Infrared Small Target Detection Based on Gradient-Intensity Joint Saliency Measure

Yongsong Li, Zhengzhou Li, Member, IEEE, Weite Li, and Yuchuan Liu

Abstract—Small target detection is an arduous mission in the infrared search and tracking system, especially when the target signal is disturbed by high-intensity background clutter. In view of this situation, this article presents a robust target detection algorithm based on gradient-intensity joint saliency measure (GISM) to gradually eliminate complex background clutter. Because of thermal remote sensing imaging, the infrared target usually occupies a small area that accords with the optics point spread function, so it can be distinguished from the background clutter in both gradient and intensity properties. According to this, first, the original image is transformed into a gradient map, and the gradient saliency measure (GSM) is calculated to highlight the target signal and suppress the sharp edge clutter, so the candidate targets can be reliably extracted by using the maximum entropy principle. Second, the local intensity saliency measure (LISM) is obtained by calculating the gray difference between each candidate region and its local surroundings, so as to preserve the real target and remove intense structural clutter such as black holes or corners. Finally, by fully integrating the gradient and intensity properties, the GISM defined by LISM-weighted GSM map can efficiently identify the real target signal and eliminate false alarms. Experimental results prove that the proposed method not only has advantages in background clutter suppression and small target enhancement, but also has reasonable time consumption.

Index Terms—Gradient saliency measure (GSM), infrared imaging, local intensity saliency measure (LISM), small target detection.

I. INTRODUCTION

INFRARED small target detection plays an important role in homeland security systems such as coastal defense and air defense, which requires high robustness and accuracy [1], [2]. But unfortunately, in real infrared video surveillance, the complex background is often very chaotic and the target contrast is very low [4], which leads to the performance of existing algorithms being seriously weakened. Therefore, infrared small target detection in complex background clutter is a challenging problem, which has attracted the attention of numerous researchers [5], [6]. In recent ten years, many algorithms have been developed, which can be roughly divided into four categories: filtering-based method, sparse-based method, visual contrast-based method, and segmentation-based method.

In the filtering-based method, the reasonable filter is constructed in terms of the distribution law of the target and background to suppress background clutters and check the target signal. Traditional filters, such as max-median/max-mean filters [7] and Butterworth filter [8] are widely used for small target detection because of their simple design. In the literature [9], the modified wavelet filter performs well in the detection of small targets a sea-sky background by calculating the local singularity of horizontal and vertical sub-bands. In addition, the morphological filters are also very successful in the field of infrared small target detection [10], [11]. By modifying morphological operation rules and structuring elements, the new top-hat filter reported by Bai et al. [12], [13] can protrude small target signals with gray difference from the surrounding background. However, when weak targets are interfered by complex background clutter, the above algorithms will have poor clutter suppression ability. On the basis of Bai’s research, Deng et al. [2] modified structuring elements by adaptively fusing local background information, and Li and Li [14] constructed directional structuring elements by considering edge clutter elimination, which further improved the robustness of morphological filtering method for infrared small target detection.

The sparse-based method usually reconstructs sparse matrix and low-rank matrix to separate target components according to the sparsity of targets and the nonlocal similarity of background clutter. Gao et al. [15] put forward the infrared-patch-image (IPI) method based on sparse and low-rank matrix reconstruction, which has high accuracy in characterizing the small target structure. Furthermore, by imposing more constraints, Dai and Wu [16] introduced the reweighted IPI model and Zhang et al. [17] designed the nonconvex rank approximation minimization model, which significantly improved the target detection ability of the IPI model under complex background interference. In another way, Liu et al. [18] and Wang et al. [19] constructed the sparse representation methods via learning the over-complete dictionaries of sky background or sea-sky background,
respectively, which are effective in detecting small targets in specific scenes. However, some intense clutter may be as sparse as the target signal in complex background, which easily leads to false detection or missed detection due to the aliasing of the target and background. Besides, the above algorithms spend a lot of time on target-background decomposition and iterative processing, which declines their feasibility in real-time detection.

The visual contrast-based method realizes target enhancement and background suppression according to the selective attention mechanism of the human visual system. Kim and Lee [20], [21] established the scale space of Laplacian-of-Gaussian operator to protrude the point-like targets from the background clutter. Moradi et al. [22], [23] reported the average absolute gray difference methods, which are suitable for enlarging the small target and suppressing high-intensity background clutter. Chen et al. [24] designed a local contrast measure (LCM) method to enhance small targets by calculating the gray differences between the central subblock and the surrounding subblocks in eight directions. Furthermore, Han et al. [25] revised the LCM method and reported tri-layer LCM and weighted strengthened LCM [26], which are effective for target detection in high-intensity background clutter. Besides, Liu et al. [27] presented the gradient direction diversity weighted multiscale flux density method by transforming the infrared image into gradient vector field, which can well distinguish small targets from background clutter. Bai et al. [28], [29] calculated the derivatives of the image in multiple directions based on the facet model and fused all derivatives to enhance the target and suppress background clutter. Based on the intensity and gradient properties of small target, Zhang et al. [30] calculated the local intensity and gradient (LIG) map to detect small targets and mitigate the background clutters. The visual contrast-based algorithms can obtain excellent detection ability in many scenarios, but they may have large responses for some high-contrast undulating clutters.

In the segmentation-based method, the original image is first divided into regions, and then each region is analyzed according to the regional characteristics of the small target and background, so as to check the real target region with maximum likelihood. In the literature [31], the top-hat transformation was adopted to filter and segment the infrared image to extract the candidate target regions, and then a classifier based on multifeature learning was presented to identify the real small target. Qin et al. [32] constructed a target detection algorithm based on image segmentation, which segmented the candidate targets after facet kernel filtering, and then calculated the local contrast descriptor of each candidate to recognize real small targets. Subsequently, Qin and Li [33] also employed difference-of-Gaussian (DoG) filter to obtain the segmentation result, and based on this, a novel local contrast method was constructed. Moreover, Huang et al. [34] selected 20 pixels with the largest peak density in the image as seed points and then utilized the maximum-gray region growing algorithm to recognize the real target. Recently, Chen et al. [35] used the DoG filter to highlight and segment candidate target regions and then detected the small targets by analyzing the intensity and gradient information of each candidate region. In the literature [36], a sliding window was designed to quickly segment salient regions, and then an improved fuzzy C-means method was developed, which can distinguish the real target from the background. In summary, these above segmentation-based methods can further optimize the robustness of the algorithm by focusing and analyzing the local characteristics of each candidate region.

