this section can effectively extract target areas and obtain the least false targets.

5.3. Detection Results and Analysis. Sequence images of three real scenes are employed to verify the effectiveness of the detection algorithm proposed in this paper. The targets of the three sequence images are selected in a dynamically

changing scene. The target in the scene A move diagonally and pass through the unstable edge contour area during the motion; the target in the scene B moves from top to bottom and also passes through the unstable edge contour area during the motion, while the target in the scene C moves randomly in the center of the image and is affected by the strong background interference in the center of the image. Furthermore, the detection algorithm proposed in this paper is compared with (TDLMS) [3], Top-Hat [24], improved bilateral filtering [25], and anisotropic filtering [26]. Detection results of all these algorithms in three scenes are presented in Figures 6–8. Of which, Figures 6(a)–6(f), 7(a)–7(f), and 8(a)–8(f) in order are the original images, the TDLMS detection results, the Top-Hat detection results map, the improved bilateral filtering detection results, the anisotropic filtering detection results, and the result of the detection algorithm proposed in this paper.

As can be seen from Figures 6–8, since the three scene targets are seriously disturbed by background, the three

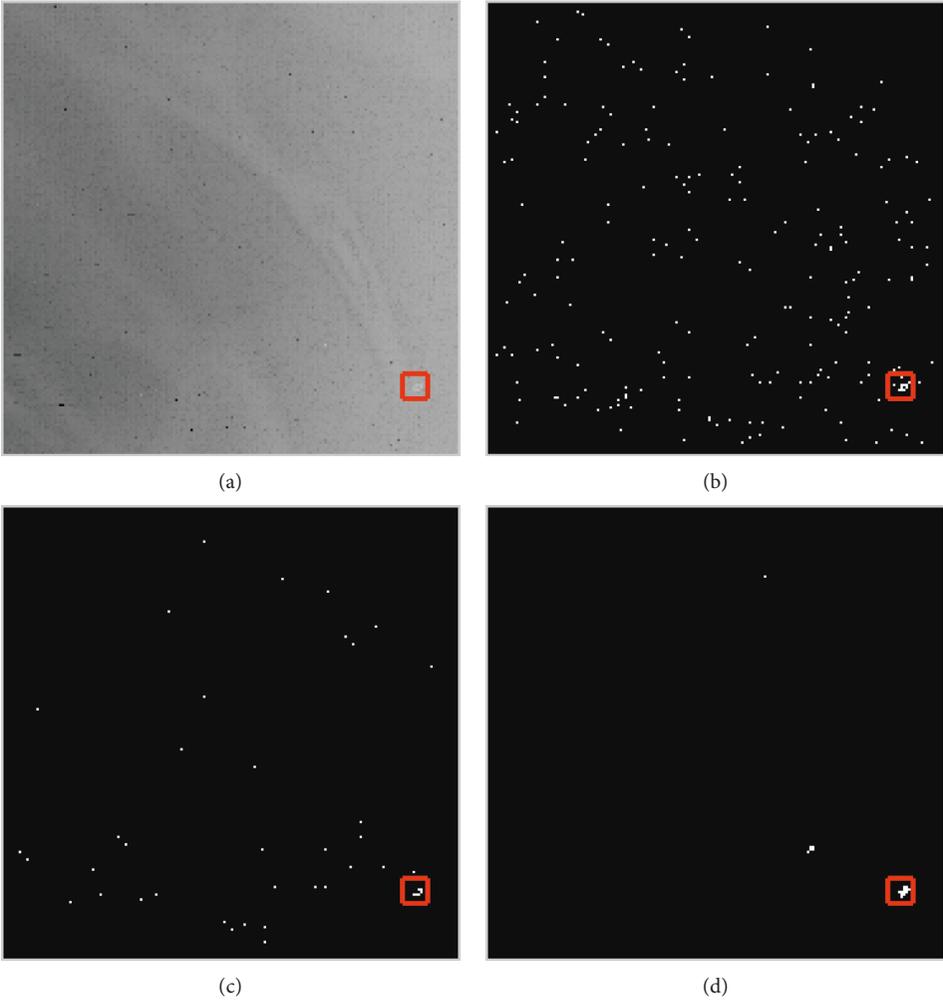


FIGURE 3: Segmentation results of different algorithms in scene 1.

