A Difference-Based Local Contrast Method for Infrared Small Target Detection Under Complex Background

KAI ZHANG[1], KE YANG[2], SHAOYI LI[ID][1], AND HAI-BAO CHEN[ID][2]

[1]Institute of Aeronautics and Astronautics, Northwestern Polytechnical University, Xi'an 710129, China
[2]Department of Micro/Nano-Electronics, Shanghai Jiao Tong University, Shanghai 200240, China

Corresponding author: Hai-Bao Chen (haibaochen@sjtu.edu.cn)

ABSTRACT Infrared small target detection under complex background is of great significance in the field of remote sensing, such as optical remote sensing, infrared precise guidance, infrared surveillance, and night navigation. Because of low-contrast and complex-background clutters of infrared images, infrared small target detection under complex background is difficult and has a serious false alarm. In this paper, a difference-based local contrast method is proposed to improve the detection performance. First, a median filtering process is used to reduce pixel-sized noises, and infrared image after filtering is divided into a series of sub-images. Then, a difference-based local contrast measure based on contrast mechanism is calculated for generating the saliency map, which can exceedingly improve the detection rate and reduce the false alarm rate. Eventually, an adaptive threshold is used to extract the target sub-image. Experiments on five real infrared small target image datasets show that the proposed method is robust and effective with great respect to detection accuracy. Our method can achieve a higher detection rate and lower false alarm rate under complex background with better performance in target enhancement and background suppression compared with the traditional algorithms.

INDEX TERMS Infrared small target, complex background, median filter, difference-based local contrast method.

I. INTRODUCTION

In recent years, infrared imaging systems have been widely used in military and civilian fields such as optical remote sensing [1], infrared precise guidance, infrared surveillance [2] and night navigation [3], [4] due to the development of infrared imaging technology. As one of the key techniques in these fields, the task of infrared small target detection takes a matter of great concern. However, infrared small targets have their own characteristics which makes the detection task quiet difficult. Different from general visible images and objects, infrared small targets are always missing texture features (such as size, shape, etc.) due to the long sensing distance. Besides, the contrast between target and background is always very low in the infrared image and the small infrared target sometimes could be submerged in kinds of

backgrounds such as cloud, sea and so on. Meanwhile, limited by the sensitivity, resolution, and systematic electronic noises of the infrared imaging device [5], the image quality will be much degraded, which will adversely affect the subsequent image processing.

Over the past few decades, a great number of researchers focus on the task of the infrared small target detection and have proposed many effective algorithms. These algorithms can be roughly classified into the following two categories: the filtering method and the detection method. The filtering method extracts the target directly from the filtered image. For example, Wang *et al.* designed the DoG filter [6], i.e, the difference of Guassians which skillfully reduces the background clutters and enhance targets by removing the low frequency components of infrared images. As well, many works have been done with the filter using mathematical morphology [7], [8]. Khan and Alam proposed a small target detection method for FLIR imagery based on morphological

The associate editor coordinating the review of this manuscript and approving it for publication was Ravibabu Mulaveesala.
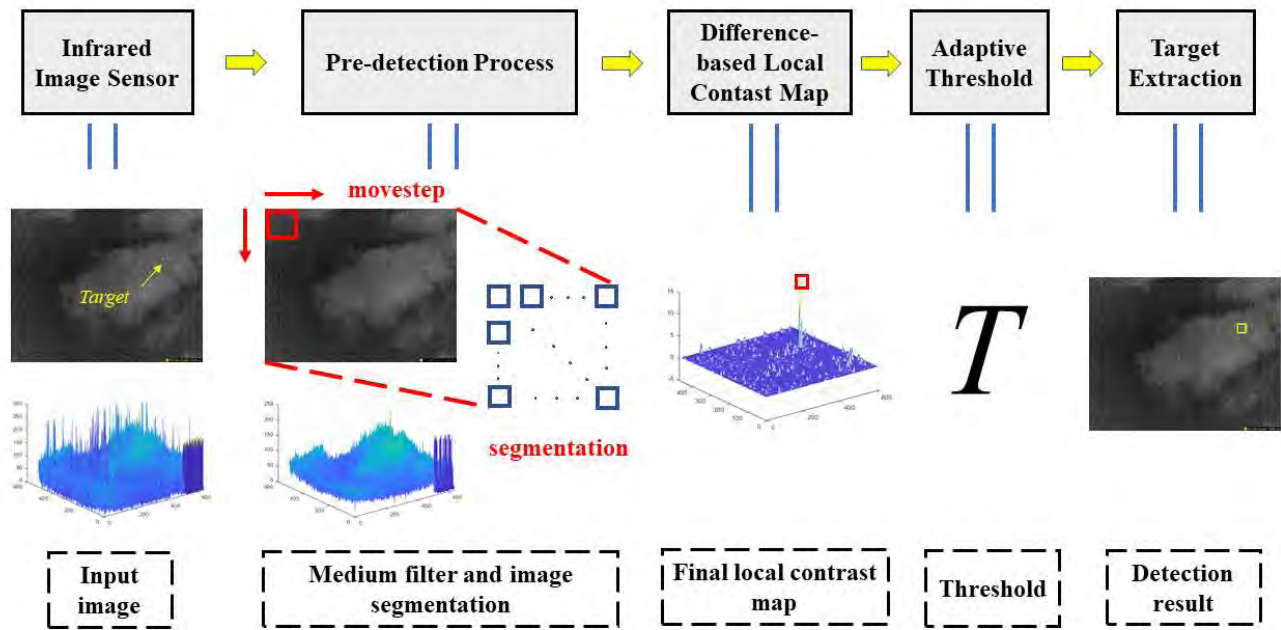
FIGURE 1. The proposed infrared small target detection system.

signal processing [8]. In the literature [7], a Top-Hat morphological filter was proposed and has been applied to the task of small target detection. While the detection method uses different techniques to compute saliency map, such as background regression [9], [10], neural network [11]–[13], *contrast mechanism*, etc. A kernel-based regression method is reported in [14] which achieves nonparametric background regression. Based on the background regression, the target component will be the difference between original image and regression background. Besides, neural network makes a breakthrough development especially in recent years and convolutional neural network (CNN) has widely applied to the field of image and video recognition [15]. The classic convolutional neural networks include ResNet [16], [17], AlexNet [18], VGG [19] and GoogleNet [20] which provide a new idea for small target detection.

