An Efficient Infrared Small Target Detection Method Based on Visual Contrast Mechanism

Yuwen Chen and Yunhong Xin

***Abstract*—Robust and efficient detection of an infrared (IR) small target is very important in the IR search and track system. Based on the contrast mechanism of the human visual system, an IR small target detection method with high detection rate, low false alarm rate, and short processing time is proposed in this letter. This method consists of two stages. At the first stage, with the top-hat filter and an adaptive threshold operation based on the constant false alarm rate applied to the original image, the suspicious target regions are obtained. In this way, the computing time of the following steps would be reduced a lot; meanwhile, the desired and predictable detection probability with the constant false alarm probability is maintained. At the second stage, we first define a new efficient local contrast measure between the target and the background, and the local self-similarity of an image is introduced to calculate the local saliency map. With the combination of the local self-similarity and local contrast, an efficient saliency map is obtained, which cannot only increase the signal-to-clutter ratio but also suppress residual clutter simultaneously. Then, a simple threshold operation on the saliency map is used to get the true targets. Experimental results indicate that the proposed method is superior in detection rate, false alarm rate, and processing time compared with the contrast algorithms, and it is an efficient method for IR small target detection in a complex background.**

***Index Terms*—Detection, infrared (IR) small target, saliency map, visual contrast mechanism.**

I. Introduction

IT IS well known that infrared (IR) imaging technology has found wide applications in military, remote sensing, aerospace, etc. As an important technique in IR precision guidance, prewarning, video surveillance, search and tracking, and so on, the IR dim and small target detection has received much attention [1]–[5].

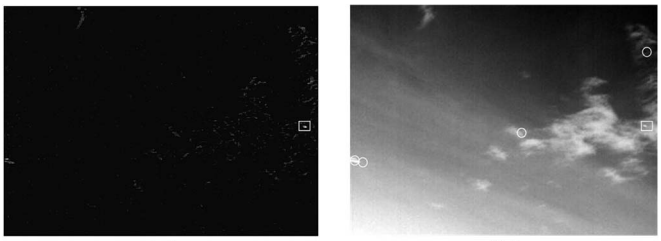
Generally, the IR target is small and occupies a few pixels in the image to show few distinct features [5], which is submerged in the background clutter easily. Although most targets are brighter than the background, the contrast between the target and the background is very low in a raw IR image. In addition, due to the influence of inherent sensor noise and natural factors, there are usually many pixel-sized noises which are similar to targets with high brightness in IR images. These effects make the IR target detection more difficult [6], [7]. Therefore, detecting IR small targets of unknown position and speed at low signal-to-clutter ratio (SCR) is a challenging task. Over the past few decades, many researchers have paid much attention to it and made much progress. The methods such as wavelet-transform-based algorithms [8], max-mean and max-median filters [9], matched filter [10], high-pass filter [11], Bayes estimation [12], and so on are widely used in the target detection.

In recent years, the human visual system (HVS) has been introduced to IR small target detection for getting favorable results. According to the human visual properties, a salient region where targets might exist can often attract attention of human eyes from complex background clutter quickly, and the robust properties of the HVS, such as contrast mechanism, multiresolution representation, size adaptation, and pop-out phenomena, show efficiency and robustness to improve the capabilities of the object detection [13]–[17]. Based on the principle of the HVS, there is a series of improved methods. Kim *et al.* [18] achieved the target signal enhancement and background clutter suppression by tuning and maximizing the SCR in Laplacian scale-space based on the HVS properties. Li *et al.* [19] computed the visual saliency map by using scale-space representation for target detection and can solve the problems caused by salient dark parts of targets. Wang *et al.* [20] introduced the difference of Gaussians filter to get salient regions and selected the true targets based on the visual attention with high detection precision and efficiency. Qi *et al.* [21] proposed a new directional saliency-based method incorporating with visual attention theory and showed great performance and robustness for IR target detection. Chen *et al.* [22] presented a new contrast measuring method named local contrast measure (LCM) to enhance the targets and improve the SCR of the image significantly. These methods are more effective than those traditional detection methods. However, some of them cannot get the satisfying results when it comes to the images with heavy noise and clutter. In addition, the computational complexity and time cost cannot be ignored in some methods. In this letter, an efficient method for IR small target detection is presented to pursue good performance in detection rate, false alarm rate, and computing time simultaneously. It consists of two stages. At the preprocessing stage, the suspicious target regions are screened out. In this way, the program running time is dramatically reduced, and it is very important for practical applications. At the detection stage, we utilize local self-similarity [23]–[25] as weight value and combine it with the local contrast to obtain a local saliency map (LSM), from which the targets are extracted out by a threshold operation which can remove the regions with small contrast. The experimental results show that our method has higher detection rate, lower false alarm, and less cost of time compared with four contrast detection algorithms, and it is an efficient method to the IR dim and small target detection in the complex background.