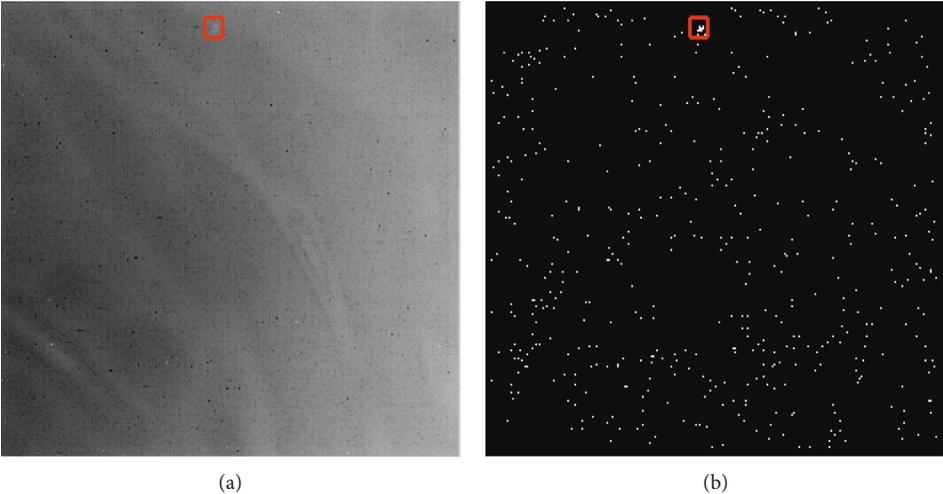


FIGURE 4: Continued.

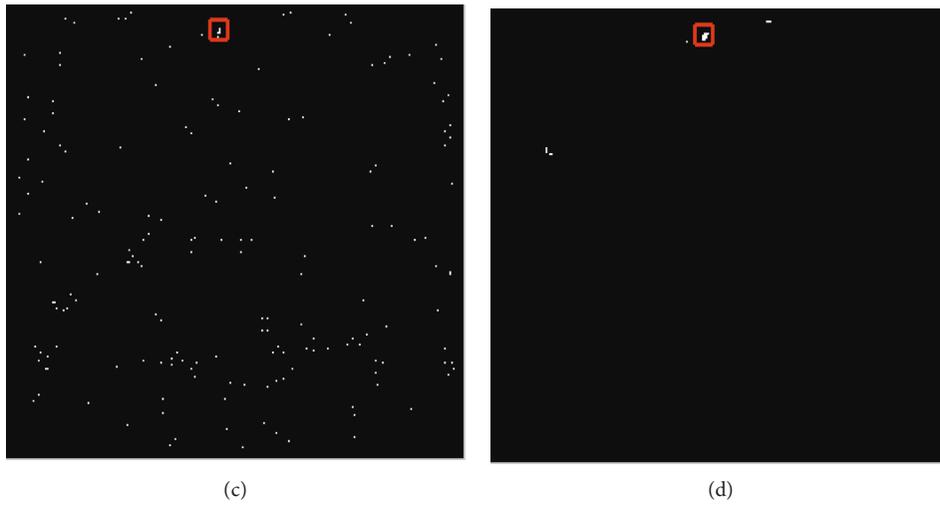


FIGURE 4: Segmentation results of different algorithms in scene 2.