Although researchers have done a lot of infrared small target detection, it is still an open topic that deserves further exploration. This article proposes a simple but effective scheme based on gradient-intensity joint saliency measure (GISM) to extract dim small targets under complicated backgrounds. First, according to the gradient property of small target, the horizontal and vertical derivatives of the image are calculated and fused, which can well suppress sharp edge clutter and enhance candidate target regions. Second, in view of the intensity property of small target, the local gray difference between each candidate target region and its surrounding background is computed, which can stably remove black holes and strong corners. Experimental results demonstrate that by fully considering the gradient and intensity properties, the proposed small target detection method is superior to several state-of-the-art algorithms and has a good effect on suppressing sharp edges and strong corners in complex background. The contributions of this article can be summarized as follows.

1) By deeply mining the gradient characteristic of small target, the gradient saliency measure (GSM) is constructed to evaluate the derivative differences between target signal and edge clutter in \(x\) and \(y\) directions, which can effectively eliminate edge clutter and enhance the target signal.

2) By analyzing the center-surround gray difference of small target, the local intensity saliency measure (LISM) is designed to make up for the deficiency of GSM in suppressing black holes and corner clutters, so as to further remove high-intensity structural clutter while retaining the target signal.

3) Combining the above approaches and their advantages, an effective small target detection scheme is proposed to eliminate background clutters one by one, i.e., edge clutter, black hole, and corner clutter. Hence, the proposed method is superior to some advanced algorithms in weak target detection under complex background with reasonable time consumption.

The rest of this article is as follows: Section II analyzes the characteristics of infrared small target. The overall target detection algorithm is included in Section III. Section IV shows the experimental results and discussions. Section V gives the conclusions of this article.

II. CHARACTERISTICS OF INFRARED SMALL TARGET

An infrared small target image can be modeled as

\[
f(x, y) = f_t(x, y) + f_b(x, y) + f_n(x, y)
\]

where \(f(x, y)\), \(f_t(x, y)\), \(f_b(x, y)\), and \(f_n(x, y)\) are the original image, target component, background component, and noise component at pixel coordinate \((x, y)\), respectively. Due to the energy attenuation, optical defocus, lens aberration, and other factors in thermal remote sensing imaging system, the infrared
small target is usually a spot region that accords with the optics point spread function (PSF), and its size range is $2 \times 2 \sim 9 \times 9$ pixels [15], [20], [28]

$$f_t(x, y) = A^+_{\text{peak}} \times \exp \left\{ -\frac{1}{2} \left( \frac{x - \bar{x}}{\sigma_x} \right)^2 + \left( \frac{y - \bar{y}}{\sigma_y} \right)^2 \right\}$$

(2)

where $A^+_{\text{peak}}$, $(\bar{x}, \bar{y})$, and $\sigma_x$ and $\sigma_y$ indicate the peak amplitude, centroid position, and horizontal and vertical scales of the small target, respectively. Therefore, the infrared small target is a Gaussian-shaped bright spot that is discontinuous with the surrounding environment, and it has two spatial attributes that can be distinguished from the background clutter.

**Attribute 1:** The gradients of small target have great convergence, while the gradients of strong edge clutter are usually directional and have weak convergence [27], [30].

**Attribute 2:** The amplitude of small target is usually brighter than its local surroundings, while the amplitude of dark holes and corners is smaller than or close to the local surroundings [14], [26].

To sum up, the real target is predictable in both gradient and intensity attributes, and this concept can be used for the design of infrared small target detector, which is beneficial to distinguish small targets from complex background.

### III. PROPOSED SMALL TARGET DETECTION ALGORITHM

According to the gradient and intensity characteristics of small target, a robust target detection algorithm is proposed in this section, which mainly includes three stages: First, the GSM is designed to suppress most edge clutter and enhance candidate targets; second, the LISM is proposed to highlight the target signal and suppress black holes, noisy pixels, and corner clutters. Finally, the real target can be reliably identified by fusing GSM map and LISM map. The details are presented in the following parts.

#### A. Gradient Saliency Measure

According to **Attribute 1**, the gradients of small target have great convergence, while the gradients of strong edge clutter are usually directional and have weak convergence. Given an infrared image $f$, the corresponding gradient vector can be easily computed by

$$\nabla f(x, y) = [f'_x(x, y), f'_y(x, y)]$$

(3)

here $f'_x(x, y)$ and $f'_y(x, y)$ are the derivatives along $x$- and $y$-directions

$$f'_x(x, y) = \frac{f(x + 1, y) - f(x - 1, y)}{2}$$

(4)

$$f'_y(x, y) = \frac{f(x, y + 1) - f(x, y - 1)}{2}.$$  

(5)

Fig. 1 shows that an infrared image is decomposed into derivative maps in the $x$- and $y$-directions, and small target is labeled by a red circle and strong oblique edge clutter is marked by a purple ellipse. It can be seen that the small target has a relatively large response values in both $x$- and $y$-directional derivative maps.

In Fig. 1(b), the horizontal edges in the $x$-directional derivative map have relatively small responses, while their responses in the $y$-directional derivative map are relatively large, as shown in Fig. 1(c). Similarly, the vertical edge has a relatively small response in the $y$-directional derivative map, but a large response in the $x$-direction derivative map.

From the above observation, it can be found that the small target has a greater response than most background clutters in the gradient map. Meanwhile, considering that the target has gradient convergence and the edge clutter has directivity, in order to amplify the target signal and smooth the background noise in the gradient map, we define the horizontal, vertical, and diagonal derivative amplitude maps as follows, which are recorded as $f'_h$, $f'_v$, and $f'_d$ in turn

$$\begin{align*}
  f'_h(x, y) &= (f'_x(x, y))^2 \ast \text{DoG}_{\sigma_1, \sigma_2} \\
  f'_v(x, y) &= (f'_y(x, y))^2 \ast \text{DoG}_{\sigma_1, \sigma_2} \\
  f'_d(x, y) &= (f'_x(x, y) \times f'_y(x, y)) \ast \text{DoG}_{\sigma_1, \sigma_2}
\end{align*}$$

(6)

where $\ast$ is the convolution operation. The reason why DoG kernel filtering is adopted here is that, unlike the Gaussian kernel, the DoG kernel is based on the visual contrast mechanism, which can not only smooth the noise components, but also enhance the central area and suppress the surrounding background [37], [38]. The DoG$_{\sigma_1, \sigma_2}$ with two standard deviations $\sigma_1$ and $\sigma_2$

$$\text{DoG}_{\sigma_1, \sigma_2}(x, y) = G_{\sigma_1}(x, y) - G_{\sigma_2}(x, y)$$

$$= \frac{1}{2\pi\sigma_1^2} \exp \left( -\frac{x^2 + y^2}{2\sigma_1^2} \right) - \frac{1}{2\pi\sigma_2^2} \exp \left( -\frac{x^2 + y^2}{2\sigma_2^2} \right).$$

(7)

Generally, $\sigma_1 < \sigma_2$ and $\sigma_2/\sigma_1 = 2$ can better guarantee the positive center and negative periphery. $\sigma_1$ determines the size of the positive center of DoG kernel, and the setting of parameter $\sigma_1$ will be discussed in Section IV. Fig. 2 shows the three-dimensional (3-D) mesh plot of DoG kernel when $\sigma_1 = 3$. Fig. 3 gives the schematic diagram of derivative waveform $f'_d$ at a small target region. The original small target signal is displayed in Fig. 3(a), and its directional derivative with positive and negative lobes is shown in Fig. 3(b). Fig. 3(c) is the square of the derivative, which can highlight the derivative amplitude.
Fig. 2. 3-D mesh of the DoG kernel ($\sigma_1 = 3$).

