

Gaussian Scale-Space Enhanced Local Contrast Measure for Small Infrared Target Detection

Xuewei Guan^{ID}, Zhenming Peng^{ID}, Suqi Huang^{ID}, and Yingpin Chen

Abstract—Robust small-target detection plays an important role in the infrared (IR) search and track system, but it is still a challenge to detect small IR target under complex background. In this letter, an effective method inspired by the scale-space theory and the contrast mechanism of the human vision system is proposed. First, Gaussian scale-space (GSS) of an IR image is constructed by the convolution of a variable-scale Gaussian function. Second, the gray features of the local image can be directly represented by downsampling in a scale image, and enhanced local contrast measure (ELCM) is defined to enhance small target and suppress complex background. Then, the saliency map is obtained by using max-pooling operation, and an adaptive threshold is adapted to segment real targets. Experimental results on a test set with three real IR sequences demonstrate that the proposed method has a good performance in target enhancement and background suppression, and shows strong robustness under complex background. Especially, the proposed method has high computational efficiency, which can improve detection speed.

Index Terms—Human vision system (HVS), infrared (IR) target detection, local contrast, scale-space.

I. INTRODUCTION

SMALL-TARGET detection is one of the key techniques in infrared search and tracking (IRST) system [1]–[4]. The operating distance of the IRST system greatly depends on the detection capability of small target. At a long imaging distance, targets usually have no concrete shapes and textures with small size and are probably submerged in complex background and noise [3]. Typical IR small-target detection methods can be fallen into two categories: single-frame detection and sequential detection. Single-frame detection is the foundation of sequential detection and is usually a solution for many practical applications such as fast moving platforms [4]. Thus, we will focus on single-frame detection. Conventional single-frame detection methods, such as top-hat filter [5], can achieve good performance with a uniform background, but they have high false alarm ratio under complex background.

Manuscript received January 4, 2019; revised April 13, 2019; accepted May 6, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61571096 and Grant 61775030, in part by the Sichuan Science and Technology Program under Grant 2019YJ067, and in part by the Key Laboratory Fund of Beam Control, Chinese Academy of Sciences, under Grant 2017LBC003. (Corresponding author: Zhenming Peng.)

The authors are with the School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: guanxuewei2005@126.com.cn; zmpeng@uestc.edu.cn; huangsuci@uestc.edu.cn; 110500617@163.com.cn).

Color versions of one or more of the figures in this letter are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LGRS.2019.2917825

In order to overcome complex background interference, the methods based on the robust principal component analysis have been proposed, like the IR patch-image model [3] and the nonconvex rank approximation minimization (NRAM) [6]. These methods can reduce the false alarm rate in the complex background, but they are usually time-consuming. Anomaly detection methods, such as the cluster kernel Reed–Xiaoli (CKRX) algorithm [7], have been proven effective for detecting small targets, yet they are sensitive to abnormal pixels and time-consuming. Lately, a novel method via modified random walks (MRW) is proposed to detect small IR target with quite low signal-to-noise ratio [8]. However, it is still a difficult and challenging task to detect small IR target with high detection rate, low false alarm rate, and high speed under complex background.

In recent years, methods based on the contrast mechanism of human vision system (HVS) have been widely introduced to single-frame detection. Chen *et al.* [2] introduced the local contrast measure (LCM) to small IR target detection. Since then, several methods, including the multiscale patch-based contrast measure (MPCM) [9], the relative local contrast measure (RLCM) [10], and the high-boost-based multiscale local contrast measure (HBMLCM) [11], have been proposed. The common ground of these methods is to extract the local image region by a rectangular sliding window, which can be used to calculate the local contrast of the center location. They usually represent the gray features of local image by the mean gray, the maximal gray, or the average of maximal gray values. However, neither the mean value nor the maximal value can accurately represent the substantial gray features of local image. This problem may undermine the robustness of the local contrast methods and lead to false alarm and missed detection. In order to detect the small target with changing size, most of the local contrast methods use variable size window for multiscale detection, but these windows of different sizes are applied to the same image. Intuitively, it should be more reasonable that the window of a certain size is used to the corresponding scale image. In addition, how to define local contrast between the target and background is also important to effectively enhance small target and eliminate interference of various background clutter.