Human visual system (HVS) based method has been introduced to infrared small target detection including *contrast mechanism*. *Contrast mechanism* is deeply mentioned in the local contrast method (LCM) [21]. Generally, the infrared small target is a pixel-sized image region whose brightness is higher than its neighborhood on the whole. Based on this characteristic, LCM measures every pixel's contrast relative to its neighborhood in the whole image. But LCM is time consuming because it is calculated pixel by pixel and calculation on every pixel is complicated and cyclic. However, real-time target output is much more important and sometimes even necessary in a large number of actual applications especially in the military field [22]. To overcome this weakness, Han *et al.* proposed the improved LCM (ILCM) [23] which saved much time but achieved higher detection possibility and lower false alarm. However, both of two algorithms made an

excessive process on the single pixel-sized electronic noises with high brightness, which still leads to high false alarm rate under complex background.

In this paper, a difference-based local contrast measure which combines the filtering method and the detection method is proposed to improve the detection performance including the increase of detection rate and decrease of false alarm rate. The proposed detection system is shown in Fig. 1. First, during the image pre-detection process, a median filter is used to reduce the point-sized electronic noises, and infrared image after filtering is divided into a series of sub-images with same size. Then, based on *contrast mechanism*, a difference-based local contrast map is calculated as saliency map. Eventually the target sub-image will be extracted using a threshold operation.

This paper is organized as follows. Section II introduces the image pre-detection process. In Section III-B, a difference-based local contrast measure is proposed to extract targets. Finally the experimental results and conclusion are respectively presented in Section IV and Section V.

II. PRE-DETECTION PROCESS

Infrared sensor is the key technology and core component of infrared systems. Limited by the existing manufacturing and material, as well as the electronic noises, there are always many pixel-sized noises with high brightness in the actual infrared image. Actually, these noises are much similar to the impulse noise, which can be reduced by the median operation [24]. In order to improve the robustness of our system, a median filter is used first, i.e., replace the gray of a pixel in the infrared image with the median of the pixel gray values in its neighborhood. However, median filtering operation also
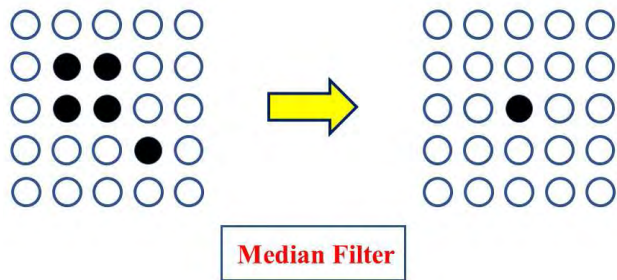
FIGURE 2. Median filter can not remove all the pixel-sized noises. Black circles represent pixels with gray 255 while white circles represent pixels with gray 0.

FIGURE 3. Difference-based LCM calculation.

destroys the details of the image which sometimes leads to the burred image. Therefore, we set the median filter template to 3×3 in order to minimize the loss of image information.

To all appearances median filtering operation can not remove all the pixel-sized noises which can be shown in Fig. 2. After the filtered image is obtained, we segment it to a serious of image sub-images based on human visual system *size-adaptation process*. Human visual system size-adaptation process means that human visual system can extract some salient regions at the scale adaptive to the target size for further process instead of image process pixel by pixel, which contributes to a fast detection speed. Target detection on a region level has been widely studied [23], [25]. A similar method with the literature [23] is used. Let a window with a fixed size of $W \times H$ slide on the image (after filtering) at a certain step from left to right and up to down, and we get a serious of sub-images and the gray average of a sub-image m can be written as follows:
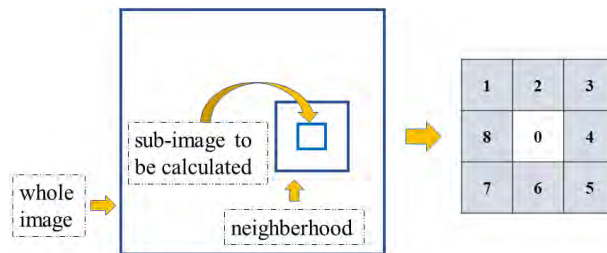
$$m(\text{sub-image}) = \frac{1}{N} \sum I(x, y), (x, y) \in \text{sub-image} \quad (1)$$

where $I(x, y)$ is the gray value of the pixel located in the coordinate (x, y) in the sub-image, and N is total number of pixels in the sub-image. It is important to note that the size of the slide window ($W \times H$) should be roughly equal with the target size in order to achieve the better performance of the next detection step the difference-based local contrast measure). Generally the infrared small target has a small image area which has less than 100 pixels [6] and usually is shaped with rectangle or circle. Based on this idea, we set the size of the slide window to 10×10, and its moving step should be no more than the slide width for overlaying entire image. Here we set the moving step to half of the window side length, i.e., the moving step is set to be 5.

III. DETECTION USING DIFFERENCE-BASED LCM
A. DIFFERENCE-BASED LCM
Contrast mechanism measures the contrast between target and background instead of brightness of target directly, which pursues to prevent highlights from being misidentified as targets. Based on the image segmentation in Section II and derived from ILCM, we construct a difference-based method

to measure the contrast which is proven to have a better performance. The details will be described below.

We first consider a specific sub-image, L_n is its gray maximum, i.e.,

$$L_n = max(I(x, y)), (x, y) \in \text{sub-image} \quad (2)$$

where (x, y) is the coordinate of a pixel and $I(x, y)$ represents its gray value. Usually a rectangle whose sides are parallel to the coordinate axis can be represented by a four-dimensional vector (x, y, w, h), where (x, y) is coordinate of the rectangle's upper left vertex; w, h are its width and height, respectively. We take the upper left corner of the image as the coordinate origin and assume that the concerned sub-image is written as (x_0, y_0, W, H). By applying an image patch whose side length is three times to the concerned sub-image's side length on the image and taking the concerned sub-image as the central sub-image, the image patch can be written as $(x_0 - W, y_0 - H, 3W, 3H)$. Furthermore, it can be seen from Fig. 3 that the image patch can be divided into nine sub-images with same size including the central concerned sub-image. By finding eight adjacent sub-images of the concerned sub-image, the difference-based local contrast measure of the concerned sub-image can be written as:

$$\text{DLCM} = \frac{L_n m_0}{max(m_i)} \frac{m_0 - max(m_i)}{255} \varepsilon(m_0 - max(m_i)) \quad (3)$$

where $m_i(i = 1, 2, \ldots, 8)$ is the average gray value of the sub-image which represents the concerned sub-image's background, m_0 is the average gray of the concerned sub-image, $\varepsilon(m_0 - max(m_i))$ is a step function. More clearly, if $m_0 \geq max(m_i)$, $\varepsilon(m_0 - max(m_i))$ is set to 1, instead if $m_0 < max(m_i)$, $\varepsilon(m_0 - max(m_i))$ is set to 0. According to the formula (3), the value of $max(m_i)$ can not be zero because of its denominator position. When the situation $max(m_i) = 0$ occurs, we directly set $max(m_i)$ to 0.001 and continue to perform the calculation.