Fig. 3. Schematic diagram of derivative waveform $f'_h$ at a small target region. (a) Original small target signal. (b) $x$-directional derivative. (c) Square of the derivative. (d) Waveform after DoG filtering on (c).

of the target with a positive value, and then the target signal can be completely amplified by DoG filtering as shown in Fig. 3(d).

Fig. 4 shows the difference between small target and oblique edge clutter in directional derivative maps. It can be known that both the small target and the oblique edge clutter have larger response values in the $x$- and $y$-directional derivative maps. Accordingly, they have relatively high values in both $f'_h$ and $f'_v$ maps, as shown in Fig. 4(e1), (f1), (e2), and (f2). Nevertheless, from Fig. 4(d1), the small target area has some positive and negative response values with four petal shapes in the map of $f'_x \times f'_y$, while in Fig. 4(d2) the oblique edge clutter has a relatively large positive response. Hence, the response value of the oblique edge clutter is very large, and much greater than that of the small target area in the $f'_{\text{diag}}$ map obtained from $f'_x \times f'_y$ map after DoG filtering. This difference is beneficial to distinguish the small target signal from the strong oblique edge clutters.

According to the above analysis, the small target has larger response values in both $f'_h$ and $f'_v$ maps, while horizontal or vertical edge clutters have little response in the $x$- and $y$-directional derivatives respectively. Afterward, although the oblique edge also has great responses in $f'_h$ and $f'_v$ maps, and its response in the $f'_{\text{diag}}$ map is also very large, the small target has little response in the $f'_{\text{diag}}$ map. Consequently, in order to reasonably depict the

Attribute 1 to amplify the target signal and eliminate background clutter, the GSM is defined as follows:

$$\text{GSM}(x, y) = \max \left\{ \left( f'_h(x, y)f'_v(x, y) - (f'_{\text{diag}}(x, y))^2 \right) ; 0 \right\}$$

where $\max(\bullet)$ is the maximum operation. By the definition of (8), the small target can be reliably enhanced and the horizontal or vertical edge clutters can be suppressed by using $f'_h$, $f'_v$; Furthermore, in the $f'_{\text{diag}}$ map, the response of small target is very small, while the response of the oblique edge clutter is relatively large, so the oblique edge clutters can be greatly eliminated by using $f'_h f'_v - (f'_{\text{diag}})^2$. Obviously, $f'_h f'_v \gg (f'_{\text{diag}})^2$ (i.e., $f'_h f'_v - (f'_{\text{diag}})^2 \gg 0$) should be satisfied for the small target as illustrated in Fig. 5(a1)–(c1), but this condition is generally not satisfied for most strong edge clutters as illustrated in Fig. 5(a2)–(c2), so adding a non-negative operation to (8) can effectively protrude small targets and suppress most of the background clutters.

Fig. 6 gives the calculation process of GSM. As shown in Fig. 6(a), most of the background clutter can be suppressed, while small target and oblique edge clutter have a larger response.
so there are still some clutter residues at the black holes and corners that cannot be ignored in the processed results. Recall that Attribute 2, the infrared small target is usually brighter than the adjacent background, so by measuring the gray difference between each candidate target and its local surroundings, the GSM map can be further refined to remove the strong clutter residues.

Inspired by solving the small target detection problem from the perspective of image segmentation [31, 32, 35], the Kapur’s entropy [39] is employed to extract candidate targets in the GSM map. First, we regularize the magnitude of GSM map into the range \([0, 255]\) to facilitate the calculation and segment the regularized GSM map into candidate target class and background class by using the threshold \(t \in [0, 255]\). Their information entropy is derived as follows:

\[
\begin{align*}
H_B (t) & = - \sum_{j=0}^{255} \left[ \frac{p(j)}{p(t)} \ln \left( \frac{p(j)}{p(t)} \right) \right], \\
H_{CT} (t) & = - \sum_{j=0}^{255} \left[ \frac{p(j)}{1-p(t)} \ln \left( \frac{p(j)}{1-p(t)} \right) \right],
\end{align*}
\]

where \(j\) is the gray level of the regularized GSM map, and \(p(B) = \sum_{j=0}^{255} p(j)\) is the probability distribution of background class at threshold \(t\). \(H_{CT}(t)\) and \(H_B(t)\) are the information entropy of candidate target class and background class, and the sum of information entropy is \(H(t) = H_B(t) + H_{CT}(t)\).

According to maximum entropy principle [39], the optimal threshold is

\[
t^* = \arg \max_{t \in [0,255]} \{ H(t) \}.
\]

After the binarization of GSM map with \(t^*\), all candidate targets can be extracted and labeled with a connected component labeling algorithm [40], and the candidate target regions are grouped into a pixel set \(\mathbb{R} = \{ R_1, \ldots, R_K \} \). Then, the centroid position and major axis length of each candidate region can be easily obtained, which are recorded as \((\bar{x}_k, \bar{y}_k)\) and \(MajorAxis_k\) respectively. In order to enhance small targets and eliminate clutter residues, two guard windows are generated to perceive the surrounding background information of each candidate target, which are defined by

\[
\begin{align*}
\Omega^{(1)}_k & = \left\{ (m, n) \mid \max \{ |\bar{x}_k - m|, |\bar{y}_k - n| \} = \frac{MajorAxis_k}{2} \right\}, \\
\Omega^{(2)}_k & = \left\{ (p, q) \mid \max \{ |\bar{x}_k - p|, |\bar{y}_k - q| \} = \frac{MajorAxis_k}{2} + 2 \right\}
\end{align*}
\]

where \(\Omega^{(1)}_k\) and \(\Omega^{(2)}_k\) represent the pixel coordinate sets of first and second guard windows of \(k\)th candidate target. Afterward, since small target is usually brighter than its local surroundings, the average of the maximum gray values in the two guard windows can be estimated as the local background value of \(k\)th candidate target, which is expressed as

\[
g^{BG}_k = \text{mean} \left( \max \left( f (\Omega^{(1)}_k) \right), \max \left( f (\Omega^{(2)}_k) \right) \right)
\]

where \(\text{mean}(\cdot)\) is the average operation. \(f(\cdot)\) is the gray-scale extraction operation, and its output is the gray values of the input pixel coordinate set corresponding to the original image.