Scale-space theory provides a basic tool for representing and handling image structures at different scales and has a profound influence in computer vision [12], [13]. Based on the scale-space theory, the small target of changing size is detected in the Laplacian scale-space [1].

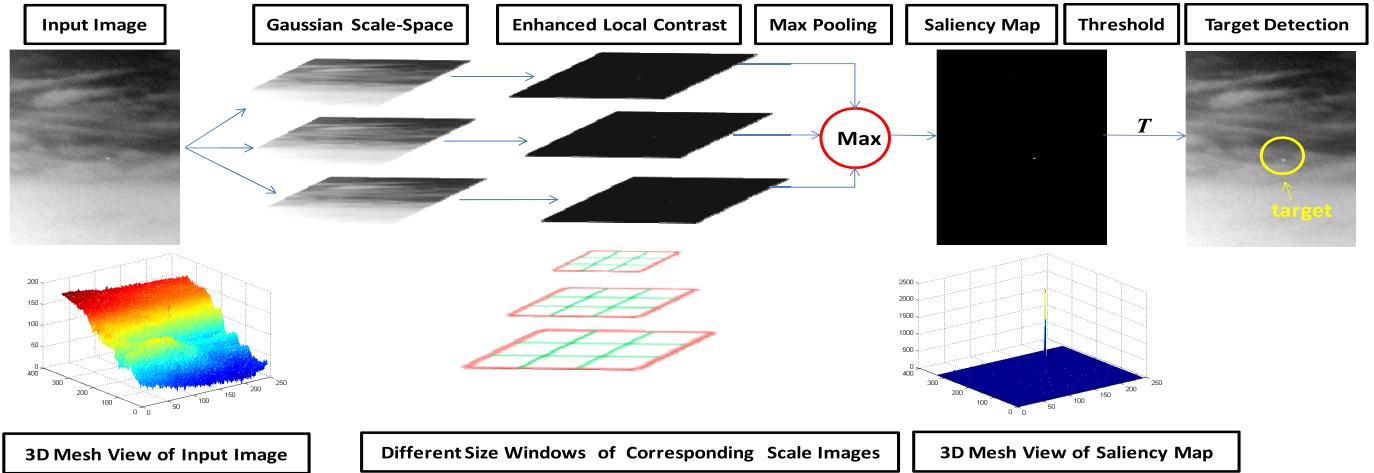


Fig. 1. Flowchart of the proposed small target detection method.

With these considerations, an IR small-target detection method based on the Gaussian scale-space enhanced local contrast measure (GSS-ELCM) is proposed in this letter. The proposed method calculates the local contrast in the GSS, where different size windows are used to the corresponding scale image and the gray features of local image can be represented more accurately by downsampling. ELCM is defined to enhance small target and suppress various background clutters more effectively. Moreover, the local extreme attribute of the small target is introduced to improve detection speed.

II. PROPOSED METHOD

The proposed small-target detection method has a pipeline structure. First, the GSS is constructed. Second, the enhanced local contrast is calculated. Then, the saliency map is obtained by using max-pooling, and the target is extracted by an adaptive threshold. The flowchart of the proposed GSS-ELCM method is shown in Fig. 1.

A. Gaussian Scale-Space

The scale-space theory is a general framework that has been developed for representing image data and the multi-scale nature of it at the very earliest stages in the chain of visual processing [12]. In the scale-space, an image can be represented as a set of different resolution images with the same spatial size [12], [13]. It has been demonstrated that under a variety of reasonable assumptions, the only possible scale-space linear kernel is the Gaussian function [12]–[14]. Consequently, a scale image $L(x, y, \sigma)$ can be produced from the convolution of a variable-scale Gaussian kernel, $G(x, y, \sigma)$

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y) \quad (1)$$

where $I(x, y)$ is an input image, \otimes represents the convolution operation in the pixel (x, y) , σ denotes a scale factor, and

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right). \quad (2)$$

The Gaussian kernel is a particularly efficient function to compute [14]. The 2-D Gaussian filter can be separated into the 1-D filter to reduce the computational complexity.