Similar to ILCM, if the concerned sub-image contains the target, usually $m_0 > max(m_i)$, thus $m_0 - max(m_i) > 0$, $\varepsilon(m_0 - max(m_i))$ is 1; meanwhile, $L_n \geq max(m_i)$, therefore DLCM $> m_0/255$, the target will be enhanced. On the other hand, if the concerned sub-image is background, there will be $m_0 < max(m_i)$. Hence $\varepsilon(m_0 - max(m_i))$ is set to 0 directly and DLCM will be 0. Therefore, the background will be suppressed.
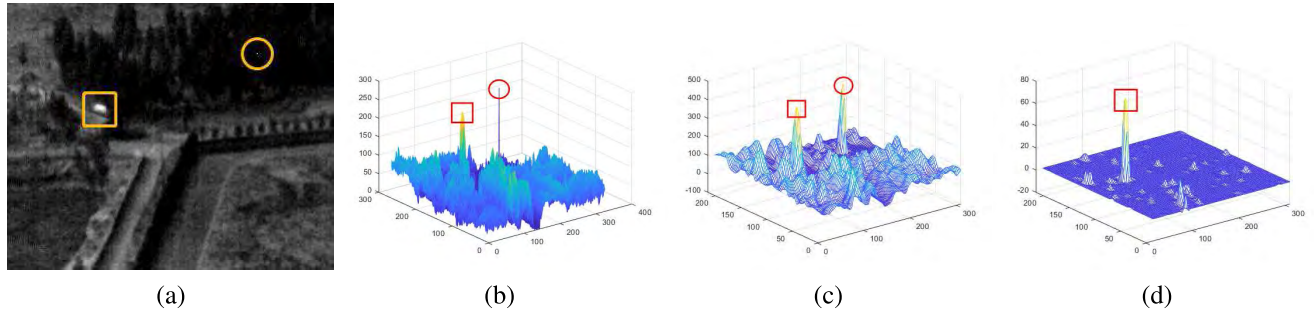
FIGURE 4. Resistance of single point-sized electronic noises with high brightness versus ILCM and DLCM. (a) The simulation image that contains one small target and one single point-sized electronic noise. (b) 3-D gray map of the simulation image. (c) The saliency map obtained by ILCM. (d) The saliency map obtained by DLCM.

According to [23], we should point out that ILCM can be calculated by the following formula:

$$ILCM = L_n \frac{m_0}{max(m_i)}, \quad i = 1, 2, \ldots, 8 \tag{4}$$

Obviously, if $m_0 \approx max(m_i)$, $m_0/max(m_i) \approx 1$, gray maximum L_n will determine high or low of the ILCM value. Let us consider a special situation, an image patch which doesn't contain target has a low gray value on the whole, generally speaking it should be seen as background. However, if a sub-image and its adjacent sub-images satisfy $m_0 \approx max(m_i)$ in this patch, a highlight with single pixel will set L_n to a high value even up to 255 which directly leads to high ILCM value of the sub-image. In other words, ILCM misunderstands the background with a noise as a target in this case. Therefore, we need to be aware of this problem and have to make targeted improvements. We construct a factor β (compare (3) and (4)) which is related to the difference of average gray between the concerned sub-image and its background, i.e. $(m_0 - max(m_i))$:

$$\beta = \frac{m_0 - max(m_i)}{255} \varepsilon(m_0 - max(m_i)) \tag{5}$$

It can be seen from (5) that the factor β always satisfies $0 \leq \beta \leq 1$. When m_0 is close to $max(m_i)$, β approaches zero and it will be responsible for reducing the value of the DLCM even if L_n takes a large value. Besides, DLCM measure gives much more significant difference between the case $m_0 > max(m_i)$ and $m_0 < max(m_i)$ than ILCM. And the ILCM method can not be calculated in the situation $max(m_i) = 0$.

In order to further prove the robust resistance of single point-sized electronic noises of the proposed method compared with ILCM, we select an infrared image that contains one small target and one single point-sized electronic noise which is shown in Fig. 4 (a) labeled in a rectangle and a circle respectively. The simulations are done by using the traditional ILCM method and the proposed method, respectively. We first remove all the noises in the image and construct artificially a pixel with high brightness located in the upper right of the simulation image under background with low gray value. The simulation results of generating saliency map are presented in Fig. 4 (c) and Fig. 4 (d). It can be seen from

these figures that ILCM measure maintains the target and the noise almost the same value while the proposed DLCM method enhances only the target and can achieve a prefect resistance nearly to zero towards the single point-sized noise.

B. TARGET EXTRACTION
We calculate the value of DLCM of every sub-image using formula (3) after image segmentation described in Section II and make up them as saliency map. Then we respectively figure out the mean μ and variance σ of the saliency map, and take them as the references setting the threshold Th:

$$Th = \mu + k\sigma \tag{6}$$

where parameter k is a constant. In our experiments, k is suitable in [20, 30]. Once the DLCM of sub-image is higher than Th, it will be predicted as target, instead it will be predicted as background. If needed, the coordinate of target (x_t, y_t) can be written as the grey centroid of the obtained target sub-image:

$$x_t = \frac{\sum_{(x,y)} x \times I(x, y)}{\sum_{(x,y)} I(x, y)} \tag{7}$$

$$y_t = \frac{\sum_{(x,y)} y \times I(x, y)}{\sum_{(x,y)} I(x, y)} \tag{8}$$

where (x, y) is one pixel located in the target sub-image.

IV. EXPERIMENTAL RESULTS AND ANALYSIS
In order to show the effectiveness of the proposed infrared small target detection system, the experiments have been performed by using five typical infrared small target image datasets against complex and noisy backgrounds. All experiments are implemented in PyCharm Community IDE by Python program with python-opencv, ImageFilter, Image and numpy libs on a PC with 8-GB RAM and 3.6-GHz Intel i7 Dual-Core CPU.