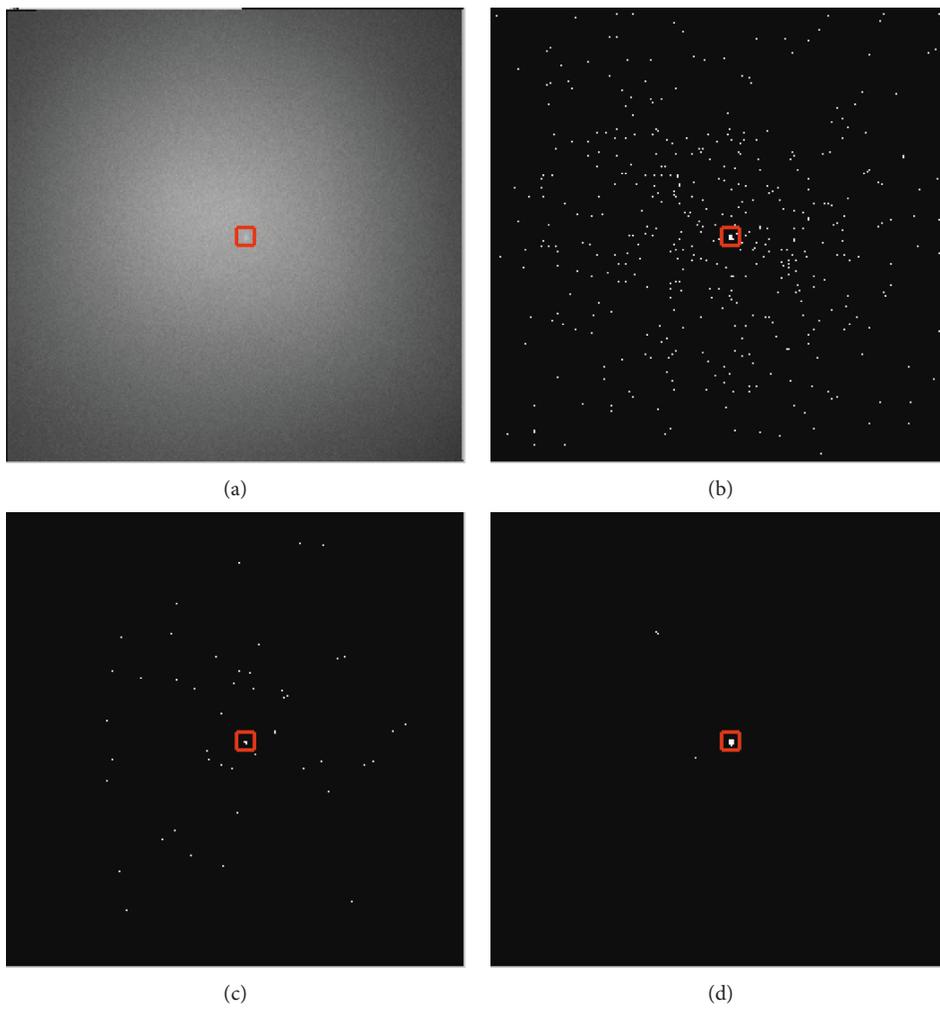


FIGURE 5: Segmentation results of different algorithms in scene 3.

