A Robust Infrared Small Target Detection Algorithm Based on Human Visual System

Jinhui Han, Yong Ma, Bo Zhou, Fan Fan, Kun Liang, and Yu Fang

***Abstract*—Robust human visual system (HVS) properties can effectively improve the infrared (IR) small target detection capabilities, such as detection rate, false alarm rate, speed, etc. However, current algorithms based on HVS usually improve one or two of the aforementioned detection capabilities while sacrificing the others. In this letter, a robust IR small target detection algorithm based on HVS is proposed to pursue good performance in detection rate, false alarm rate, and speed simultaneously. First, an HVS size-adaptation process is used, and the IR image after preprocessing is divided into subblocks to improve detection speed. Then, based on HVS contrast mechanism, the improved local contrast measure, which can improve detection rate and reduce false alarm rate, is proposed to calculate the saliency map, and a threshold operation along with a rapid traversal mechanism based on HVS attention shift mechanism is used to get the target subblocks quickly. Experimental results show the proposed algorithm has good robustness and efficiency for real IR small target detection applications.**

***Index Terms*—Human visual system (HVS), improved local contrast measure (ILCM), infrared (IR) small target.**

I. Introduction

In THE field of infrared (IR) small target detection, most targets are brighter than background, but the contrast between target and background is usually very low in a raw IR image and small weak target may be easily submerged in a bright background [1], which leads to a low detection rate. Meanwhile, due to the IR focal plane array material properties, as well as the electronic noises, there are usually many pixel-sized noises with high brightness (PNHB) in IR images [2], which are similar to targets and introduce a high false alarm rate. In addition, real-time target output is needed in many applications [3]; thus, detection algorithms with fast speed are much more popular.

In recent years, robust human visual system (HVS) properties, such as *size-adaptation process* [4], *contrast mechanism* [5], and *attention shift mechanism* [6], [7] have been introduced to IR small target detection [8]–[11], for they can effectively improve the detection capabilities, such as detection rate, false alarm rate, speed, etc. However, current algorithms based on HVS usually improve one or two of the aforementioned detection capabilities while sacrificing the others. For example, based on *size-adaptation process*, Wang *et al.* [8] segmented the whole IR image into small regions, in addition, *attention shift mechanism* was applied when they searched for candidate target regions, both of the aforementioned two steps can improve the detection speed, but they chose the brightest region as the candidate target region; thus, when the background is bright, they cannot achieve satisfied detection rate and false alarm rate. Based on *contrast mechanism*, Kim *et al.* [9] introduced the Laplacian of Gaussian (LoG) filter to the IR small target detection field to improve the image contrast, Shao *et al.* [10] combined the LoG filter with morphological processing, and Chen *et al.* [11] gave a new contrast measuring method named local contrast measure (LCM), all the three algorithms can improve the detection rate, but the LoG filter is hard to calculate and LCM is calculated pixel by pixel, which are both time consuming. Moreover, LCM made an excessive enhancement on PNHB, which results in a high false alarm rate.

In this letter, a robust IR small target detection algorithm based on HVS is proposed to pursuit a good performance in detection rate, false alarm rate, and speed simultaneously. It consists of two stages. In predetection stage, *a size-adaptation process* is used and the IR image after preprocessing is divided into subblocks to improve detection speed. In detection stage, based on *contrast mechanism*, the improved LCM (ILCM), which can improve detection rate and reduce false alarm rate, is proposed to calculate the saliency map, then a threshold operation along with a rapid traversal mechanism based on *attention shift mechanism* is used to get the target subblocks quickly.

This letter is organized as follows. In Section II, the predetection stage is introduced briefly. In Section III, the detection stage is described in detail. Experimental results are presented in Section IV, and this letter concludes in Section V.

II. Predetection Stage

HVS *size-adaptation process* means HVS can extract some salient regions at the scale adaptive to the target size for further process, which contributes to a fast detection speed. Target detection on a region level has been widely studied [8], [12]. In this letter, a similar method with the literature [8] is used first, i.e., slide a window on the IR image (after preprocessing) from up to down and left to right at a certain step, to obtain a series of subblocks. However, considering that PNHB usually emerges as single pixel, whereas a real IR small target usually has a small area (less than 9×9 according to the Society of Photo-Optical Instrumentation Engineers, SPIE. [13]), to utilize

Manuscript received February 26, 2014; revised April 4, 2014; accepted May 7, 2014. Date of publication May 22, 2014; date of current version June 20, 2014. This work was supported in part by the National Natural Science Foundation of China under Grant 61078062, Grant 61108074, and Grant 61275098 and in part by the Research Fund for the Doctoral Program of Higher Education of China under Grant 20100142120012.