B. Local Intensity Saliency Measure

By reasonably exploiting the gradient characteristics, the GSM map can be adopted to enhance the target signal and suppress most of the background and edge clutters. However, the GSM does not sufficiently consider the intensity information,
Next, in order to effectively delineate Attribute 2, the LISM is defined by calculating the gray difference between candidate targets and their local background values

\[
\text{LISM}(x, y) = \max \left\{ \left( w_k \times (f(x, y) - g_k^{BG}) \right), 0 \right\}, \forall (x, y) \in R_k \subseteq \mathbb{R}
\]

(13)

where \( w_k \) is the weighting factor of the \( k \)th candidate target, which can be computed as

\[
w_k = \left\{ \begin{array}{ll}
\frac{\sum_{i=1}^{N_k} g_k(i) / N_k}{g_k^{BG}} \times \left( \sum_{i=1}^{N_k} g_k(i) / N_k - g_k^{BG} \right), & N_k \geq 3 \\
0, & \text{Otherwise}
\end{array} \right.
\]

(14)

here \( N_k \) is the total number of pixels whose gray level in the \( k \)th candidate target area is greater than \( g_k^{BG} \), and \( g_k(i) \) is the \( i \)th gray value greater than \( g_k^{BG} \) in the candidate target area. Because the small target is usually brighter than the surrounding background, if the pixel is located in the target area, then \( f(x, y) > g_k^{BG} \); otherwise, if the pixel \( (x, y) \) in the black hole or corner regions, then \( f(x, y) \leq g_k^{BG} \). Accordingly, the non-negative constraint in (13) can be used to suppress black hole or corner clutters. Furthermore, considering that the local contrast of the target is usually more salient compared with the background clutter, constructing a weighting factor \( w_k \) can not only further correctly enhance the small target signal and weaken the false alarm clutters, but also remove some isolated noisy pixels with \( N_k < 3 \) in the candidate target region.

Fig. 7 shows the calculation process of the LISM map. As the GSM is based on gradient characteristic, it may amplify some black holes and corner clutters while enhancing the target signal, so it can be seen from Fig. 7(b) that after the entropy threshold segmentation, not only the real target area can be located, but also some corner areas will be remained. Comparing Fig. 7(c1)—(c3) and (d1)—(d3), we can find that the gray value of the real target will be higher than the surrounding background, while the gray value corresponding to the corner area located by GSM is often not greater than the maximum intensity of the surroundings, which is due to the fact that the background clutter has grayscale extension in some specific directions. Consequently, by calculating the gray difference between each candidate target area and the local environment, the LISM map can further enhance the target signal and remove intense structural clutters, such as black holes or corners, as Fig. 7(e) and (f) demonstrates.

C. Small Target Detection Using GISM

According to the discussions in the previous parts, we can see that the GSM map can delineate the convergence characteristic of small targets in the gradient map, and then the LISM map depicts that small targets have certain contrast characteristic in the local image. Consequently, the GISM is defined as the LISM weighted GSM map, which can greatly enhance the reliability of target detection and eliminate a large number of false alarms:

\[
\text{GISM}(x, y) = \text{LISM}(x, y) \times \text{GSM}(x, y).
\]

(15)

Fig. 7. Calculation process of LISM map. (a) Original image. (b) Candidate targets extracted from the binarized GSM map, in which real target and corner clutter are marked by red box and purple box respectively; (c1)—(c3) and (d1)—(d3) are the target region, local image, and gray matrix of candidate target in turn; (e) and (f) are the computed LISM map and corresponding 3-D mesh plot.

In the GISM map, the small target will become more salient compared with background clutter, as shown in Fig. 8, so the real target region can be segmented by an adaptive threshold \( \tau_{\text{GISM}} \):

\[
\tau_{\text{GISM}} = u_{\text{GISM}} + \varepsilon \times \sigma_{\text{GISM}}
\]

(16)

here \( u_{\text{GISM}} \) and \( \sigma_{\text{GISM}} \) denote the average and standard deviation values of the GISM map, respectively. \( \varepsilon \) is the experimental
Algorithm 1: Target Detection Using Gradient-Intensity Joint Saliency Measure (GISM).

**Input:** The raw image $f$.

**Output:** The segmented target image.

**Step 1:** Compute the gradient map of the raw image $f$.

**Step 2:** Obtain the horizontal, vertical and diagonal derivative amplitude maps (i.e., $f'_h$, $f'_v$, and $f'_{\text{diag}}$) by using square operation and DoG filtering, according to (6) and (7).

**Step 3:** Acquire the gradient saliency measure (GSM) to enhance the target signal and suppress background clutters with (8).

**Step 4:** Search the candidate targets in the GSM map by using the maximum entropy principle.

**Step 5:** Generate two guard windows around each candidate target, and calculate the local background value with (11) and (12).

**Step 6:** Obtain the local intensity saliency measure (LISM) to eliminate the black hole and corner regions using (13) and (14).

**Step 7:** Achieve the gradient-intensity joint saliency measure (GISM) by integrating LISM weighted GSM map.

**Step 8:** Separate the real target from the GISM map.

constant, which can be selected from the interval of 6–10 in most cases. Once the pixel intensity in the GISM map is bigger than $\tau_{\text{GISM}}$, it is divided into target pixel. Algorithm 1 summarizes the whole process of the proposed method.

IV. EXPERIMENTAL RESULTS

In this section, eight image sequences with different background clutters are tested, and nine related baseline algorithms are used for performance comparison to verify the accuracy and robustness of our method. This series of experiments were carried out in Matlab 2016a, which was installed on a computer with 16 GB memory and Intel i5-6500 processor.

A. Test Dataset and Metrics

1) **Test Dataset:** In this section, eight infrared sequences are employed to evaluate the performance of the algorithm, and each sequence is a typical scene, which is labeled as Seq_1–8 respectively. The representative images of these sequences are shown in Fig. 9, in which real targets are marked with red circles. As shown in Fig. 9(a), the infrared point target in sequence 1 is located in a relatively mild sky background. In Fig. 9(b), a small aircraft with low contrast is maneuvering into a large cloud. Seq_3 is the situation where two small targets are buried in the background of high-brightness clouds. In Seq_4, two small boats are driving in the river of the urban scene, and the infrared radiation of artificial objects, such as bridges and buildings in the background is higher than the target signal. Fig. 9(e) shows a fast-flying aircraft target appearing in a great number of trivial clouds. In Seq_6, there are many hot pixels with high brightness and sparsity, which will seriously interfere with the detection of small target. Fig. 9(g) shows that a target appears on the cloud layer, but there are some stains in the background, which are similar to the target signal in shape to a certain extent. In Fig. 9(h), two small ship targets with different sizes and shapes are disturbed by high fluctuating and chaotic ocean background clutter. Table I lists the details of the test dataset.