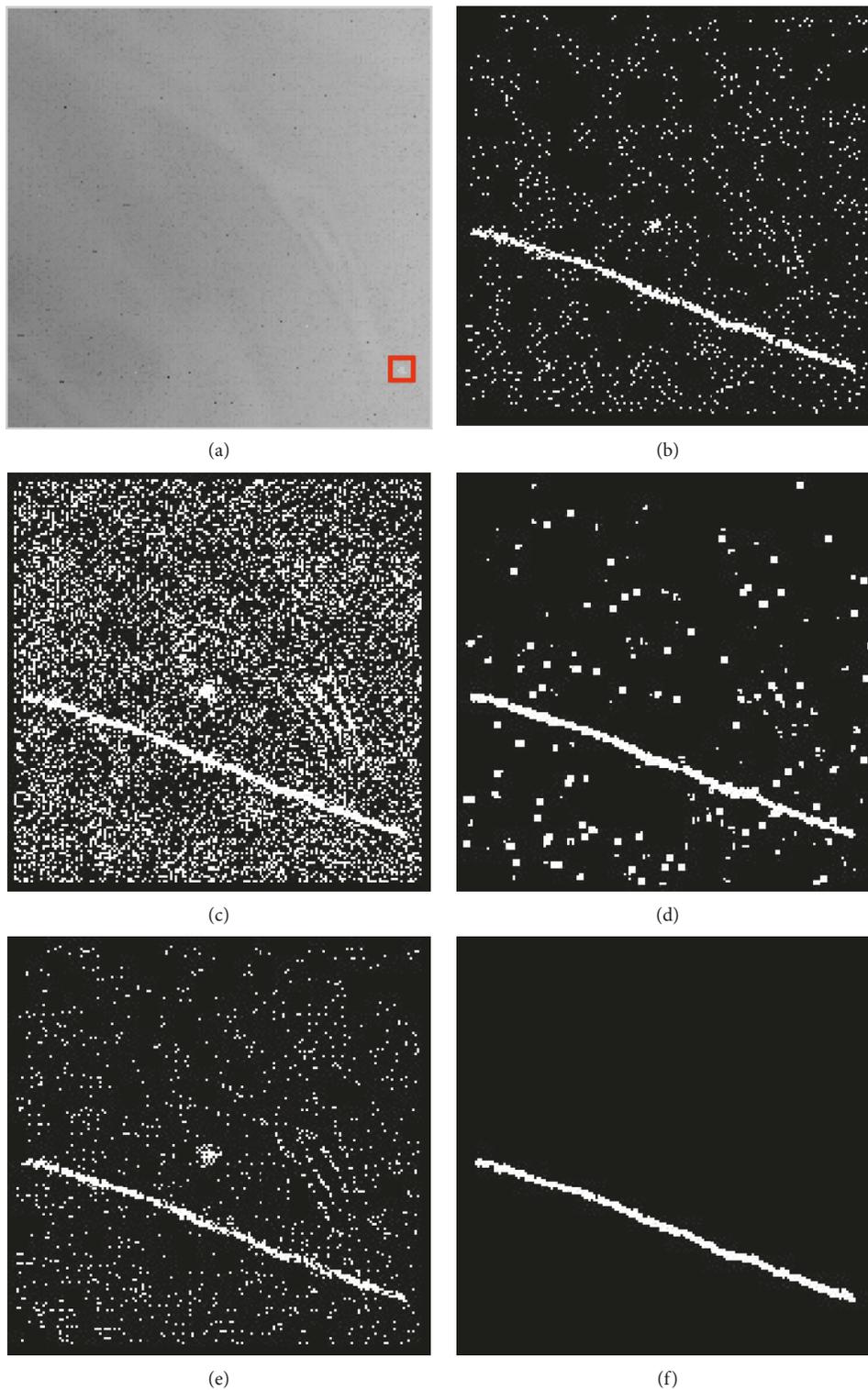


FIGURE 6: Detection results of different algorithms in scene 1.

algorithms of TDLMS, Top-Hat, and anisotropic filtering have more false targets in the detection results. On the basis of the gray characteristics and space information of the pixels, adding a filtering template to the improved bilateral filtering can effectively decrease the number of

false targets. The detection algorithm proposed in this paper can effectively eliminate background interference while well reserving target signals on the background modeling at the early stage and the multidirection gradient maximum segmentation algorithm, which can also