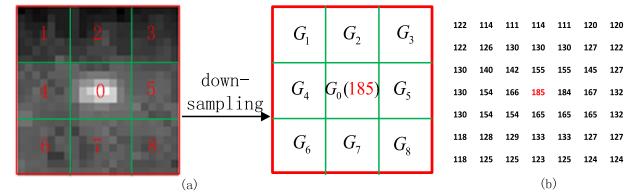


Fig. 2. (a) Local window region and downsampling. (b) Pixels gray of cell 0. It is well known that Gaussian filter can effectively suppress high-frequency components, such as pixel-sized noises with high brightness (PNHB) in IR image.

B. ELCM

Similar to LCM, ELCM can be calculated by a sliding window, which is of size $3d \times 3d$. The sliding window is divided into nine small cells, and the size of each cell is $d \times d$. In the GSS, we can use the sliding window to process the corresponding scale image with the factor σ . Because the Gaussian convolution can suppress most fine-scale information in the input image [12], the scale image with the factor σ can be downsampled by the step size d and the downsampled image retains substantial information of coarse-scale in the input image. This is true for the local window region, as for the whole image. Therefore, we can use the downsampling operation with a step size d to the local window region. That is to say, the gray feature of each cell can be represented by itself central pixel gray value in the scale image. In Fig. 2, the local window region with nine cells and downsampling are shown, where G_i ($i = 0, 1, \dots, 8$) denotes the gray feature of the i th cell. Taking G_0 as an example, the center pixel gray of cell 0 is 185, so we make G_0 equals 185. After that, the calculation of ELCM will be based on the nine downsampled points.

When the sliding window locates at a pixel (x, y) , considering that the small target in the local region has a higher gray value than the background, the target pixels belong to the extreme points. Therefore, the potential target pixels and background pixels can be discriminated fleetly by the local extreme attribute of the small target, as shown in (3). The cost of this check is reasonably low due to the fact that

most sample pixels will be eliminated following the first few checks:

$$\begin{cases} G_0 > \max\{G_1, G_2, \dots, G_8\} & (x, y) \text{ is potential target pixel} \\ \text{else} & (x, y) \text{ is background pixel.} \end{cases} \quad (3)$$

For a background pixel, we directly define the local contrast of that equals 0. For a potential target pixel, we need to further describe the local difference between the central cell and the surrounding cells. There are two definitions to describe the local difference: one is the difference form definition D , which can effectively eliminate highlight background and another is the ratio form definition R , which can enhance small target [10]. D and R are defined in the following equation:

$$D = \min(G_0 - G_i)^2 \quad i = 1, 2, \dots, 8 \quad (4)$$

$$R = \min \frac{G_0}{G_i} \quad i = 1, 2, \dots, 8. \quad (5)$$

Based on (3), there is $G_0 - G_i > 0$ for the potential target pixel. Then, taking into account the fact that both D and R are the increasing function of G_0 and the decreasing function of G_i , we can combine the two definitions together by product. Consequently, ELCM is defined as

$$\text{ELCM} = \min \frac{(G_0 - G_i)^2 G_0}{G_i} \quad i = 1, 2, \dots, 8. \quad (6)$$

Comparing D , R , and ELCM of the real target pixel with that of the background pixel, the small target is usually local salient and is not spatially correlated with the gray feature of the local neighborhood, yet the background is usually local continuous and is spatially correlated, we can get

$$\begin{cases} D_T > 0, \quad R_T > 1, \quad \text{ELCM}_T > 0 \text{ for a real target pixel} \\ D_B \approx 0, \quad R_B \approx 1, \quad \text{ELCM}_B \approx 0 \text{ for a background pixel.} \end{cases} \quad (7)$$

From (7), it can be seen that the real small target can be enhanced and the background can be suppressed by using the ELCM. In (6), ELCM is sensitive to information of different directions by using the minimum operation. Since background edges usually distribute along a particular direction in a local window region while the distribution of the small target in all directions is very similar, ELCM can effectively suppress complex background edges.