The authors are with the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan 430074, China (e-mail: mayong@hust.edu.cn).

Digital Object Identifier 10.1109/LGRS.2014.2323236

the size difference between PNHB and real target, in this letter, we define the value of a subblock as the gray average of all pixels in it. Thus

$$m\left(\text{sblk}(s,t)\right) = \frac{1}{mb \times nb} \sum_{i} \sum_{j} I\left(\text{pix}(i,j)\right) \tag{1}$$

where $\text{sblk}(s,t)$ is a subblock, m is its value, mb and nb are the subblock sizes, $\text{pix}(i,j)$ is a pixel that belongs to $\text{sblk}(s,t)$, and $I(\text{pix}(i,j))$ is the gray value of $\text{pix}(i,j)$.

Form all the subblock values as a new matrix, which is denoted by M.

$$M(i,j) = m\left(\text{sblk}(i,j)\right), \quad i = 1,2,\ldots,MB, \quad j = 1,2,\ldots,NB \tag{2}$$

where MB and NB are the subblock numbers in vertical and horizontal, respectively. With a proper choice of window's size and its moving step, the element number in M will be far less than the pixel number in the IR image; thus, the detection speed can be improved significantly. It is necessary to point that the window's size should be approximated to the size of the small target, and the moving step is suggested to set to 1/2 of the window's side length.

III. Detection Stage

A. ILCM Calculation

A subblock may contain the whole (or the msost part of) small target or totally the background, which can be denoted by TB and BB, respectively (note that since the moving step is smaller than the subblock's side length, around a TB, there must be some subblocks overlapped to it but only containing a smaller part of the small target, which can be denoted by PB). Since real IR small targets are usually brighter than background, the simplest way to obtain TB is to apply a threshold operation on M directly [8]; however, the results may be easily influenced by bright background.

HVS *contrast mechanism* means that it is the contrast but not the brightness, which occupies the most important part in the streams of our visual system. This fact is true in the whole detection process, which will benefit finding the target from the whole IR image with a high detection rate. In this letter, based on *contrast mechanism* and derived from LCM [11], the ILCM is proposed. The calculation of ILCM is as follows.

If $\text{sblk}(s,t)$ is a subblock and its gray average is m_0 in M, L_n is the maximum of the gray value in $\text{sblk}(s,t)$, i.e.,

$$L_n = \max\left(I\left(\text{pix}(ij)\right)\right), \quad \text{pix}(i,j) \in \text{sblk}(s,t). \tag{3}$$

Apply an image patch whose side length is three times to the subblock's side length on the image and take $\text{sblk}(s,t)$ as the central subblock, and find the eight adjacent subblocks of $\text{sblk}(s,t)$ in this patch, as well as their gray average $m_1 \sim m_8$ in M. The ILCM of $\text{sblk}(s,t)$ will be defined as

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

Similar to LCM, if $\text{sblk}(s,t)$ is TB, usually $\max(m_i) < L_n$; thus, ILCM(TB) $> m_0$, then target can be enhanced. If $\text{sblk}(s,t)$ is BB, there may be $\max(m_i) \geq L_n$; thus, ILCM(BB) $\leq m_0$, then the background can be suppressed. It can be seen that after the ILCM calculation, target enhancement and background suppression can be achieved simultaneously, which is robust to resist bright background and will help to achieve a high detection rate.

It is necessary to point that it is hard to decide whether enhancement or suppression is applied on PB, but since PB only emerges around TB, and obviously for a same target ILCM(PB) $<$ ILCM(TB), it will not disturb the detection of TB. PB will be eliminated using the *inhibition-of-return mechanism* (see Section III-B).