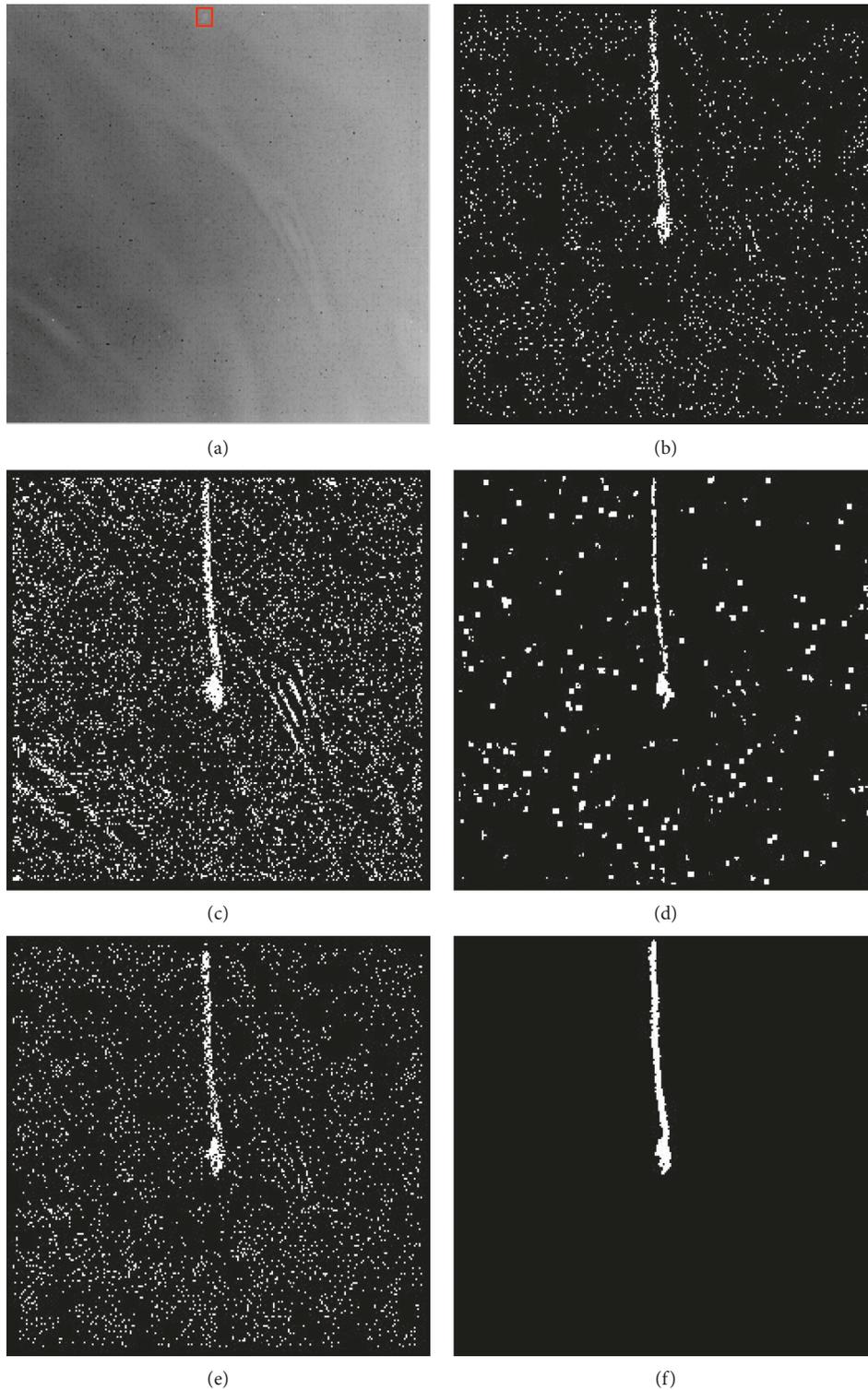


FIGURE 7: Detection results of different algorithms in scene 2.

effectively detect real target points associating with the continuous multiframe motion coherence in the neighborhood.

The detection rate (Pd) and false alarm rate (Pf) are adopted to evaluate the performances and differences of

these algorithms. Also, ROC curves of the three scenes are shown in Figure 9. From which, we can see that the detection algorithm proposed in this paper has an efficient achievement, followed by the improved bilateral filtering. Based on Figure 9(a), the Pd of the algorithm proposed in