It is worth pointing out that all our five datasets used are real data captured from an infrared sensor device called IRCAM GmbH with continent-ground (see Fig. 5 (a1)),
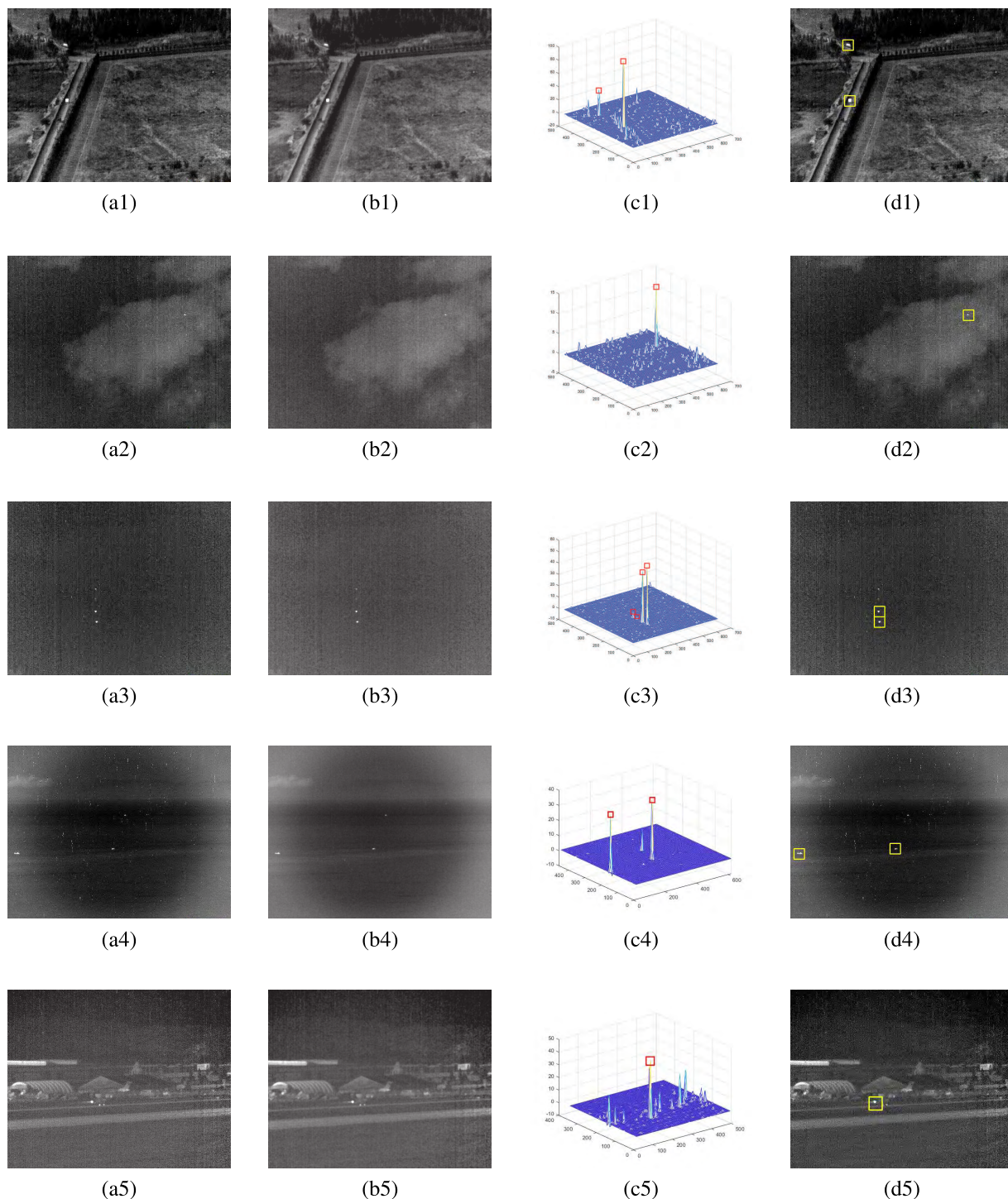
FIGURE 5. Experimental results using the proposed small infrared target detection system. The rows 1-5 are respectively corresponding with the five real small infrared target image datasets: (a1) continent-ground, (a2) sky-cloud, (a3) sky, (a4) sky-sea and (a5) ground-house. (a) The original infrared images from five datasets, respectively. (b) Images after median filtering process. (c) 3-D Saliency maps after DLCM calculation. (d) Detection results (labeled in yellow rectangles).

sky-cloud (see Fig. 5 (a2)), sky (see Fig. 5 (a3)), sky-sea (see Fig. 5 (a4)) and ground-house (see Fig. 5 (a5)) representing five kinds of complex backgrounds. The details are shown in Table 1.