There are two main differences between ILCM and LCM. First, ILCM is calculated for each subblock, whereas LCM is calculated pixel by pixel; thus, ILCM will have a faster speed than LCM. Second, both the maximum brightness (L_n) and the size information of the central subblock [shown in the calculation of m_0 in (1)] are used in ILCM, whereas LCM only uses the maximum brightness of the central cell. When a PNHB exists in a PB/BB (denoted by NPB/NBB), its influence on the mean estimation of the central subblock is weak; thus, it can be easily deduced that

$$\text{ILCM}_{\text{NBB}} < \text{ILCM}_{\text{TB}}, \qquad \text{ILCM}_{\text{NPB}} < \text{ILCM}_{\text{TB}}. \tag{5}$$

It means that ILCM is robust to resist the PNHB and can achieve lower false alarm rate than LCM (see Fig. 1), here, the small target is labeled in a rectangle and the PNHB with similar brightness to the target is labeled in a circle.

B. Traversal Mechanism and the Threshold Operation

For each element of M, calculate its ILCM according to (4), and form them as the saliency map SM. The ILCM of TB can be found by applying a simple threshold operation on SM, here, threshold Th is defined as

$$\text{Th} = \mu_{\text{SM}} + k\sigma_{\text{SM}} \tag{6}$$

where μ and σ is the mean and standard deviation of SM, respectively, k is a parameter, experimental results show that 10 to 20 is quite suitable in our work.

The subblocks with larger ILCM than Th are taken as TB, the target center (ti, tj), if needed, can be calculated as

$$ti = \frac{\sum_{i,j} i \times I\left(\text{pix}(i,j)\right)}{\sum_{i,j} I\left(\text{pix}(i,j)\right)}, \quad tj = \frac{\sum_{i,j} j \times I\left(\text{pix}(i,j)\right)}{\sum_{i,j} I\left(\text{pix}(i,j)\right)} \tag{7}$$

where $\text{pix}(i,j)$ is a pixel that belongs to TB, and $I(\text{pix}(i,j))$ is the gray value of $\text{pix}(i,j)$.

HVS *attention shift mechanism* means that the region with the most saliency will gain the highest priority (also called *pop-out phenomena* [8]), and regions been processed will not be processed repetitively (also called *inhibition-of-return mechanism* [11]), which is helpful to reduce the calculation redundancy, thereby further improving detection speed. Similarly, a

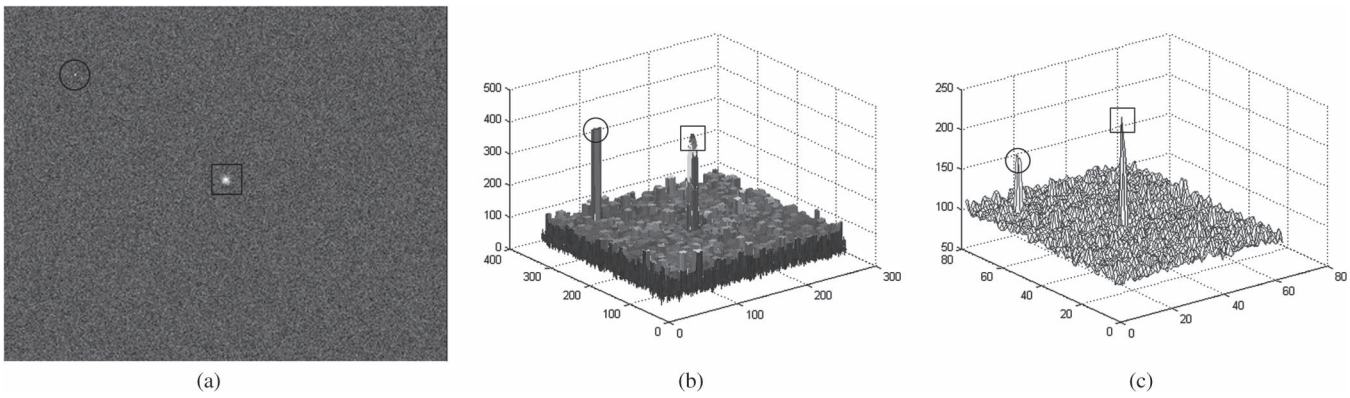


Fig. 1. Resistance to PNHB of ILCM versus LCM. (a). A simulation image that contains a small target and a PNHB. (b). The saliency map obtained by LCM. (c). The saliency map obtained by ILCM.