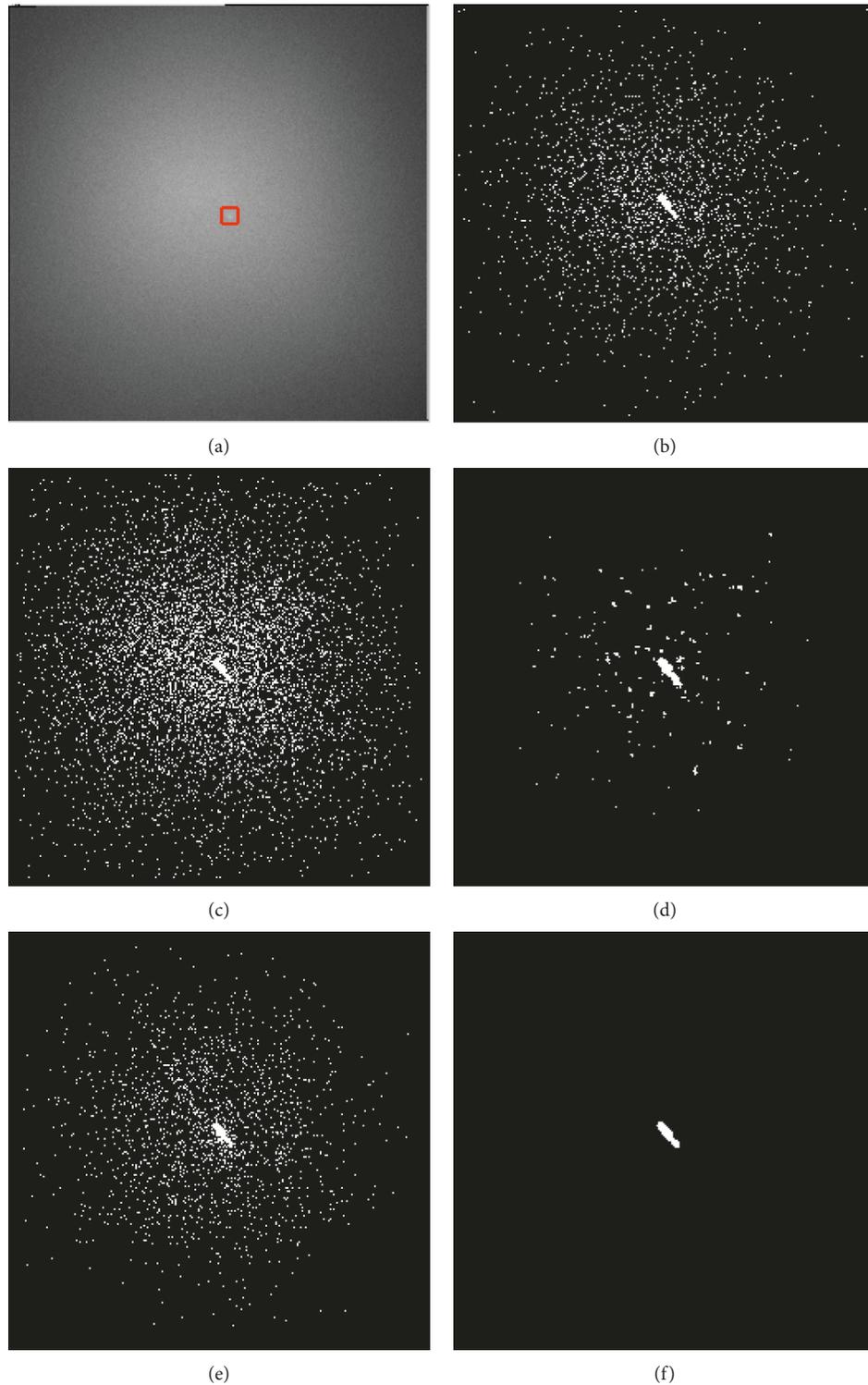


FIGURE 8: Detection results of different algorithms in scene 3.

this section is greater than 90%, while other algorithms are lower than 83% when the false alarm rate $P_f=0.01$. According to Figure 9(b), the P_d of the algorithm proposed in this section is greater than 92%, while other algorithms are lower than 85% when the false alarm rate

$P_f=0.032$. In Figure 9(c), the detection rates obtained by other detection algorithms are lower than that of the algorithm proposed in this section. To be specific, the P_d of the algorithm proposed in this section is 91%, while other algorithms are lower than 88% when $P_f=0.026$.