Because of the limitation of materials and techniques of the infrared sensor device such as photon fluctuations of infrared background radiation, processing circuit additional noises and so on, all our five kinds of infrared images are rich
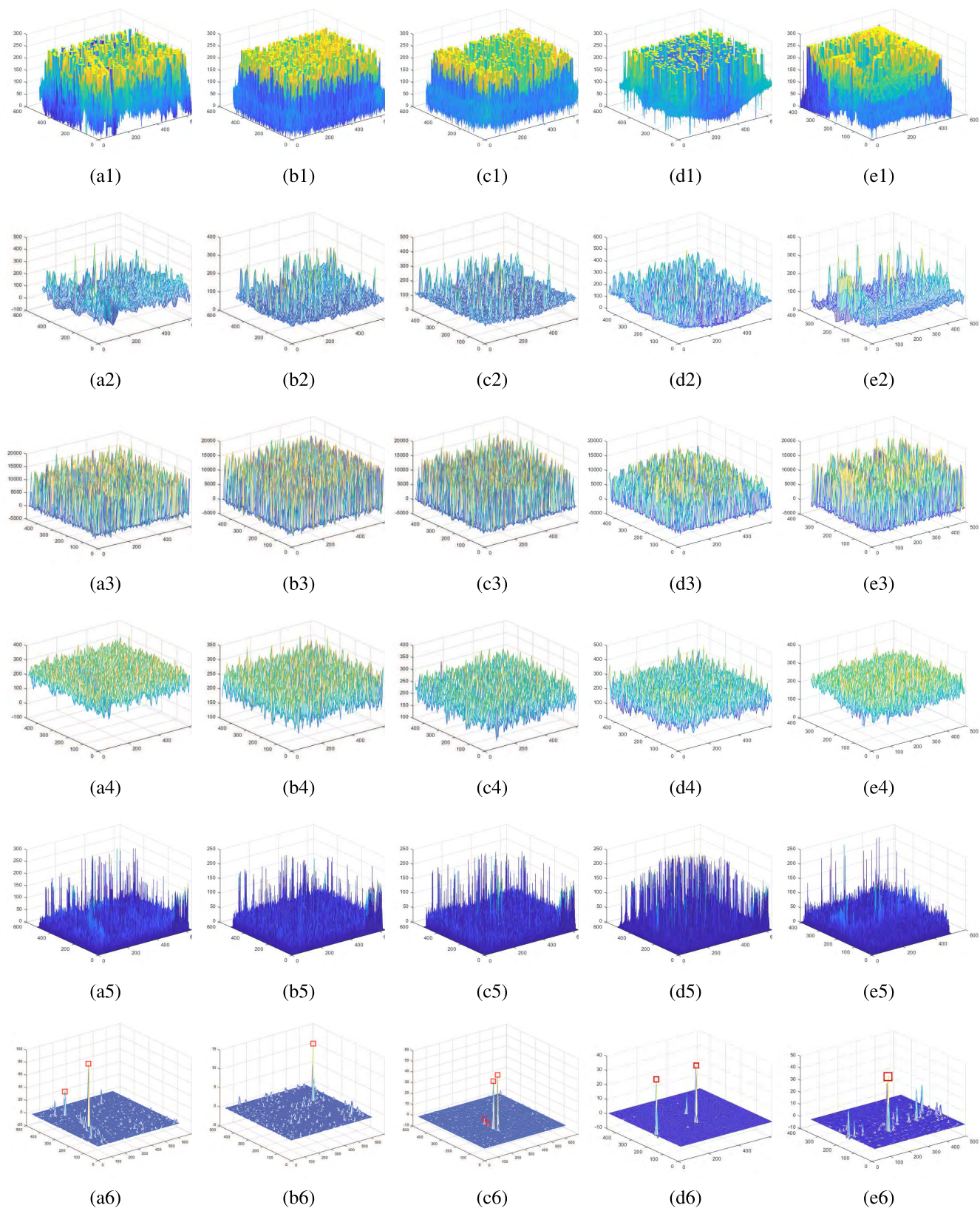
FIGURE 6. 3-D distributions of saliency map obtained through different detection methods. The rows (1) LCM, (2) ILCM, (3) DoG, (4) DoG-ILCM, (5) Kernel regression, (6) Proposed method. The columns (a)-(e) are corresponding with the datasets 1-5 respectively.

TABLE 1. Five real small infrared target image datasets captured from an infrared sensor device called IRCAM GmbH with different complex and noisy backgrounds in our experiments.

Datasets	Image Resolution	Number of images	Total number of small targets	Image description
Dataset 1 (see Fig. 5 (a1))	640×512	130	130×2	Heavy single pixel-sized noises. Undulating ground background.
Dataset 2 (see Fig. 5 (a2))	640×512	39	39×1	Low SCR value. Heavy cloudy-sky background clutters.
Dataset 3 (see Fig. 5 (a3))	640×512	39	39×4	Sky background without cloud. Many noises with pixel level size.
Dataset 4 (see Fig. 5 (a4))	640×512	30	30×2	Sky-sea background with pixel-sized noises.
Dataset 5 (see Fig. 5 (a5))	512×409	7	7×1	Natural background with houses and pixel-sized noises.

of point-sized electronic noises including fixed pattern noise (FPN) such as multiplicative noise and random noise such as gaussian noise. In a real infrared image, the noise type could be one or a mixture of the noises mentioned above.
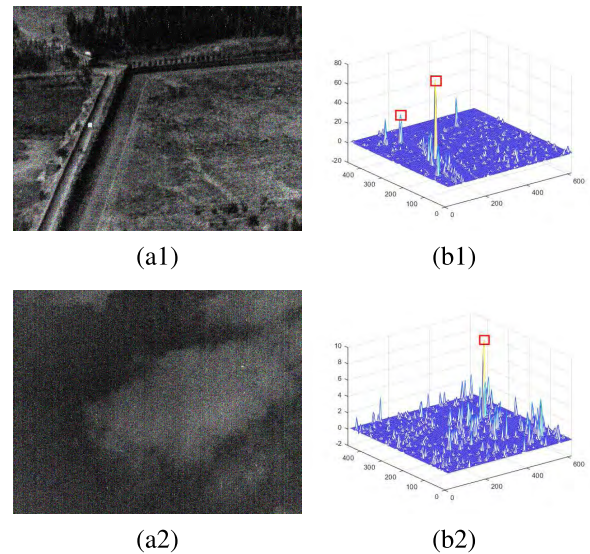
A. DETECTION RESULTS USING DLCM

Experimental results using the proposed detection method are shown in Fig. 5. Fig. 5 (a) is the original infrared image selected from five used datasets, respectively. Fig. 5 (b) is the image after median filtering. It can be seen that there is a significant drop in amount of noises with single pixel size. Fig. 5 (c) is the 3-D saliency map after DLCM calculation and Fig. 5 (d) is the final detection result where small targets are labeled in yellow rectangles automatically by the program. Since the adjacent image blocks need to be analyzed in the DLCM calculation process, we do not calculate the DLCM of the sub-images close to the image boundary. In the threshold operation, we set k to 30 in the formula (6). It can be seen from these figures that the targets are enhanced enormously while the background and noises are suppressed even sometimes directly to zero. Besides, the results of the target recognition are also demonstrated for the simulated mixture noise with multiplicative noise and gaussian noise to clear the resistance of multiplicative and gaussian noise with the proposed method. We add artificially the mixture of multiplicative and gaussian noise to the infrared images of dataset 1 and dataset 2. The multiplicative noise can be written as follows:

$$f(x, y) = g(x, y) + n(x, y) \times g(x, y) \qquad (9)$$

where g is the original image and f is the image with multiplicative noise. Here we set the parameter n obey the gaussian distribution with a mean of 0 and a variance of 0.04. Meanwhile, the mean and variance of gaussian noise are respectively set to 0 and 0.01. Fig. 7 shows the detection results with multiplicative and gaussian noises using the proposed method. From Fig. 7 (b1) and (b2), we can see that the DLCM of the target sub-image always maintains the highest value which means that the target can be extracted without any false detections if the threshold takes a suitable value. As a result, the proposed method has a good resistance of both multiplicative and gaussian noises.