2) **Metrics:** First, the background suppression factor (BSF) and signal-to-clutter ratio gain (SCRG) are used to quantitatively evaluate the clutter suppression capability of the algorithm, which are defined as follows [14], [23]:

$$\text{BSF} = \frac{C_{\text{in}}}{C_{\text{out}}}$$

$$\text{SCRG} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}}}$$

(17)

where $C$ and SCR denote the standard deviations and signal-to-clutter ratio (SCR) levels of the image, $\text{in}$ and $\text{out}$ represent the images before and after background suppression processing. Generally, the larger the obtained SCRG and BSF values, the better the clutter suppression effect of the algorithm.

Moreover, in order to quantitatively evaluate the target detection accuracy of the algorithm, the false positive rate (FPR) and the true positive rate (TPR) are obtained by using a series of thresholds to segment the processed results, and then the receiver operating characteristic (ROC) curve can be defined as the ratio diagram of TPR and FPR [25], [27]. The TPR and FPR are computed as follows:

$$\text{TPR} = \frac{\text{The number of correctly detected targets}}{\text{The number of actual targets}}$$

(18)

$$\text{FPR} = \frac{\text{Pixel number of false alarms}}{\text{Pixel number of the whole image}}$$

(19)

The ROC curve is widely used to analyze the detection ability of algorithms, and the closer the ROC curve is to the upper left corner, the more robust the target detection effect is.

B. Parameters Selection

The parameter $\sigma_1$ of DoG kernel is an important parameter when calculating the GSM map, which is directly related to the target size in the test dataset. To reasonably select the parameter $\sigma_1$, Table II counts the BSF and SCRG values of the GSM map with different parameter $\sigma_1$ configuration on each image of Fig. 9. It can be seen that when the $\sigma_1 = 1$, the DoG kernel cannot completely cover the target region, so the target enhancement
TABLE I
DETAILS OF TEST DATASET

| Sequences | Total frames | Resolution | Target area | Target Characteristic | Background Characteristic |
|-----------|--------------|------------|-------------|-----------------------|---------------------------|
| Seq_1     | 500          | 640 × 480  | 24–30       | ● Point shape.         | ● Stationary background.  |
|           |              |            |             | ● Slow motion.         | ● Big and gentle cloud.   |
| Seq_2     | 800          | 640 × 480  | 32–56       | ● Point shape.         | ● Slow changing background. |
|           |              |            |             | ● Low contrast.        | ● Scattered bright clouds. |
| Seq_3     | 1200         | 640 × 480  | 21–36       | ● Point shape.         | ● Heavy sky clutters.     |
|           |              |            |             | ● Fast motion.         | ● High-intensity clouds.  |
| Seq_4     | 1000         | 640 × 480  | 42–64       | ● Circular shape.      | ● Urban background clutter.|
|           |              |            |             | ● High contrast.       | ● High contrast edges.    |
| Seq_5     | 500          | 320 × 240  | 15–32       | ● Strip shape.         | ● Changing background.    |
|           |              |            |             | ● Fast motion.         | ● Chaotic cloud clutters. |
| Seq_6     | 1500         | 320 × 240  | 3–14        | ● Point shape.         | ● Many noisy pixels.      |
|           |              |            |             | ● Small in size.       | ● Sharp cloud edge.       |
| Seq_7     | 500          | 320 × 240  | 21–37       | ● Ellipse shape.       | ● Blurred sky background. |
|           |              |            |             | ● Fast motion.         | ● Some stripe stains.     |
| Seq_8     | 1000         | 320 × 240  | 25–72       | ● Ellipse/Point shape. | ● High fluctuating background. |
|           |              |            |             | ● High contrast.       | ● Heavy sea clutters.     |

TABLE II
SCRG AND BSF VALUES OF THE GSM MAP WITH DIFFERENT PARAMETER $\sigma_1$ CONFIGURATION FOR THE TEST IMAGES

| Metrics       | Parameter settings | Fig. 9(a) | Fig. 9(b) | Fig. 9(c) | Fig. 9(d) | Fig. 9(e) | Fig. 9(f) | Fig. 9(g) | Fig. 9(h) |
|---------------|--------------------|------------|------------|------------|------------|------------|------------|------------|------------|
| SCR | $\sigma_1=1$     | 194.620    | 45.636     | 25.108     | 1.927      | 16.291     | 15.681     | 4.108      | 8.298      |
|    | $\sigma_1=2$     | 287.830    | 86.524     | 51.905     | 7.920      | 19.443     | 24.579     | 29.225     | 21.874     |
|    | $\sigma_1=3$     | 313.283    | 90.457     | 47.995     | 25.646     | 22.679     | 36.912     | 39.294     | 28.172     |
|    | $\sigma_1=4$     | 263.536    | 71.791     | 34.319     | 22.459     | 14.942     | 38.810     | 33.054     | 26.369     |
|    | $\sigma_1=5$     | 158.326    | 62.054     | 27.532     | 13.976     | 11.856     | 31.194     | 13.573     | 20.818     |
| BSF| $\sigma_1=1$     | 64.621     | 65.269     | 30.796     | 22.179     | 26.511     | 25.583     | 43.317     | 10.568     |
|    | $\sigma_1=2$     | 76.904     | 69.010     | 49.060     | 21.896     | 25.912     | 27.038     | 47.459     | 14.239     |
|    | $\sigma_1=3$     | 79.383     | 87.539     | 42.419     | 16.290     | 29.355     | 32.494     | 43.603     | 16.565     |
|    | $\sigma_1=4$     | 55.515     | 74.782     | 35.217     | 10.345     | 10.848     | 30.640     | 26.552     | 7.601      |
|    | $\sigma_1=5$     | 46.390     | 49.253     | 24.204     | 9.827      | 7.605      | 21.454     | 12.927     | 5.139      |

ability is very limited, and the SCRG values are usually small. In addition, if the $\sigma_1$ value is too small, the DoG kernel is sensitive to noise, so the BSF value of the GSM map is also small. On the contrary, when the $\sigma_1$ value is too large, the DoG kernel may mistakenly amplify some irrelevant background clutters when enhancing the target signal, so the SCRG and BSF values will also decrease. For Fig. 9(d), since the processed result of DoG kernel is almost no output when $\sigma_1=1$, the BSF value of the GSM map can get the largest value, while its SCRG value is relatively small. Overall, when $\sigma_1=3$, the GSM map can guarantee relatively good background suppression and target enhancement capabilities in most cases according to the BSF and SCRG values. Consequently, the parameter $\sigma_1$ is set to 3 in subsequent experiments.