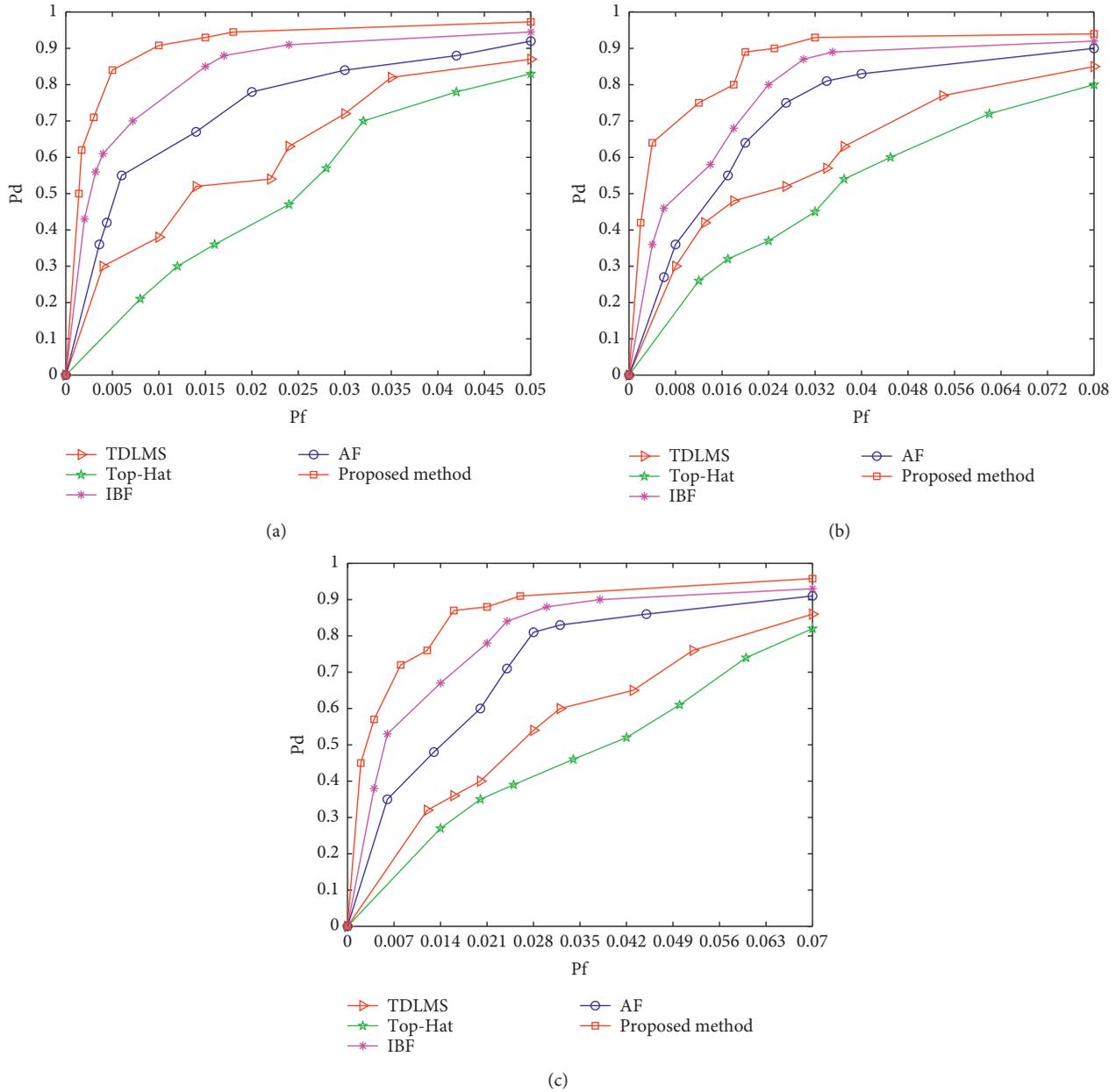


FIGURE 9: ROC curves of different scenes.

Tests show that the algorithm proposed in this section achieves a better detection result than other algorithms.

6. Conclusion

An algorithm is put forward in this paper in order to improve the capability of detecting dim-small targets under dynamic changes and low SNR scenes. Based on the imaging mechanism of point spread function, gradient differences of blocks in different directions are firstly constructed with the spatial-temporal information. Secondly, the central block is filled by the mean of pixels of the corresponding positions of two neighborhood blocks with the least gradient difference in four directions. After that, the image background is estimated through obtaining the mean of the images in the

temporal domain. On this basis, the multidirection gradient maximum segmentation algorithm is constructed through fully utilizing neighborhood blocks, which segments dim-small targets through the search of the maximum gradient of neighborhood blocks by mean of transforming step lengths. Finally, combined with the coherence of the target motion in the sequence image, the average gray of the target in the local neighborhood of the continuous multiframe is used to define the energy aggregation degree of the target, and the detection of dim and small targets is realized.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

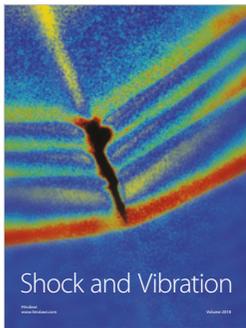
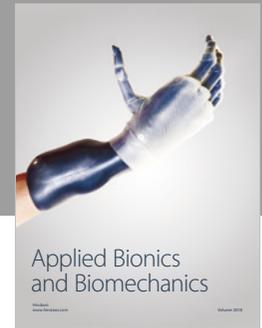
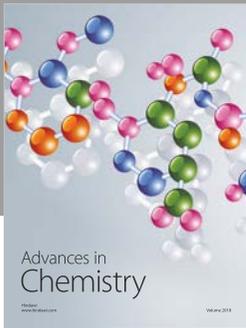
The authors declared that they have no conflicts of interest to this work.

Acknowledgments

This work was partly supported by the Guangxi Science and Technology Base and Talent Project (Acceptance No. 2019AC20147) and Doctoral Fund of the Guangxi University of Science and Technology (No. 19Z31).