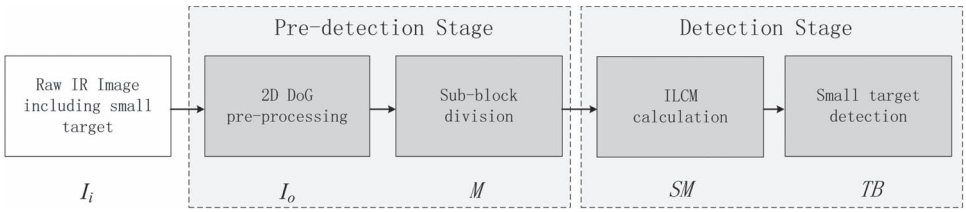


Fig. 2. Block diagram of the proposed target detection method.

winner-take-all mechanism is adopted to the traversal of all SM elements: the largest element will be processed first, then other elements will be processed from large to small until it is smaller than Th. Meanwhile, in order to avoid repetitive detection, once a TB is detected, the TB and its adjacent overlapped PBs will be all inhibited.

IV. Experimental Results and Analysis

Experiments on three real IR image sequences have been done. Each sequence contains 100 images obtained around Wuhan Tianhe International Airport using a cooled HgCdTe IR detector with a resolution of 320 × 256 and involved one aircraft with a small target flying in sunny weather (Seq. 1), one target flying in cloudy weather (Seq. 2), and one target submerged in bright cloud (Seq. 3). All experiments were implemented by MATLAB software on a PC with 4-GB memory and 3.3-GHz Intel i5 dual processor.

A. Experimental Results Using the Proposed Algorithm

The block diagram of the proposed target detection method is shown in Fig. 2. Here, the 2-D difference of Gaussians (DoG) filter [14] is chosen as the preprocessing method since it can enhance target and suppress background simultaneously and is very consistent with HVS *contrast mechanism*, too. For the sake of simplicity, the same approximate binomial kernel with literature [8] is adopted in this letter.

The experimental results using the proposed algorithm are shown in Fig. 3. Fig. 3(a) is the raw IR image I_i, for each sequence one image is selected, all the targets are labeled in rectangles. Fig. 3(b) is I_o after 2-D DoG preprocessing. Fig. 3(c) is the saliency map SM after ILCM calculation, here, the subblock size is set to 8×8, and its moving step is set to 4. It can be seen that the TB has the most saliency, which will help to improve detection rate and reduce false alarm rate. Fig. 3(d) is the detection result, here, k in (6) is set to 19. The proposed algorithm can extract TB precisely without any false alarms for all three images.

B. Comparison With Other Three Algorithms

To validate the detection effect of the proposed algorithm, we compare it with other three IR small target detection methods, including traditional top-hat algorithm, Wang's algorithm, and Chen's LCM algorithm. The same three images in Fig. 3(a) are used, and the comparison is shown in Fig. 4.

Fig. 4(a) is the detection result of top hat. It can successfully detect the target only when the sky is clean and the target is bright. Fig. 4(b) is the detection result of Wang's algorithm. It can successfully detect the targets for the first and the second image, but for the third image, it failed because the brightness of the background may be much higher than the real target. Fig. 4(c) is the detection result of Chen's LCM algorithm. It can detect all the three targets precisely, but false alarms emerge in all the three images as the former two algorithms, labeled in circles. Fig. 4(d) is the detection result of the proposed algorithm. All the three targets are successfully detected without any false alarms.

To further reveal the advantages of the proposed algorithm to the other three algorithms, the receiver operator characteristic (ROC) curve [15] is applied for the three sequences, i.e., for each algorithm, set its threshold to different values and draw the relationship between detection rate [true positive rate (TPR) defined in (8)] and false alarm rate [false positive rate (FPR) defined in (9)], shown in Fig. 5. From the figure, we can see that