Manuscript received January 21, 2016; revised April 12, 2016; accepted April 14, 2016. Date of publication May 6, 2016; date of current version June 10, 2016.

The authors are with the School of Physics and Information Technology, Shaanxi Normal University, Xi'an 710119, China (e-mail: yuwench@snnu.edu.cn; xinyh@snnu.edu.cn).

Digital Object Identifier 10.1109/LGRS.2016.2556218



(a) (b)

Fig. 1. (a) Image after morphology filtering. (b) Suspicious pixels in the IR image after the preprocessing stage.

The rest of this letter is organized as follows. In Section II, the preprocessing stage for suspicious target regions is introduced. In Section III, we describe the detection stage in detail, including local contrast combined with similarity weight value for calculating saliency map and extracting true targets by the threshold operation. In Section IV, experimental results are presented. Finally, some concluding remarks are given in Section V.

II. Preprocessing Stage

In order to suppress the background clutter in the original image and reduce the cost of computation, we introduce a preprocessing stage to the raw image. At this stage, we choose the traditional top-hat morphological filter [26] to process the raw image first, with which the SCR of the image can be improved effectively. We set the structuring element as 3×7. Fig. 1(a) shows the image after morphology filtering and the true target labeled by a rectangle. Then, an adaptive threshold operation based on the constant false alarm rate (CFAR) [27], [28] is utilized, which is designed for target detection under complex clutter, aiming at acquiring the desired detection probability and maintaining constant false alarm probability. At last, we can obtain some possible target pixels. The threshold got from CFAR is defined as follows:

$$\text{Th} = \mu - \phi^{-1}(P)\sigma \tag{1}$$

where Th is the adaptive threshold, P is the false alarm probability, ϕ is the standard normal distribution function, and μ and σ are the mean value and the standard deviation of the preprocessed image, respectively. The size of the preprocessed image is $W \times H$. Let $f(i,j)$ be the gray value of the pixel at coordinate (i,j), and μ and σ can be calculated as

$$\mu = \frac{1}{W \times H} \sum_{1}^{W} \sum_{1}^{H} f(i,j) \tag{2}$$

$$\sigma^2 = \frac{1}{W \times H} \sum_{1}^{W} \sum_{1}^{H} \left[f(i,j) - \mu \right]^2. \tag{3}$$

The threshold is modified as follows:

$$\text{Th} = \mu - a\phi^{-1}(P)\sigma \tag{4}$$

where a is an adjustive coefficient. The pixel whose gray value is larger than the threshold is considered as a suspicious target pixel. By means of the adaptive threshold segmentation, we can get some suspicious target pixels and remove most background clutter.

III. Detection Stage

Human visual contrast mechanism means that it is the contrast between the target intensity and the background intensity, but not the brightness of the target, that plays an important role in human visual attention and the whole target detection process [18]. What we need to do is to obtain an efficient definition of local contrast to make targets prominent in the background. Due to the possible target, pixels have been selected at the preprocessing stage. Here, we only need to tackle these selected pixels. In the following, we describe the calculation of LSM and the threshold operation.

A. Calculating LSM

Based on the contrast mechanism and inspired by LCM [22], we proposed an efficient way to calculate the saliency map. Fig. 1(b) shows the selected candidate targets labeled by rectangles. The number of selected target pixels is N, the selected target pixels are named as p_k $(k = 1, 2, \ldots, N)$, and the region centered at p_k is named as V_0. Because a real target usually has a small area with a few pixels which is less than 80 pixels [29], the size of patch V_0 is not larger than $9 \times 9 = 81$ in the experiment. Let $I(i,j)$ be the gray value of pixel (i,j) and L_{p_k} be the maximum value of the gray values in V_0, i.e.,

$$L_{p_k} = \max\left(I(i,j)\right), I(i,j) \in V_0. \tag{5}$$

We define an image patch named V whose side length is three times to that of V_0 and V_0 is at the center of V. Then, around patch V_0, we can get eight patches named V_n $(n = 1, \ldots, 8)$ with the same size as V_0 in patch V. These adjacent regions of V_0 in patch V have their mid-value m_n $(n = 1, \ldots, 8)$, respectively

$$m_n = \text{median}\left(I_n(i,j)\right), I_n(i,j) \in V_n,\ n = 1, \ldots, 8. \tag{6}$$