References

- [1] Y. H. Wang and W. N. Liu, "Dim target enhancement algorithm for low-contrast image based on anisotropic diffusion," *Opto-Electronic Engineering*, vol. 35, no. 6, pp. 334–339, 2008.
- [2] B. Wang, *Time-Domain Detection Method for Moving Dim Small Targets in Infrared Sequences*, Xi'an University of Electronic Science and Technology, Xi'an, China, 2010.
- [3] L. Wan and M. Wang, "Infrared small target detection based on neighborhood information TDLMS filter," *Journal of Huazhong University of Science and Technology*, vol. 43, no. 1, pp. 178–181, 2015.
- [4] T.-W. Bae, "Small target detection using bilateral filter and temporal cross product in infrared images," *Infrared Physics & Technology*, vol. 54, no. 5, pp. 403–411, 2011.
- [5] C. Yang, J. Ma, S. Qi, J. Tian, S. Zheng, and X. Tian, "Directional support value of Gaussian transformation for infrared small target detection," *Applied Optics*, vol. 54, no. 9, pp. 2255–2265, 2015.
- [6] A. Sobral and A. Vacavant, "A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos," *Computer Vision and Image Understanding*, vol. 122, no. 1, pp. 4–21, 2014.
- [7] M. H. Sigari, N. Mozayani, and H. R. Pourreza, "Fuzzy running average and fuzzy background subtraction: concepts and application," *International Journal of Computer Science and Network Security*, vol. 8, no. 2, pp. 253–259, 2008.
- [8] L. Dong, B. Wang, M. Zhao, and W. Xu, "Robust infrared maritime target detection based on visual attention and spatiotemporal filtering," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 5, pp. 3037–3050, 2017.
- [9] X. Dong, X. Huang, Y. Zheng, S. Bai, and W. Xu, "A novel infrared small moving target detection method based on tracking interest points under complicated background," *Infrared Physics & Technology*, vol. 65, no. 1, pp. 36–42, 2014.
- [10] Y. He, M. Li, J. Zhang, and Q. An, "Small infrared target detection based on low-rank and sparse representation," *Infrared Physics & Technology*, vol. 68, pp. 98–109, 2015.
- [11] X. Wang, S. Shen, C. Ning, M. Xu, and X. Yan, "A sparse representation-based method for infrared dim target detection under sea-sky background," *Infrared Physics & Technology*, vol. 71, pp. 347–355, 2015.
- [12] C. Q. Gao, D. Y. Meng, Y. Yang, Y. T. Wang, X. F. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 124–138, 2013.
- [13] X. Wang, Z. Peng, D. Kong, P. Zhang, and Y. He, "Infrared dim target detection based on total variation regularization and principal component pursuit," *Image and Vision Computing*, vol. 63, no. 1, pp. 1–9, 2017.
- [14] J. Zhong, T. Lei, and G. Yao, "Robust vehicle detection in aerial images based on cascaded convolutional neural networks," *Sensors*, vol. 17, no. 12, p. 2720, 2017.
- [15] X. Yang, H. Sun, K. Fu et al., "Automatic ship detection in remote sensing images from google earth of complex scenes based on multiscale rotation dense feature pyramid networks," *Remote Sensing*, vol. 10, no. 1, pp. 132–139, 2018.
- [16] K. Kang, H. S. Li, J. Yan et al., "T-CNN: Tubelets with convolutional neural networks for object detection from videos," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 10, pp. 127–139, 2018.
- [17] K. Qian, H. Zhou, H. Qin, S. Rong, D. Zhao, and J. Du, "Guided filter and convolutional network based tracking for infrared dim moving target," *Infrared Physics & Technology*, vol. 85, no. 1, pp. 431–442, 2017.
- [18] L. R. Moyer, J. Spak, and P. Lamanna, "A multi-dimensional Hough transform-based track-before-detect technique for detecting weak targets in strong clutter backgrounds," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 4, pp. 3062–3068, 2011.
- [19] W. K. Dong, J. Q. Zhang, D. L. Liu et al., "Pipeline filtering algorithm based on motion direction estimation," *Acta Photonica Sinica*, vol. 42, no. 4, pp. 471–474, 2014.
- [20] P. S. Xiu, *Small Targets Detection and Tracking in the Sky Background*, University of Chinese Academy of Sciences, Beijing, China, 2015.
- [21] Q. Zhang, J. Cai, Q. Zhang, and R. Zhao, "Small dim infrared targets segmentation method based on local maximum," *Infrared Technology*, vol. 33, no. 1, pp. 41–44, 2011.
- [22] X. M. Zhao, S. C. Yuan, X. L. Ma, and L. L. Qi, "Infrared small target detection technique based on moving pipeline filtering," *Infrared Technology*, vol. 33, no. 5, pp. 295–297, 2009.
- [23] H. L. Qin, H. X. Zhou, S. Liu, and F. Li, "Dim and small target background suppression using bilateral filtering," *High Power Laser and Particle Beams*, vol. 21, no. 1, pp. 25–28, 2009.
- [24] F. X. Lu, Y. J. Li, X. Chen, G. Chen, and P. Rao, "Weak target detection for PM model based on Top-hat transform," *Systems Engineering and Electronics*, vol. 40, no. 7, pp. 1417–1422, 2018.
- [25] Y. Q. Zeng and Q. Chen, "Dim and small target background suppression based on improved bilateral filtering for single infrared image," *Infrared Technology*, vol. 33, no. 9, pp. 537–540, 2011.
- [26] H. X. Zhou, Y. Zhao, H. L. Qin et al., "Infrared dim and small target detection algorithm based on multi-scale anisotropic diffusion equation," *Acta Photonica Sinica*, vol. 44, no. 9, pp. 146–150, 2015.



Hindawi

Submit your manuscripts at
www.hindawi.com