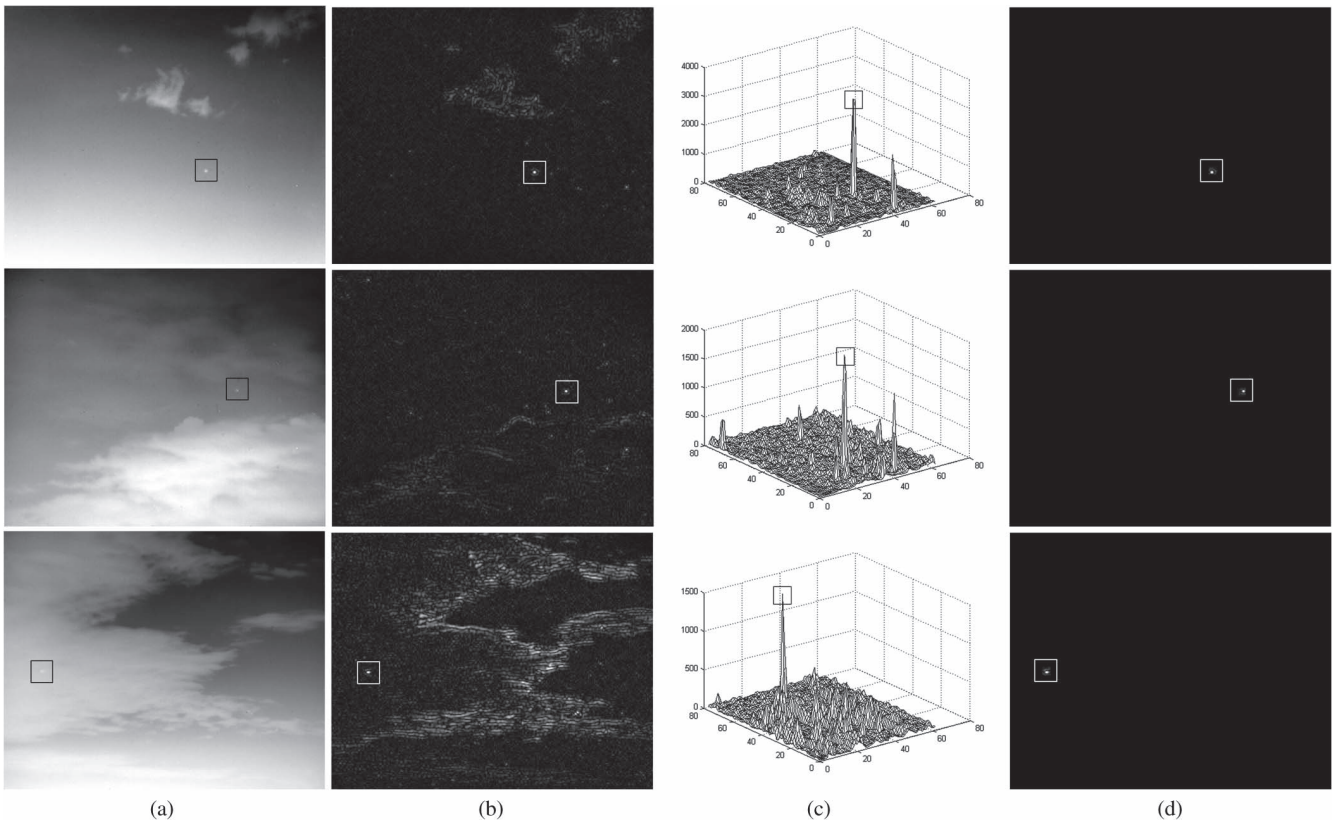


Fig. 3. IR small target detection results using the proposed algorithm. (a) The raw IR image I_i, for each sequence one image with representative is provided, and all the targets are labeled in rectangles. (b) I_o after 2-D DoG preprocessing. (c) Saliency map SM after ILCM calculation. (d) Detect results (without target center calculation).