C. Visual Comparisons

In order to assess the performance of our method, nine state-of-the-art algorithms are selected for comparative experiments, including the multidirectional improved top-hat filter (MITHF) [14], the reweighted infrared patch-tensor (RIPT) method [16], nonconvex rank approximation minimization (NRM) [17], absolute directional mean difference (ADMD) [23], and three-layer LCM (TLLCM) [25], the gradient direction diversity weighted multiscale flux density (GDDMFD) [27], the LIG [30], the facet kernel and random walker (FKRW) [32], and the fast adaptive masking and scaling with iterative segmentation (FAMSIS) [35]. The MITHF is a filtering-based method, and the RIPT and NRM are sparse-based methods. The ADMD, TLLCM, GDDMFD, and LIG are visual contrast-based methods, and the FKW and FAMSIS are segmentation-based methods. Among them, the GDDMFD, LIG, and FAMSIS are developed according to the gradient and intensity characteristics, which is similar to the basic assumptions of our algorithm.

Fig. 10 is the processed results of representative images by different algorithms. As shown in Fig. 10(b), the MITHF method measures the gray difference between inside and outside areas by constructing multidirectional structural elements, which can effectively suppress the edge clutters and highlight the target signal. Whereas, because the MITHF method does not consider the interference of noisy pixels well, for the hot pixels in Seq_6, a great number of false detections will occur in the resulting
map. The RIPT method has a good effect on suppressing the background components with strong autocorrelation, but for some complex scenes with many fluctuating clutters, the high-intensity sparse clutters will be mistakenly divided into the target image, such as Seq_3 and Seq_6–9. Furthermore, by introducing an extra regularization term with $l_{2,1}$ norm, the NRAM method can eliminate the strong residual edge clutters and noisy pixels more robustly, and its clutter suppression ability is much improved compared with the RIPT method. As can be seen from Fig. 10(e), the ADMD method shows a reasonable effect in removing sharp edge clutters by integrating the direction information. However, when the small target is buried in the broken background clutters, the ADMD method cannot stably highlight the target signal, and there may be a large number of false alarms in the processed results, such as Seq_2, 3, and 5–7. The TLLCM method is also based on visual contrast and has achieved better detection performance than the ADMD method by adding an alert layer, as shown in Fig. 10(f). Even so, some structural clutter also has high local contrast characteristics, so the processed results of ADMD and TLLCM methods will be greatly affected, as for Seq_2, 5, and 7.

As illustrated in Fig. 10(g), the GDDMFD method uses the divergence and structure tensor of the gradient vector to enhance the target signal, which can suppress the edge clutter well. In
addition, by combining the gradient and intensity information of small targets, the LIG method has a considerable effect on target enhancement and clutter removal in common scenes, but for Seq_4, 6, and 8, this method has poor adaptability to suppress high-intensity sharp edge clutter, and its performance needs to be further optimized. The segmentation-based FKRW method can achieve good target enhancement ability in the Seq_1–7, but it is difficult to characterize a small target signal in Seq_8 which contains a lot of high-intensity undulating sea clutter, resulting in missed detection. It can be observed in Fig. 10(j) that FAMSIS method can enhance small targets in all sequences by using gradient and intensity information to characterize the target signal after image segmentation. However, because the DoG filter directly used in FAMSIS method is sensitive to sharp edges, there are some edge clutter residues in the processed results of Seq_4 and Seq_8.

Nevertheless, from Fig. 10(k), it can be found that the proposed method can stably highlight small target signals with extremely low false alarm rate on all representative images. The success of this method is attributed to the reasonable representation and integration of the gradient and intensity characteristics of small target, so that the strong background clutter such as edge clutter, black hole, or corner clutter can be reliably eliminated one by one. First, according to the gradient characteristic of the small target, the horizontal and vertical derivatives of the image are calculated and fused, which can suppress sharp edge clutters and enhance the candidate targets, especially for Seq_4 and Seq_8. Afterward, according to the intensity characteristic of small target, the local intensity difference between each candidate target and the surrounding background is calculated, which can stably eliminate black holes, noisy pixels and strong corner clutters, as for Seq_6 and Seq_7. The proposed method fully integrates the advantages of the above two measures, so it can effectively highlight small targets and eliminate complicated background clutters and is superior to the compared algorithms in visual comparisons.

D. Quantitative Comparisons

Then, the SCRG, BSF, and ROC metrics are employed for quantitatively analyzing the effectiveness of the proposed algorithm. Table III lists the SCRG and BSF values of different algorithms on the test images in Fig. 9, where **bold** indicates the largest value and *italic-underlined* indicates the second largest value. It can be seen that the MITHF method can obtain large SCRG and BSF values on most images, but the values in Seq_6 are relatively small because this method is not effective in suppressing noise pixels. The RIPT and NRAM methods can obtain relatively large SCRG and BSF values when the background is stationary and the target contrast is high, such as for Fig. 9(a)–(c) and (e). By introducing an additional regularization term with $l^2,1$ norm, the background suppression effect of NRAM is more robust, such as eliminating noisy pixels in Seq_6 and strong edges in Seq_8. Overall, the SCRG and BSF values of the ADMD method are relatively low among the comparison algorithms, while the background suppression effect of the TLLCM method is better than that of ADMD, but it still needs to be improved. The LIG method can achieve considerable SCRG and BSF values for Fig. 9(a)–(c), (e), and (g), but its background suppression effect is not good for Fig. 9(d), (f), and (h) with sharp edge clutters. In addition, the SCRG and BSF values of the FKRW method in Fig. 9(a), (e), and (g) can reach relatively large values, but for Fig. 9(h), its target enhancement and clutter suppression performance are still not satisfactory. The FAMSIS method uses both contrast and gradient information to enhance target characteristics after image segmentation,
so it has better background suppression ability than ADMD, TCLLM, and GDDMFD methods in most situations. But it is also necessary to promote the background suppression effect of the FAMSIS method for Fig. 9(h). Due to the integration of GSM map and LISM map with a multiplication operation, the proposed algorithm has much higher SCRG and BSF values than other compared algorithms for all cases in Fig. 9, which indicates that the proposed algorithm is robust in target enhancement and clutter suppression.