There are three points to distinguish ELCM from other methods based on local contrast. First, ELCM uses the down-sampled values in the scale image to represent the gray features of local window, whereas other methods represent the features by gray mean or gray maximal value. Second, by using the extreme attribute of the small target, the large area of the background can be eliminated in advance and ELCM is calculated only for the potential target pixels, which effectively reduces the computational complexity. Third, ELCM can enhance the small target while suppressing the background by combining the two definitions of local difference, and uses the square gray difference to suppress the high brightness background more effectively.

C. Saliency Map Calculation

In practice, the size of the small target is changing due to different imaging distances. To solve this problem, we calculate the ELCM in each scale image of the GSS and use maximum pooling on multiple scales ELCM to output a saliency map, because the maximum pooling is the scale invariance [15]

$$I_{SM}(x, y) = \max(\text{ELCM}_s(x, y)) \quad s = 1, 2, \dots, L \quad (8)$$

where I_{SM} is the saliency map, (x, y) denotes a pixel location, s represents the level number of scale, and L is the largest scale. It is crucially important how to determine the cell size d and the corresponding scale factor σ . For the ideal case, the size of each cell d should be approximated to the size of the small target h , as $d = h$. The typical small target presents as a Gaussian spot feature in IR image, and the size of target can be defined as the zero-crossing diameter of the second derivative of Gaussian function [1], [16], so the scale factor σ is determined by given small-target size h , as $h = 2\sqrt{2}\sigma$. And then, there is $d = 2\sqrt{2}\sigma$. A small target usually occupies several pixels, and spatial size is less than 9×9 . The appropriate cell sized should be no more than 9. In this letter, we choose $L = 3$ and $d = 3, 5, 7$, which can accommodate most small targets and can balance computational efficiency and detection performance.

D. Target Segmentation

In the saliency map, the most salient pixels will be likely real targets. Furthermore, the pixel values in the saliency map are assumed to be in accordance with the Gaussian distribution. Thus, it is easy to segment targets by an adaptive threshold T , which is defined as

$$T = \mu + k * \text{std} \quad (9)$$

where μ and std are the mean and standard deviation of the saliency map, respectively. The parameter k usually ranges from 50 to 70 is effective in our experiments.

III. EXPERIMENTS AND ANALYSIS

In this section, the proposed GSS-ELCM method is validated through experiments on a test set. First, we introduce the evaluation metrics. Then, experimental results and comparisons with other state-of-the-art methods are shown and analyzed.

A. Evaluation Metrics

To evaluate the performance of the proposed method, we introduce several common metrics. The signal-to-clutter ratio (SCR) gain and the background suppression factor (BSF) are two of them. The SCR, the GSCR (SCR gain), and the BSF are defined as

$$\text{SCR} = \frac{|P_t - \mu_b|}{\sigma_b}, \quad \text{GSCR} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}}}, \quad \text{BSF} = \frac{\sigma_{\text{bin}}}{\sigma_{\text{bout}}} \quad (10)$$

where P_t , μ_b , and σ_b are the gray peak of the target, the gray average of the background, and the standard deviation of