Let M_{p_k} be the average value of these mid-values, and it can be obtained as

$$M_{p_k} = \frac{1}{N_v} \sum_{n=1}^{N_v} m_n \tag{7}$$

where N_v is the number of these mid-values, and it is equal to 8 here. Different from LCM [22] which adopts maximum average intensity of the surrounding patches to local contrast, we use average median intensity around patch V_0. In this way, the impacts of isolated background clutter with high brightness in the surrounding region are weaken, and the false alarm introduced by them is reduced. In addition, a weight value about the possibility of being detected as a true target is introduced to the saliency value. The local contrast value is defined as

$$C_{p_k} = W_{p_k} \cdot \frac{L_{p_k}^2}{M_{p_k}} \tag{8}$$

where W_{p_k} is the weight value related to the believability of suspicious target pixel p_k, which is applied to make the target region more salient. Here, we introduce the local similarity as the weight value. The similarity between two image patches can be easily revealed with a very simple method sum of square

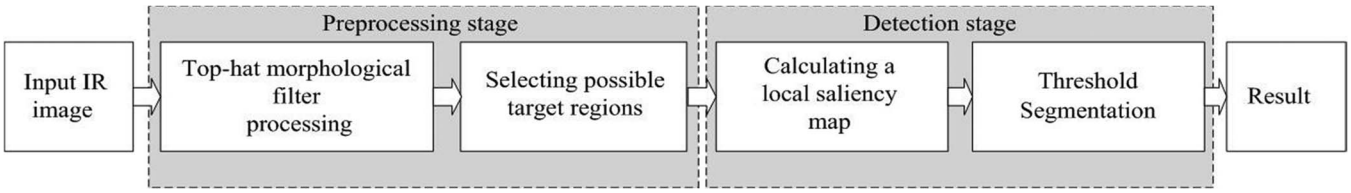


Fig. 2. Block diagram of the proposed method.

differences (SSD) [24]. The local similarity between patch V_0 centered at pixel p and its adjacent region is defined as follows:

$$\mathrm{SSD}_{p_k}(n) = \sum_{i,j} |V_0(i,j) - V_n(i,j)|^2, \ (n = 1, 2, \ldots, 8) \quad (9)$$

where (i, j) is the pixel position coordinate in an image patch V. The possibility that the region centered at p_k is the target region can be approximately calculated as

$$\bar{w}_{p_k} = \frac{1}{8} \sum_{n=1}^{8} \mathrm{SSD}_{p_k}(n) \quad (10)$$

where $\bar{w}_{p_k}(k = 1, 2, \ldots, N)$ is their believability measurements, respectively. $\bar{w}_{p_k}$ is normalized to [0, 1], and we can get the normalized believability W_{p_k}, which is defined as

$$W_{p_k} = \frac{\bar{w}_{p_k} - \bar{w}_{\min}}{\bar{w}_{\max} - \bar{w}_{\min}} \quad (11)$$

where $\bar{w}_{\max}$ and $\bar{w}_{\min}$ are the maximum and minimum values of $\bar{w}_{p_k}$, respectively. The local contrast values C_{p_k} constitute the LSM. Because the target exhibits the discontinuous features with surrounding background and it bears little resemblance to the background, the pixels which are more like the true target pixels would have larger weight values than those of pixels with less probability.

B. Threshold Segmentation

If the LSM about a set of the picked pixels is obtained, it is likely that the most salient region in the scene is a target region. In order to eliminate the rest impacts from noise and extract the targets accurately, a simple threshold operation is applied to the LSM. The threshold T is defined as follows:

$$T = g \times \max(\mathrm{LSM}) \quad (12)$$

where $\max(\mathrm{LSM})$ is the maximum value in the LSM and g is an adjustment parameter of the threshold. In our experimental work, g ranges from 0.5 to 0.9. If the local contrast value C_{p_k} of the pixel p_k were larger than threshold T, p_k would be selected as the pixel of a target.

IV. Experimental Results

Experiments on real IR images have been done. The simulations are all implemented by MATLAB (R2010a) on a personal computer with 2-GB memory, 2.94-GHz Intel Core2 processor.