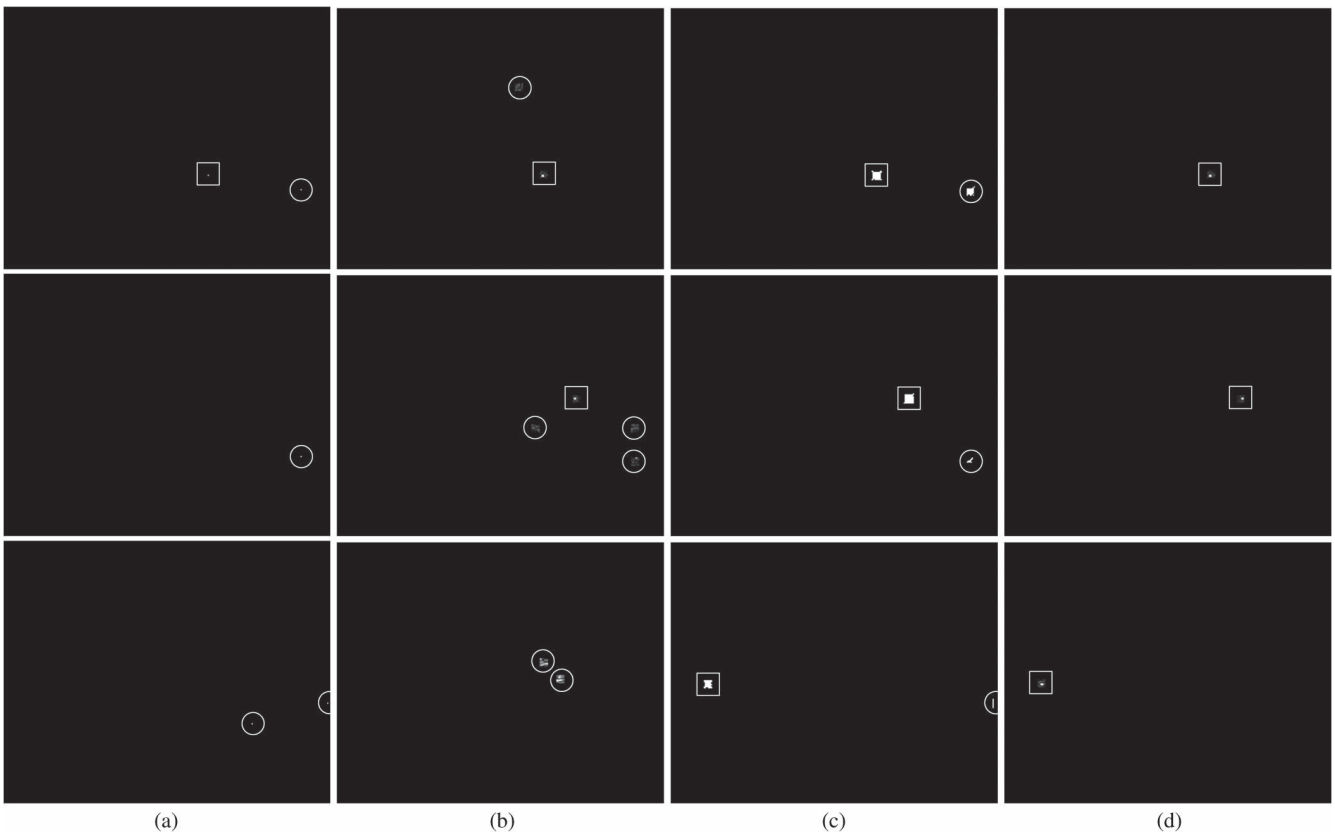
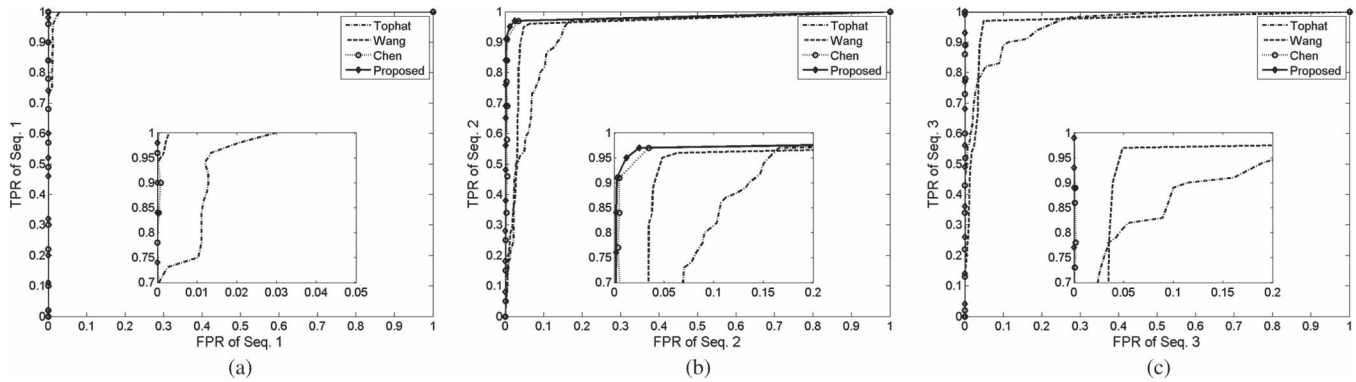


Fig. 4. Detection results of (a) top hat, (b) Wang's algorithm, and (c) Chen's LCM algorithm versus (d) the proposed algorithm. Here, the same three raw IR images with Fig. 3(a) are used.