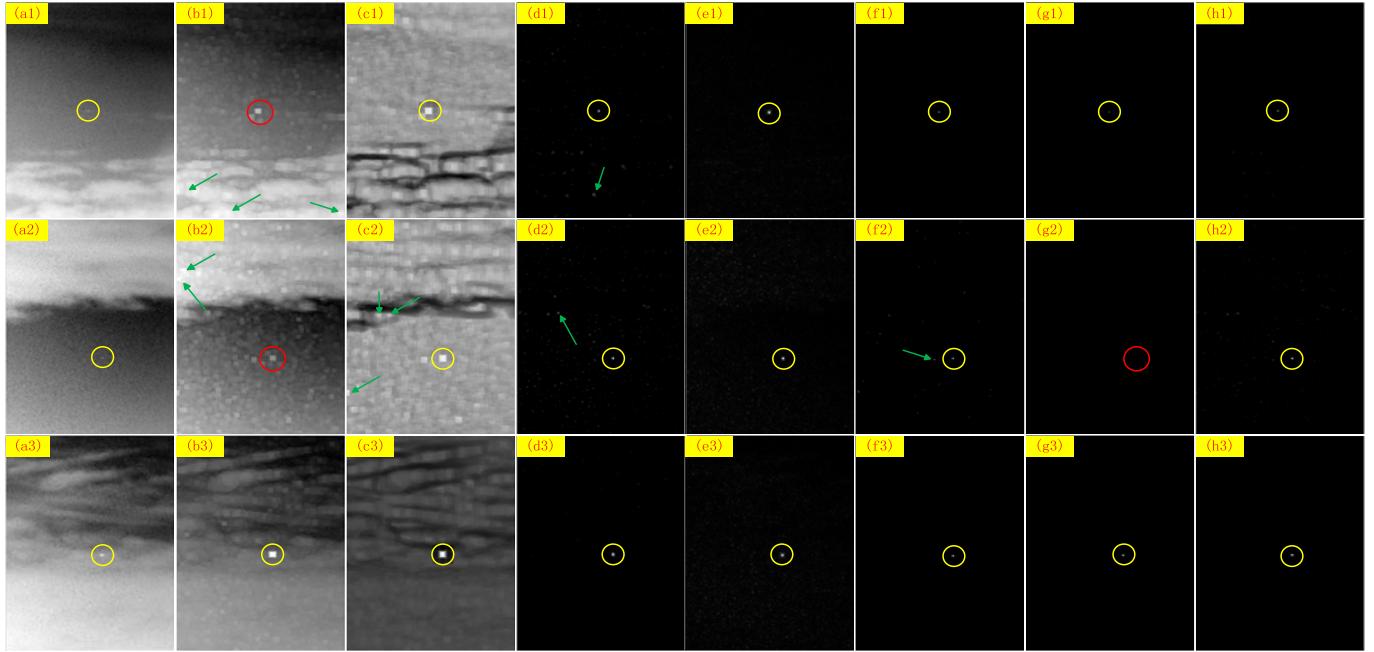


Fig. 3. Test images and detection results of different methods. (a1)–(a3) Test images of Seq. 1, Seq. 2, and Seq. 3, respectively. (b1)–(b3), (c1)–(c3), (d1)–(d3), (e1)–(e3), (f1)–(f3), (g1)–(g3), and (h1)–(h3) Saliency maps of LCM, RLCM, MPCM, HBLCM, MRW, NRAM, and the proposed method, respectively; yellow circle marks a true target; red circle represents an undetected true target; and green arrow denotes a false alarm location.

the background, respectively. SCR_{in} and SCR_{out} stand for the SCR of the input image and the processed image (the saliency map for the proposed method). σ_{bin} and σ_{bout} are the standard deviation of the background in the input image and the processed image, respectively. The other two metrics are the true positive rate (TPR) and the false positive rate (FPR), which are defined as

$$\text{TPR} = \frac{\text{number of detected true targets}}{\text{number of actual targets}} \quad (11)$$

$$\text{FPR} = \frac{\text{number of false pixels detected}}{\text{number of pixels in all test images}}. \quad (12)$$

B. Experimental Results and Comparisons

In our experiments, the test set with three real IR image sequences denoted by Seq. 1, Seq. 2, and Seq. 3, is used to demonstrate the effectiveness of the proposed method. These sequences are collected from the long-wave IR sensor, which was located on the ground to observe civil airplane in the sky. All three sequences contain a complex cloud background. The resolution of the images is 334×250 . The Seq. 1 contains 200 frames, and the target is above the clouds. The Seq. 2 contains 200 frames, and the target is below the clouds. The Seq. 3 contains 400 frames, and the target flies in the front of cloud background. The targets of Seq. 1 and Seq. 2 are extremely dim and small, with sizes from 2×2 to 3×3 . The target is slightly bright and large in Seq. 3, with changing size from 3×3 to 9×9 . Representative images of the three sequences are shown in Fig. 3(a1)–(a3), respectively. The input image in Fig. 1 is also taken from Seq. 3.