A. Experimental Results Using the Proposed Method

The block diagram of the proposed IR small target detection method is shown in Fig. 2. In order to evaluate the performance of the proposed method, three sequence images (Seq1, Seq2, and Seq3) are used. Seq1 has 340 frames, Seq2 has 300 frames, and Seq3 has 240 frames. The size of the frame in each sequence is 384 × 288 pixels. In the experiments, we set false alarm P as 10^{-4}. To Seq1, Seq2, and Seq3, we set a as 3, 2.5, and 2, respectively, and set g as 0.6 for all three sequences. The targets are marked in the illustrations. Fig. 3(a) is the raw IR image picked from each sequence. Fig. 3(b) shows possible target pixels. Fig. 3(c) is the LSM. Fig. 3(d) is the detection result. It can be seen from Fig. 3 that the targets are detected by the proposed method without any false alarms in all of the three images apparently.

B. Comparison With Other Algorithms

To test the effectiveness and practicality of the proposed method, we compare it with the methods of top-hat, max-mean filter, high-pass filter, and Chen's LCM [22]. The same three images in Fig. 3(a) are used. The results are shown in Fig. 4. The true targets are labeled by rectangles, and the false targets are signed by circles. We can see that the high-pass filter can detect targets with a few false targets, except for the third picture in Fig. 4(a). The max-mean filter cannot detect the target submerged in the cloud in Fig. 4(b). Top-hat only performs well in the clean background with little noise and clutter in Fig. 4(c). LCM can detect the true targets successfully, but when the target is submerged in the clouds, the LCM presents false alarms in Fig. 4(d). Our method detected the true targets in the three images accurately without any false alarms in Fig. 4(e).

Statistical calculations of the detection rate and false alarm rate of the average probability of the whole images in a sequence are made to verify the advantages of our presented method numerically in (11) [6]. The results are listed in Tables I and II

$$\mathrm{DR} = \frac{\mathrm{TD}}{\mathrm{TE}} \times 100\% \quad \mathrm{DF} = \frac{\mathrm{FD}}{\mathrm{NT}} \times 100\% \quad (13)$$

where TD is the number of true targets detected in images and TE is the number of true targets that existed in images. FD is the number of false targets detected in images. NT is the number of targets detected in images. Finally, we compare the average computing time of the five methods for a single frame (Table III).

From the experimental results, it is shown that the high-pass filter and top-hat method are time-saving, but the false alarm rate is a little high, and the max-mean filter has a little low detection rate and reveals a number of false targets. Moreover, their detection effects depend a lot on the selection of the