(a) (b) (c)

Fig. 5. Comparison of ROC curves for (a) Seq. 1, (b) Seq. 2, and (c) Seq. 3.

TABLE I
COMPUTATIONAL COST COMPARISON AMONG THE PROPOSED ALGORITHM AND OTHER THREE ALGORITHMS FOR A SINGLE IMAGE (IN SECONDS)

	Top-hat	Wang	Chen	Our
Time consuming(s)	0.0102	0.0248	17.4135	0.0697

in most cases, the proposed algorithm can achieve the lowest FPR for a same TPR

$$\text{TPR} = \frac{\text{Quantity of true TB detected in images}}{\text{Quantity of true TB existing in images}} \times 100\% \tag{8}$$

$$\text{FPR} = \frac{\text{Quantity of false TB detected in images}}{\text{Quantity of BB and PB in images}} \times 100\%. \tag{9}$$

Finally, we compare the average computational cost of the four detection methods for a single image (see Table I).

From the experimental results, we can get the conclusion that top-hat method and Wang's algorithm can achieve fast detection speed; however, for a given TPR, their FPR may be a little high, particularly for Seq. 2 and Seq. 3. Chen's LCM algorithm can achieve better performance in TPR and FPR, particularly when the background is complex or bright, but its detection speed is too slow. The proposed algorithm, can achieve not only a fast detection speed but also the best performance in TPR and FPR in all the four methods.

V. CONCLUSION

In this letter, a robust IR small target detection algorithm based on HVS has been proposed. First, an HVS size-adaptation process is used, and the IR image after preprocessing is divided into subblocks to improve detection speed. Then, based on HVS contrast mechanism, the ILCM, which can improve the detection rate and reduce the false alarm rate, is proposed to calculate the saliency map, and a threshold operation along with a rapid traversal mechanism based on HVS attention shift mechanism is used to get the target subblocks quickly. Experimental results show that the proposed algorithm is robust to resist bright background and PNHB and can achieve a high detection rate in less than 0.07 s with a low false alarm rate. This algorithm can be either directly used in single-frame target detection or used as a foundation module in sequential target tracking for real-time applications.

REFERENCES

[1] K. Huang and X. Mao, "Detectability of infrared small targets," *Infrared Phys. Technol.*, vol. 53, no. 3, pp. 208–217, May 2010.
[2] B. Wang, S. Liu, Q. Li, and R. Lei, "Blind-pixel correction algorithm for an infrared focal plane array based on moving-scene analysis," *Opt. Eng.*, vol. 45, no. 3, pp. 364–367, Mar. 2006.
[3] 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.
[4] B. A. Olshausen, C. H. Anderson, and D. C. Van Essen, "A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information," *J. Neuroscience*, vol. 13, no. 11, pp. 4700–4719, Nov. 1993.
[5] R. VanRullen, "Visual saliency and spike timing in the ventral visual pathway," *J. Physiol. Paris*, vol. 97, no. 2/3, pp. 365–377, Mar.–May 2003.
[6] C. Koch and S. Ullman, "Shifts in selective visual attention: Towards the underlying neural circuitry," *Human Neurobiol.*, vol. 4, no. 4, pp. 219–227, 1985.
[7] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 11, pp. 1254–1259, Nov. 1998.
[8] X. Wang, G. Lv, and L. Xu, "Infrared dim target detection based on visual attention," *Infrared Phys. Technol.*, vol. 55, no. 6, pp. 513–521, Nov. 2012.
[9] 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. Terahertz Waves*, vol. 30, no. 9, pp. 994–1011, Sep. 2009.
[10] X. Shao, H. Fan, G. Lu, and J. Xu, "An improved infrared dim and small target detection algorithm based on the contrast mechanism of human visual system," *Infrared Phys. Technol.*, vol. 55, no. 5, pp. 403–408, Sep. 2012.
[11] 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.
[12] B. Du and L. Zhang, "Target detection based on a dynamic subspace," *Pattern Recognit.*, vol. 47, no. 1, pp. 344–358, Jan. 2014.
[13] S. Nilufar, N. Ray, and H. Zhang, "Object detection with DoG scale-space: A multiple kernel learning approach," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3744–3756, Aug. 2012.
[14] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves," in *Proc. 23rd Int. Conf. Mach. Learn.*, Pittsburgh, PA, USA, 2006, pp. 233–240.