In Fig. 1, the 3-D mesh view of the input image and that of the saliency map are shown. We can observe that the small target is enhanced significantly by the proposed GSS-ELCM method. To further verify the validity of the

proposed method, we compared it with other six state-of-the-art methods, including LCM, RLCM, MPCM, HBLCM, MRW, and NRAM. We show the saliency maps and detection results of the six methods and the proposed method in Fig. 3.

In Fig. 3, it can be seen that the proposed method can effectively enhance small targets while suppressing the complex background and all targets are accurately detected without any false alarm in the three images. LCM produces a strong response at the high brightness background. When the target is very dim (possibly much lower brightness than the background), LCM can lead to a large number of false alarm generated and real targets missed, such as Fig. 3(b1) and (b2). RLCM can detect all three targets, but the saliency maps of it remain strong residual clutter in some background areas, and the false alarms emerge in Fig. 3(c2). Both MPCM and MRW can effectively suppress the background and detect the targets, but the false alarms still emerge in Fig. 3(d1), (d2), and (f2). HBLCM can detect all targets without false alarm, but a large amount of noise appears in its saliency maps. NRAM can suppress the background well, but it may over-shrink dim small target and cause missed detection, such as Fig. 3(g2).

The GSCR and BSF of seven different methods for three sequences are shown in Table I, where $\overline{\text{GSCR}}$ and $\overline{\text{BSF}}$ are the average GSCR and BSF in each sequence, respectively. It can be seen that our method achieves the largest GSCR and has much larger BSF than most of the other methods. Although the BSF of MPCM is slightly large in Seq. 3, we can observe that the GSCR of our method is about 2 to 3 times as much as that of MPCM. In addition, we observe that the BSF of NRAM is higher than that of other methods, including ours, in Seq. 2. However, NRAM would over-shrink dim small target; the GSCR of our method is still much larger than it.

TABLE I
GSCR AND BSF COMPARISON OF SEVEN DETECTION METHODS

Detection Methods	Evaluation indicators	Seq. 1	Seq. 2	Seq. 3
LCM	<u>GSCR</u>	2.0762	0.9872	2.2150
	<u>BSF</u>	1.0160	1.0076	1.0197
RLCM	<u>GSCR</u>	6.7078	12.1314	9.2557
	<u>BSF</u>	9.3261	7.2428	8.1094
MPCM	<u>GSCR</u>	162.3705	224.0110	233.4943
	<u>BSF</u>	33.8072	20.0715	23.9821
HBMLCM	<u>GSCR</u>	143.180	188.365	162.6438
	<u>BSF</u>	18.5438	12.1823	15.7431
MRW	<u>GSCR</u>	262.3705	312.3812	232.6837
	<u>BSF</u>	34.7322	18.1285	19.072
NRAM	<u>GSCR</u>	241.560	162.372	214.754
	<u>BSF</u>	31.8072	39.924	18.7618
Ours	<u>GSCR</u>	587.8744	451.3987	647.9121
	<u>BSF</u>	35.7520	21.9202	21.391

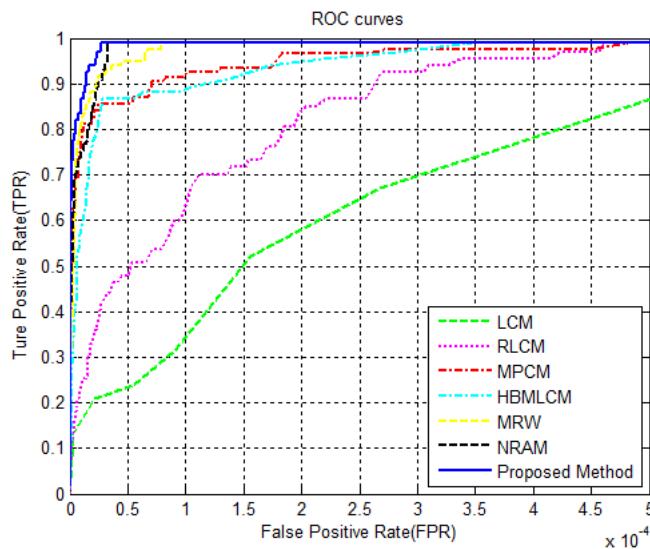


Fig. 4. ROC curves of seven detection methods for the test set.