B. COMPARISON WITH OTHER ALGORITHMS

ILCM and LCM are the classical and representative algorithms based on *contrast mechanism* which are devoted to

(a1)   (b1)

(a2)   (b2)

FIGURE 7. Detection results using the proposed method for the simulated mixture noise with multiplicative and gaussian noise: (a) simulation image with multiplicative and gaussian noise; (b) 3-D distributions of saliency map. The rows 1-2 are corresponding with the datasets 1-2, respectively.

infrared small target detection. The 2-D difference of Gaussians (DoG) is selected as the pre-process method of the ILCM detection system which highlights the target to a certain extent. Therefore, in order to validate the effectiveness of our proposed method, we compare our algorithm with other five algorithms including LCM, ILCM, DoG, ILCM with DoG preprocess (DoG-ILCM) and machine learning-based kernel regression by using the same infrared small target image datasets. Kernel regression method is an unsupervised machine learning-based recognition method [14]. It is based on background prediction with the purpose of obtaining an exact background clutter using kernel regression. Then, the predicted background clutter can be eliminated from the original image to obtain the target image containing potential target candidates.

For quantifying the comparison results, two typical evaluation indicators are applied, i.e., background suppression factor (BSF) [26] and SCR Gain [27] [28]. The BSF generally represents the residual degree of noises and background clutters which is defined as follows

$$\text{BSF} = \frac{\sigma_{bi}}{\sigma_{bo}} \qquad (10)$$

TABLE 2. Evaluations (SCR Gain and BSF) comparison among the proposed method and other five methods using five real infrared image datasets.

Methods	Evaluations	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
LCM	SCR Gain	0.5734	1.0445	0.9286	1.2814	0.7636
	BSF	0.6938	0.5480	0.5653	0.8605	0.4041
ILCM	SCR Gain	3.3112	6.0298	17.8689	7.0888	7.7273
	BSF	1.2330	1.7912	1.9593	1.0035	1.6197
DoG	SCR Gain	0.2831	0.5926	3.9842	3.5907	6.7049
	BSF	0.0112	0.0117	0.0121	0.0140	0.0122
DoG-ILCM	SCR Gain	0.8571	0.2951	4.8102	1.6767	3.7107
	BSF	1.4587	1.9422	1.8998	1.2280	1.7367
Kernel regression	SCR Gain	3.0478	3.0993	4.8102	7.3599	4.9955
	BSF	8.3820	7.6696	1.8998	20.0479	9.4905
Proposed	SCR Gain	**38.4500**	**63.7918**	**147.6561**	**81.3295**	**37.6967**
	BSF	**37.2852**	**251.3711**	**107.9967**	**93.1656**	**50.7482**

TABLE 3. Computational cost comparison among the proposed method and other five methods for a single image (in seconds).

Time consumption (s)	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5
LCM	126.7185	125.4363	114.5567	116.6178	69.2064
ILCM	5.0935	5.8071	5.8851	5.1184	3.0522
DoG	5.0564	4.9686	4.9637	5.4234	3.1136
DoG-ILCM	5.4275	5.2881	5.4111	5.3391	3.1764
Kernel regression	1295.3669	1389.9422	1565.4194	1378.5312	1079.6583
Proposed	5.5708	5.2122	5.4148	5.4463	2.1159

where σ_{bi} and σ_{bo} are respectively the background clutter standard deviation of the original image and the saliency map generated from different algorithms. Since SCR is widely used to measure the difference between targets and background, we use SCR Gain to figure out how much the targets are enhanced through the process. The SCR Gain can be rewritten as follows

$$SCR = \frac{|m_t - m_b|}{\sigma_b} \qquad (11)$$

$$SCR\ Gain = \frac{SCR_o}{SCR_i} \qquad (12)$$

where SCR_i and SCR_o are respectively the SCR value before and after detection, m_t and m_b are respectively the gray average of the target and background, and σ_b is the background clutter standard deviation. In some special cases, if an image has more than one target, its SCR can be seen as the average SCR value of all targets in this image:

$$SCR = \frac{1}{M} \sum_{j=1}^{M} SCR_j \qquad (13)$$

where M is the number of targets.

In the task of infrared small target detection, we need to enhance the target and suppress the background to the best. Generally the higher BSF and SCR Gain value of an image are, the better the algorithm performs. The experimental results of two evaluation indicators BSF and SCR Gain are listed in Table 2. Besides, 3-D distributions of saliency map obtained through different detection methods are shown in Fig. 6. It is obvious that our proposed detection system

achieves a nearly perfect target enhancement and background suppression with highest BSF and SCR Gain value while other algorithms can not make a distinction between targets and background.

In order to further demonstrate the excellent performance of the proposed method against other five algorithms, the receiver operation characteristic (ROC) curves are drawn for each dataset through setting threshold to different values. According to the description in [29], ROC curves represent the relationship between detection probability P_d and false alarm rate P_f and it can provide a quantitative comparison of the detection performances. Here P_d and P_f are respectively defined as follows:

$$P_d = \frac{\text{the number of the true detections}}{\text{the total number of targets}}$$

$$P_f = \frac{\text{the number of false detections}}{\text{the total number of targets}} \qquad (14)$$

Fig. 8 shows that the proposed method has better ROC curves compared with the other five algorithms, which demonstrates the proposed method always achieves a higher detection probability with the same false alarm.

It should be noted that DLCM and ILCM have similar computational cost because of the same time complexity. Similar with ILCM, DLCM also takes three procedures: image segmentation, DLCM calculation (ILCM calculation for ILCM) and threshold operation. Therefore, the time complexity of DLCM and ILCM is related to the number of image blocks after the image segmentation (if the image blocks are with same size). Assuming that the number of image blocks is N, then the time complexity of DLCM and ILCM
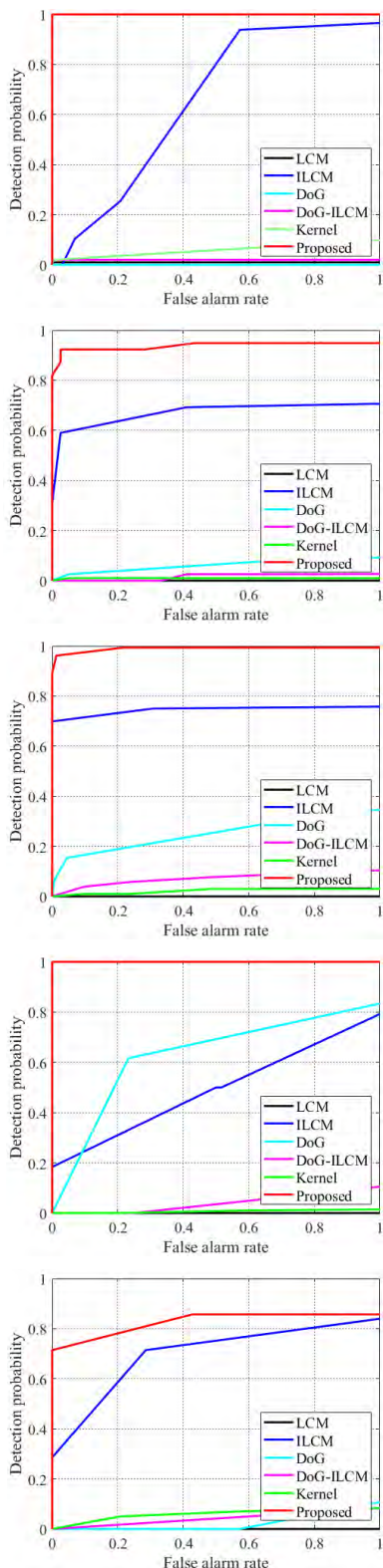
FIGURE 8. ROC curves of five datasets obtained using LCM, ILCM, DoG, DoG-ILCM, kernel regression and the proposed method. From up to down five ROC curve groups are respectively corresponding with the infrared image datasets 1-5.