Fig. 11 gives the ROC curves obtained by different algorithms on each sequence. It can be seen that in most sequences, the ROC curve of the ADMD method slowly reaches a stable value, and overall, the target detection capability of the TCLLM method is better than the ADMD method. In addition, in most cases, the ROC curves of RIPT, GDDMFD methods vary greatly among these comparison algorithms, so their target detection performance is unstable. The LIG method has similar assumptions to the proposed method, but its target detection ability and false alarm elimination effect are insufficient, especially in Seq_4 and Seq_8, and this can be seen in Fig. 11(d) and (h). The FKRW method can achieve better target detection performance with a lower false alarm rate in Seq_1, 5, and 7, and the FAMSIS method is more robust than most comparison algorithms in Seq_2 and Seq_6, but in other sequences, they are not the most effective in detecting small targets and eliminating false alarms. Compared with the above algorithms, the MITHF method and NRAM method can achieve better detection results in most cases. For Seq_2, the TPR values of the proposed method are smaller than those of the RIPT method when $1.230 \times 10^{-6} \leq \text{FPR} \leq 2.282 \times 10^{-6}$; for Seq_7, the TPR level of the proposed method is lower than FKRW at $\text{FPR} \leq 3.136 \times 10^{-6}$; for Seq_8, when $1.598 \times 10^{-6} \leq \text{FPR} \leq 2.433 \times 10^{-5}$, the TPR values of the proposed method are also lower than NRAM and GDDMFD methods. But in general, with the increase in FPR values, the TPR value of the proposed method rises faster and can reach a higher value than other comparison algorithms. Consequently, it can be concluded from the ROC curves that the proposed method has more obvious advantages in target detection accuracy and robustness for various complex scenes.

### Computational Efficiency

The second row of Table IV gives the operating complexity of different algorithms. In the proposed algorithm, the GSM operation for a specific pixel requires $O(L^2)$ cost, where $L$ is the size of local processing window, which is set to $5\sigma_1$ in this article. Accordingly, the operation cost of GSM calculation for the whole image is $O(MNL^2)$, where $M$ and $N$ are the width and height of the raw image. Then, assuming that there are $K$ candidate targets in the step of LISM calculation, and the side length of the guard window formed by each candidate target is $L_c = \frac{\text{MajorAxis}}{L_e}$, then there will be $O(L_c \log(L_c))$ cost in the sorting operation for each guard window, and the operation complexity of LISM calculation is $O(KL_c \log(L_c))$. Therefore, the total computational complexity of the proposed algorithm is $O(MNL^2 + KL_c \log(L_c))$. In addition, $m$ and $n$ represent the row
and column of the patch image in the RIPT and NRAM models, respectively. The average running time of different methods to process the test dataset is counted in the third row of Table IV. It can be seen that the ADMD method marked in bold can process an image at the fastest speed among these comparison algorithms, and the FAMSIS method marked with italic-underlined has the second fastest running speed. As shown in the table, the proposed method is slightly slower than the ADMD and FAMSIS methods in time consumption, but it saves more time cost compared with other methods. Therefore, the proposed method is also a relatively efficient target detection algorithm, which can be easily implemented in hardware systems to meet the application of real-time detection.

V. CONCLUSION

In this article, a robust and efficient small target detection method based on GISM is proposed to remove the complex background clutters one by one, such as edge clutter, black hole, or corner clutter. Hence, the core idea of this method is to fully explore the gradient and intensity characteristics to gradually eliminate background clutter and identify the small target. Depending on the gradient characteristic of the small target, the GSM map can effectively suppress sharp edge clutters and enhance candidate targets. Based on the intensity characteristic of small target, the LISIM map can highlight the target signal and remove noise pixels, black holes and corner clutters. Finally, the small target and background clutter can be well differentiated by using the LISIM weighted GSM map. Extensive experiments demonstrate that the proposed method can achieve considerable results in target detection and clutter suppression under various complex scenes. In addition, the experimental results also show that the proposed algorithm runs fast and has the potential of real-time processing.

REFERENCES

[1] W. Li, M. Zhao, X. Deng, L. li, L. Li, and W. Zhang, “Infrared small target detection using local and nonlocal spatial information,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 12, no. 9, pp. 3677–3689, Sep. 2019.

[2] L. Deng, J. Zhang, G. Xu, and H. Zhu, “Infrared small target detection via adaptive M-estimator ring top-hat transformation,” Pattern Recognit., vol. 112, 2021, Art. no. 107729.

[3] Y. Li, Z. Li, B. Xu, C. Dang, and J. Deng, “Low-contrast infrared target detection based on multiscale dual morphological reconstruction,” IEEE Geosci. Remote Sens. Lett., vol. 19, 2022, Art no. 7001905.

[4] L. Fang, W. Zhao, X. Li, and X. Wang, “A convex active contour model driven by local entropy energy with applications to infrared ship target detection based on multiscale dual morphological reconstruction,” IEEE Geosci. Remote Sens. Lett., vol. 19, 2022, Art no. 7001905.

[5] I. Song and S. Kim, “AVI.Net: A new pliable network with a novel metric for small-object segmentation and detection in infrared images,” Remote Sens., vol. 13, no. 4, 2021, Art. no. 555.

[6] D. Pang, T. Shan, W. Li, P. Ma, S. Liu, and R. Tao, “Infrared dim and small target detection based on greedy bilateral factorization in image sequences,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 13, no. 13, pp. 3394–3408, Jun. 1, 2020, doi: 10.1109/JSTARS.2020.2988822.

[7] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, “Max-Mean and max-mean filters for detection of small-targets,” in Proc. Signal Data Process. Small Targets, vol. 3809, 1999, pp. 74–83, doi: 10.1177/12.3604039.

[8] L. Yang, J. Yang, and K. Yang, “Adaptive detection for infrared small target under sea-sky complex background,” Electron. Lett., vol. 40, no. 17, pp. 1083–1085, 2004.

[9] X. Kong, L. Liu, Y. Qian, and M. Cui, “Automatic detection of sea-sky horizon line and small targets in maritime infrared imagery,” Infrared Phys. Technol., vol. 76, pp. 185–199, 2016.

[10] J. Zhao, H. Feng, Z. Xu, Q. Li, and H. Peng, “Real-time automatic small target detection using saliency extraction and morphological theory,” Opt. Laser Technol., vol. 47, pp. 268–277, 2013.

[11] X. Bai and F. Zhou, “Hit-or-miss transform based infrared dim small target enhancement,” Opt. Laser Technol., vol. 43, no. 7, pp. 1084–1090, 2011.

[12] X. Bai and F. Zhou, “Analysis of new top-hat transformation and the application for infrared dim small target detection,” Pattern Recognit., vol. 63, no. 6, pp. 2145–2156, 2010.

[13] X. Bai, F. Zhou, and T. Jin, “Enhancement of dim small target through modified top-hat transformation under the condition of heavy clutter,” Signal Process., vol. 90, no. 5, pp. 1643–1654, 2010.

[14] Y. Li and Z. Li, “Infrared maritime dim small target detection based on spatiotemporal cues and directional morphological filtering,” Infrared Phys. Technol., vol. 115, 2021, Art. no. 103657.

[15] C. 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.

[16] Y. Dai and Y. Wu, “Reweighted infrared patch-tensor model with both non-local and local priors for single-frame small target detection,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 10, no. 8, pp. 3752–3767, Aug. 2017.

[17] L. Zhang, L. Peng, T. Zhang, S. Cao, and Z. Peng, “Infrared small target detection via non-convex rank approximation minimization joint I2,1 norm,” Remote Sens., vol. 10, no. 11, 2018, Art no. 1821.