TABLE II
TIME-CONSUMPTION COMPARISON OF SEVEN DETECTION METHODS

Method	LCM	RLCM	MPCM	HBM LCM	MRW	NRAM	Ours
Time(s)	10.478	16.130	9.911	1.682	6.442	22.241	1.154

To demonstrate the robustness of our method, Fig. 4 gives the receiver operating characteristic (ROC) curves of seven detection methods for the test set. From the ROC curves, the proposed method has higher TPR than the other methods under identical FPR and achieves the best detection performance.

Table II shows the average time-consumption comparison of seven methods in the test set. All experiments were conducted on a computer with an Intel i5 processor and 8 GB memory,

and all the code was implemented in MATLAB. It can be seen that our method is better than the other six methods in computational efficiency. This is attributable to three points: the Gaussian convolution is efficient to compute, the gray features of local image can be obtained by downsampling, and the ELCM is calculated only for the potential target pixels.

IV. CONCLUSION

In this letter, a GSS-ELCM method inspired by the contrast mechanism of HVS and the scale-space theory is proposed for small IR target detection. It is an improvement of the local contrast methods. The main idea of the proposed method is to define the ELCM in the GSS, which can enhance the small target effectively and overcome various background clutters. Experimental results demonstrate that the proposed method has a good performance for small IR target detection under complex background. Especially, in terms of detection speed, the method is outstanding and shows the potential for real-time application. In the future, we will further explore the multiscale information of the image, such as treating the local contrast images as an image cube to perform a joint detection.

REFERENCES

- [1] S. Kim, Y. Yang, J. Lee, and Y. Park, "Small target detection utilizing robust methods of the human visual system for IRST," *J. Infr. Millim. THz. Waves*, vol. 30, no. 9, pp. 994–1011, Sep. 2009.
- [2] 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.
- [3] 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.
- [4] C. Gao, T. Zhang, and Q. Li, "Small infrared target detection using sparse ring representation," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 27, no. 3, pp. 21–30, Mar. 2012.
- [5] V. T. Tom, T. Peli, M. Leung, and J. E. Bondaryk, "Morphology-based algorithm for point target detection in infrared backgrounds," *Proc. SPIE*, vol. 1954, pp. 2–11, Oct. 1993.
- [6] L. Zhang, L. Peng, T. Zhang, S. Cao, and Z. Peng, "Infrared small target detection via non-convex rank approximation minimization joint $l_{2,1}$ norm," *Remote Sens.*, vol. 10, no. 11, p. 1821, 2018.
- [7] J. Zhou, C. Kwan, B. Ayhan, and M. T. Eismann, "A novel cluster kernel RX algorithm for anomaly and change detection using hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6497–6504, Nov. 2016.
- [8] C. Xia, X. Li, and L. Zhao, "Infrared small target detection via modified random walks," *Remote Sens.*, vol. 10, no. 12, p. 2004, 2018.
- [9] Y. Wei, X. You, and H. Li, "Multiscale patch-based contrast measure for small infrared target detection," *Pattern Recognit.*, vol. 58, pp. 216–226, Oct. 2016.
- [10] 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.
- [11] Y. Shi, Y. Wei, H. Yao, D. Pan, and G. Xiao, "High-boost-based multiscale local contrast measure for infrared small target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 33–37, Jan. 2018.
- [12] T. Lindeberg, "Scale-space theory: A basic tool for analyzing structures at different scales," *J. Appl. Statist.*, vol. 21, no. 2, pp. 224–270, 1994.
- [13] T. Lindeberg, "Feature detection with automatic scale selection," *Int. J. Comput. Vis.*, vol. 30, no. 2, pp. 79–116, 1998.
- [14] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.
- [15] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, "Robust object recognition with cortex-like mechanisms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp. 411–426, Mar. 2007.
- [16] T. Netsch and H. O. Peitgen, "Scale-space signatures for the detection of clustered microcalcifications in digital mammograms," *IEEE Trans. Med. Imag.*, vol. 18, no. 9, pp. 774–786, Sep. 1999.