is $O(N)$. Besides, due to none of image segmentation process, the LCM method calculates local contrast pixel by pixel resulting in much more time consumption. To show the time consumption of the proposed method quantitatively, we have compared the average computational cost of the six detection methods for a single image of the five used image datasets as shown in Table 3. As we can see, the proposed method takes good performance whose time consumption is of the same order of magnitude with ILCM, DoG and DoG-ILCM and much faster than LCM and kernel regression method.

The experimental results show that both LCM and kernel regression methods can not process the point-sized electronic noises leading to low detection probability and high false alarm rate. The kernel regression method performs well in the evaluations SCR Gain and BSF which means that it can enhance the target and suppress the background to a certain extent. In addition, because of calculation pixel by pixel, both LCM and kernel regression methods take long detection time, particularly the kernel regression method. However, ILCM can achieve a fast detection speed through the image segmentation. ILCM is the improvement of LCM mainly leading to the robustness towards the point-sized noises and the less time consumption. Indeed, it achieves better performance in the ROC curves with good enhancement of target (SCR Gain is good) but can almost not weaken the background (BSF is low). It is noted that DoG performs weakly in most of image datasets except the dataset 4 due to that the images in the dataset 4 have relative simple background without many edges. We will analyze this in detail in the next section. The proposed algorithm can achieve not only a fast detection speed but also the best performance in detection probability, false alarm rate, target enhancement and background suppression in all the six detection methods. The relative local contrast measure (RLCM) technique is introduced in [30] and [31]. RLCM is also calculated pixel by pixel which leads to a slow detection process.

C. ANALYSIS OF DoG FILTER
Difference of Gaussians (DoG) is proposed in [6] and is applied to filter the low frequency components of an image. Besides, DoG filter also catches much attention and is used to ILCM for pre-process. Here DoG is a combined band-pass filter which can be written as the following form:

$$\sum_{n=1}^{N} \mathrm{DoG}(x, y, \sigma_1^n, \sigma_2^n) \equiv G(x, y, \sigma_1^N) - G(x, y, \sigma_2^1) \quad (15)$$

where G represents the 2-D Gaussian function, (x, y) represents a pixel in an image, parameters σ_1^N and σ_2^1 respectively determine the low cut-off frequency and the high cut-off frequency. In fact, preserving high frequency components of the image is similarly an object edge extraction operation which can be shown in Fig. 9. It performs well when the

(a)

(b)
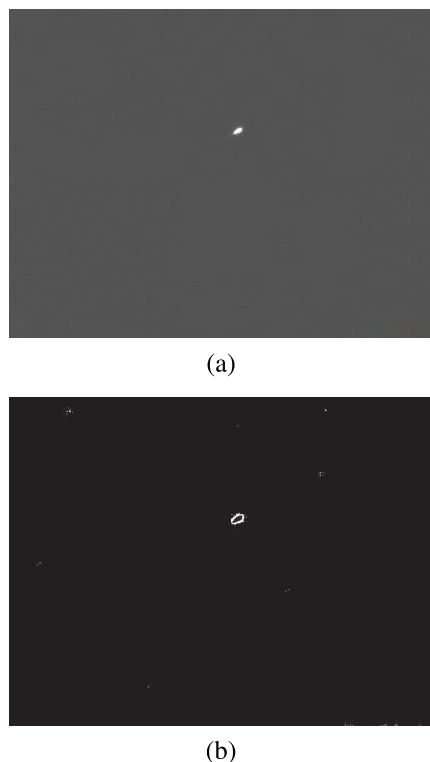
FIGURE 9. DoG process is similarly an object extraction operation: (a) simulation image that contains one small target against smooth and simple background; (b) image after DoG process.

background is simple, smooth and lack of noises. But if the background is wavy and with rich texture details, DoG process will highlight the texture details and object edges of the complex background which leads to a significant decline in target detection.

V. CONCLUSION

In this paper, a difference-based infrared small target detection algorithm based on *contrast mechanism* is proposed for infrared small target detection under complex background. First, we use a median filter to reduce electronic noises with pixel level size, and the infrared image after filtering is divided into a serious of sub-images with the same slide length. Then, a difference-based local contrast measure derived from ILCM is designed in our detection system. Finally, we apply a threshold to the saliency map and the target can be detected. Experiments on five real infrared image datasets with complex background show that the proposed detection system can achieve a high detection probability with a low false alarm rate, fast detection speed and has better performance in target enhancement and background suppression compared with the traditional algorithms.

REFERENCES

[1] S. Ma, Z. Tao, X. Yang, Y. Yu, X. Zhou, and Z. Li, "Bathymetry retrieval from hyperspectral remote sensing data in optical-shallow water," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 2, pp. 1205–1212, Feb. 2014.

[2] S. J. Krotosky and M. M. Trivedi, "Person surveillance using visual and infrared imagery," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 8, pp. 1096–1105, Aug. 2008.

[3] S. Bourennane, C. Fossati, and A. Cailly, "Improvement of target-detection algorithms based on adaptive three-dimensional filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1383–1395, Apr. 2011.

[4] M. Malanowski and K. Kulpa, "Detection of moving targets with continuous-wave noise radar: Theory and measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3502–3509, Sep. 2012.