[18] D. Liu, Z. Li, B. Liu, W. Chen, T. Liu, and L. Cao, “Infrared small target detection in heavy-sky scene clutter based on sparse representation,” Infrared Phys. Technol., vol. 85, pp. 13–31, 2017.

[19] 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 Phys. Technol., vol. 71, pp. 347–355, 2015.

[20] S. Kim and J. Lee, “Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track,” Pattern Recognit., vol. 45, no. 1, pp. 393–406, 2012.

[21] S. Kim, Y. Yang, J. Lee, and Y. Park, “Small target detection utilizing robust methods of the human visual system for IRST,” J. Infrared, Millim., THz Waves, vol. 30, no. 9, pp. 994–1011, 2009.

[22] S. Aghazayyari, S. Moradi, and H. Talebi, “Small infrared target detection using absolute average difference weighted by cumulative directional derivatives,” Infrared Phys. Technol., vol. 101, pp. 78–87, 2019.

[23] S. Moradi, P. Moallem, and M. F. Sabahi, “Fast and robust small target detection using absolute directional mean difference algorithm,” Signal Process., vol. 177, 2020, Art no. 107727.

[24] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, “A local contrast method for small infrared target detection,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 52, no. 1, pp. 574–581, Jan. 2014.

[25] J. Han, S. Moradi, I. Faramarzi, C. Liu, H. Zhang, and Q. Zhao, “A local contrast method for infrared small-target detection utilizing a Tri-layer window,” IEEE Geosci. Remote Sens. Lett., vol. 17, no. 10, pp. 1822–1826, Oct. 2020.

[26] J. Han et al., “Infrared small target detection based on the weighted strengthened local contrast measure,” IEEE Geosci. Remote Sens. Lett., vol. 18, no. 9, pp. 1670–1674, Sep. 2021.

[27] D. Liu, L. Cao, Z. Li, T. Liu, and P. Che, “Infrared small target detection based on flux density and direction diversity in gradient vector field,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 11, no. 7, pp. 2528–2554, Jul. 2018.

[28] X. Bai and Y. Bi, “Derivative entropy-based contrast measure for infrared small-target detection,” IEEE Trans. Geosci. Remote Sens., vol. 56, no. 4, pp. 2452–2466, Apr. 2018.

[29] X. Cao, C. Rong, and X. Bai, “Infrared small target detection based on derivative dissimilarity measure,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 12, no. 8, pp. 3101–3116, Aug. 2019.

[30] H. Zhang, L. Zhang, D. Yuan, and H. Chen, “Infrared small target detection based on local intensity and gradient properties,” Infrared Phys. Technol., vol. 89, pp. 88–96, 2018.

[31] Y. Bi, X. Bai, T. Jin, and S. Guo, “Multiple feature analysis for infrared small target detection,” IEEE Geosci. Remote Sens. Lett., vol. 14, no. 8, pp. 1333–1337, Aug. 2017.

[32] Y. Qin, L. Bruzzone, C. Gao, and B. Li, “Infrared small target detection based on facet kernel and random walker,” IEEE Trans. Geosci. Remote Sens., vol. 57, no. 9, pp. 7104–7118, Sep. 2019.
[33] Y. Qin and B. Li, “Effective infrared small target detection utilizing a novel local contrast method,” IEEE Geosci. Remote Sens. Lett., vol. 13, no. 12, pp. 1890–1894, Dec. 2016.

[34] S. Huang, Z. Peng, Z. Wang, X. Wang, and M. Li, “Infrared small target detection by density peaks searching and maximum-gray region growing,” IEEE Geosci. Remote Sens. Lett., vol. 16, no. 12, pp. 1919–1923, Dec. 2019.

[35] Y. Chen, G. Zhang, Y. Ma, J. U. Kang, and C. Kwan, “Small infrared target detection based on fast adaptive masking and scaling with iterative segmentation,” IEEE Geosci. Remote Sens. Lett., vol. 19, 2022, Art. no. 7000605.

[36] L. Chen and L. Lin, “Improved fuzzy C-Means for infrared small target detection,” IEEE Geosci. Remote Sens. Lett., vol. 19, 2022, Art no. 7001405.

[37] J. Han, Y. Ma, J. Huang, X. Mei, and J. Ma, “An infrared small target detecting algorithm based on human visual system,” IEEE Geosci. Remote Sens. Lett., vol. 13, no. 3, pp. 452–456, Mar. 2016.

[38] X. Wang, G. Li, and L. Xu, “Infrared dim target detection based on visual attention,” Infrared Phys. Technol., vol. 55, no. 6, pp. 513–521, 2012.

[39] J. N. Kapur, P. Sahoo, and A. K. C. Wong, “A new method for gray-level picture thresholding using the entropy of the histogram,” Comput. Vis. Graph. Image Process., vol. 29, pp. 273–285, 1985.

[40] L. He, Y. Chao, K. Suzuki, and K. Wu, “Fast connected-component labeling,” Pattern Recognit., vol. 42, no. 9, pp. 1977–1987, 2009.

Yongsong Li received the B.S. degree in electronic information science and technology from the Zhengzhou University, Zhengzhou, China, in 2016, and the Ph.D. degree from the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing, China, in 2021.

He is a lecturer with the School of Artificial Intelligence, Chongqing University, Chongqing, China. His current research interests include infrared image processing, noise estimation, dim target detection, and pattern recognition for remote sensing.

Zhengzhou Li (Member, IEEE) received the B.S. degree in optoelectronics from the Northeastern University, Shenyang, China, in 1998, the M.S. degree in physics electronics and the Ph.D. degree in optical engineering from the Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu, China, in 2001 and in 2004, respectively.

He was a postdoctoral fellow with Department of Ophthalmology, Harvard Medical School, Boston, MA, USA. He is a Professor with the School of Microelectronics and Communication Engineering, Chongqing University, Chongqing, China. His research interests include target detection and tracking and image and signal processing for remote sensing.

Weite Li received the B.S. degree in software engineering from the Dalian University of Technology, Dalian, China, in 2016, the Master’s degree and Ph.D. degree in engineering from the Graduate School of Information Science and Engineering, Ritsumeikan University, Shiga, Japan, in 2022.

He is a Lecturer with the School of Artificial Intelligence, Chongqing University of Science and Technology, Chongqing, China. His current research interests include point cloud processing, transparent visualization, deep learning, and remote sensing.

Yuchuan Liu received the B.S. and Ph.D. degrees in engineering from Southwest University of Science and Technology, Chongqing University, Chongqing, China, in 2015 and 2021, respectively.

He is currently an Assistant Professor with the School of Intelligent Technology and Engineering, Chongqing University of Science and Technology, Chongqing, China. His research interests include transfer learning (include domain adaption), computer aided diagnosis and treatment, medical signal analysis, and image processing.