[5] A. N. Enemuo, R. R. Chaudhuri, Y. Song, and S.-W. Seo, "Thermo-optic sensor based on resonance waveguide grating for infrared/thermal imaging," *IEEE Sensors J.*, vol. 15, no. 8, pp. 4213–4217, Aug. 2015.

[6] X. Wang, G. Lv, and L. Xu, "Infrared dim target detection based on visual attention," *Infr. Phys. Technol.*, vol. 55, no. 6, pp. 513–521, Nov. 2012.

[7] M. Zeng, J. Li, and Z. Peng, "The design of top-hat morphological filter and application to infrared target detection," *Infr. Phys. Technol.*, vol. 48, no. 1, pp. 67–76, Apr. 2005.

[8] J. F. Khan and M. S. Alam, "Target detection in cluttered forward-looking infrared imagery," *Opt. Eng.*, vol. 44, no. 7, 2005, Art. no. 076404.

[9] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, Feb. 2007.

[10] S. Xiang, F. Nie, C. Zhang, and C. Zhang, "Interactive natural image segmentation via spline regression," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1623–1632, Jul. 2009.

[11] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, "DSSD : Deconvolutional single shot detector," 2017, *arXiv:1701.06659*. [Online]. Available: https://arxiv.org/abs/1701.06659

[12] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems 28*, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2015, pp. 91–99.

[13] J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: Object detection via region-based fully convolutional networks," in *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, Eds. Red Hook, NY, USA: Curran Associates, 2016, pp. 379–387.

[14] Y. Gu, C. Wang, B. Liu, and Y. Zhang, "A kernel-based nonparametric regression method for clutter removal in infrared small-target detection applications," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 3, pp. 469–473, Jul. 2010.

[15] P. Le Callet, C. Viard-Gaudin, and D. Barba, "A convolutional neural network approach for objective video quality assessment," *IEEE Trans. Neural Netw.*, vol. 17, no. 5, pp. 1316–1327, Sep. 2006.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.

[18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, P. Bartlett, F. C. N. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Red Hook, NY, USA: Curran Associates, 2012, pp. 1097–1105.

[19] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556

[20] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9.

[21] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2014.

[22] Z. Wen, B. Hou, Q. Wu, and L. Jiao, "Discriminative feature learning for real-time SAR automatic target recognition with the nonlinear analysis cosparse model," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 7, pp. 1045–1049, Jul. 2018.

[23] J. Han, Y. Ma, B. Zhou, F. Fan, K. Liang, and Y. Fang, "A robust infrared small target detection algorithm based on human visual system," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2168–2172, Dec. 2014.

[24] S. Vishaga and S. L. Das, "A survey on switching median filters for impulse noise removal," in *Proc. Int. Conf. Circuits, Power Comput. Technol.*, Mar. 2015, pp. 1–6.

[25] B. Du and L. Zhang, "Target detection based on a dynamic subspace," *Pattern Recognit.*, vol. 47, no. 1, pp. 344–358, 2014.

[26] X. Qu, H. Chen, and G. Peng, "Novel detection method for infrared small targets using weighted information entropy," *IEEE Sensors J.*, vol. 23, no. 6, pp. 838–842, Dec. 2012.

[27] C. Q. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4996–5009, Dec. 2013.

[28] H. Deng, X. Sun, and X. Zhou, "A multiscale fuzzy metric for detecting small infrared targets against chaotic cloudy/sea-sky backgrounds," *IEEE Trans. Cybern.*, vol. 49, no. 5, pp. 1694–1707, May 2019.

[29] S. Qi, J. Ma, C. Tao, C. Yang, and J. Tian, "A robust directional saliency-based method for infrared small-target detection under various complex backgrounds," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 495–499, May 2013.

[30] J. Han, K. Liang, B. Zhou, X. Zhu, J. Zhao, and L. Zhao, "Infrared small target detection utilizing the multiscale relative local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 612–616, Apr. 2018.

[31] J. Han, Y. Yu, K. Liang, and H. Zhang, "Infrared small-target detection under complex background based on subblock-level ratio-difference joint local contrast measure," *Opt. Eng.*, vol. 57, no. 10, 2018, Art. no. 103105.

KE YANG received the B.S. degree in electronic science and technology from Xi'an Jiaotong University, Xi'an, China, in 2017. He is currently pursuing the M.S. degree with the Department of Micro/Nano-Electronics, Shanghai Jiao Tong University, Shanghai, China. His current research interests include integrated circuit design automation and infrared small target detection.
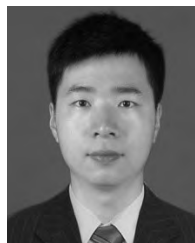
SHAOYI LI received the bachelor's degree in measurement and control technology from Southwest Jiaotong University, Chengdu, China, in 2008, and the master's and Ph.D. degrees in navigation, guidance, and control from Northwestern Polytechnical University, Xi'an, China, in 2011 and 2015, respectively, where he is currently an Assistant Researcher with the School of Aerospace. He has authored or coauthored more than ten papers in scientific journals and conference proceedings. His current research interests include ATR based on infrared image, artificial intelligence, and image processing.

HAI-BAO CHEN received the B.S. degree in information and computing sciences and the M.S. and Ph.D. degrees in applied mathematics from Xi'an Jiaotong University, Xi'an, China, in 2006, 2008, and 2012, respectively. He then joined Huawei Technologies, where he focused on cloud computing and big data. He was a Postdoctoral Research Fellow with the Electrical Engineering Department, University of California, Riverside, CA, USA, from 2013 to 2014. He is currently an Associate Professor with the Department of Micro/Nano-Electronics, Shanghai Jiao Tong University, Shanghai, China. He has authored or coauthored about 50 papers in scientific journals and conference proceedings. His current research interests include model-order reduction, system and control theory, circuit simulation, cloud computing and big data, and electromigration reliability. He received the one Best Paper Award nomination from the Asia and South Pacific Design Automation Conference (ASP-DAC), in 2015. He currently serves as an Associate Editor for *Integration, the VLSI Journal*.

KAI ZHANG received the bachelor's, master's, and Ph.D. degrees in navigation, guidance, and control from Northwestern Polytechnical University, Xi'an, China, in 2001, 2004, and 2009, respectively, where he is currently an Associate Professor with the School of Aerospace. He has authored or coauthored more than 20 papers in scientific journals and conference proceedings. His current research interests include infrared imaging, artificial intelligence, and infrared scene modeling and simulation.

• • •