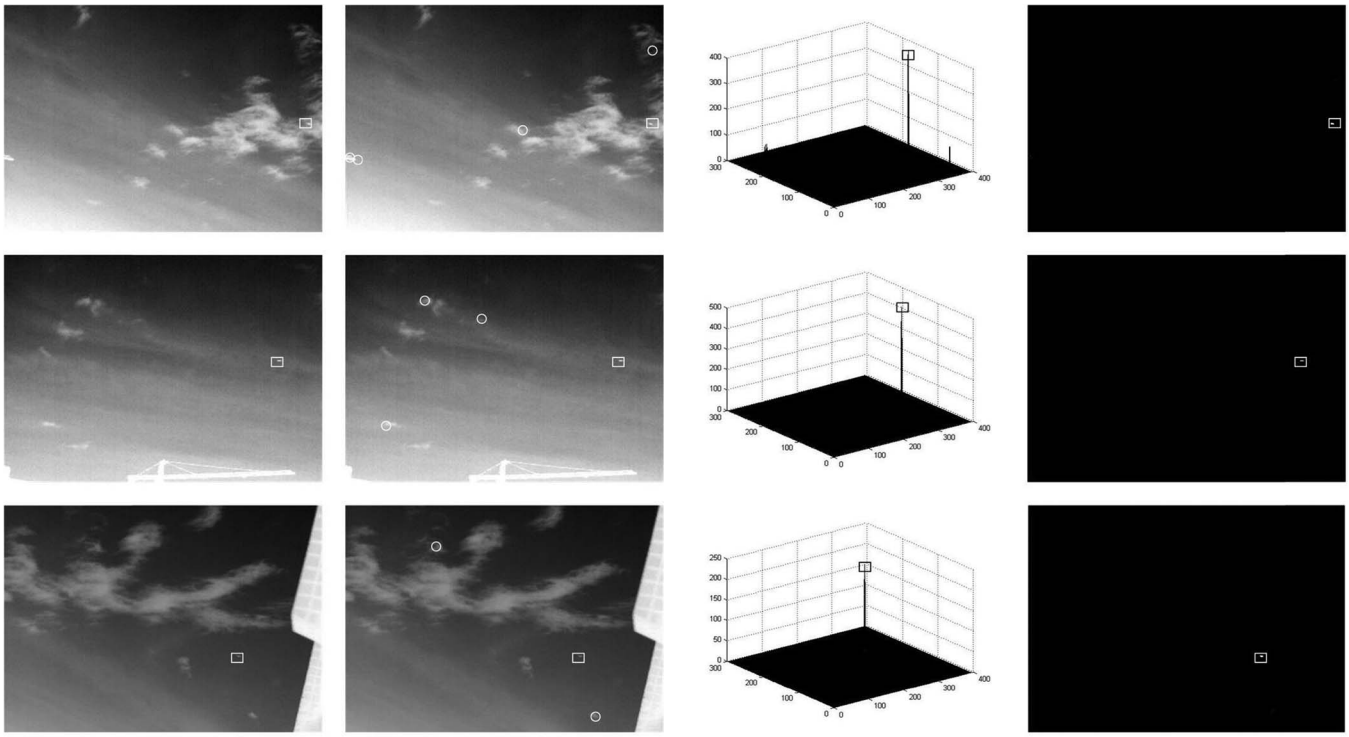


Fig. 3. IR small target detection results of the proposed method. (a) Raw IR images, from sequence 1–3, one representative is provided. (b) Selected possible target pixels after the preprocessing stage. (c) LSMs of the selected pixels. (d) Results of the detection stage.

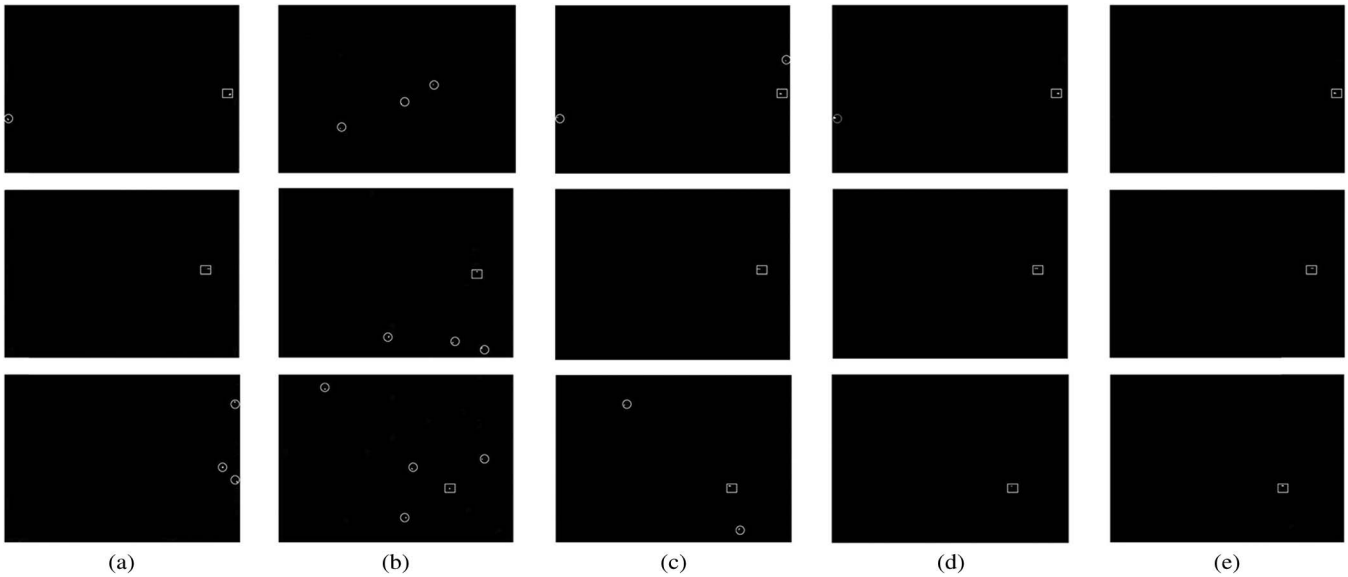


Fig. 4. Detection results of the (a) high-pass filter. (b) Max-mean filter. (c) Top-hat. (d) Local contrast method. (e) Proposed method; the same raw IR images in Fig. 3(a) are used.

TABLE I
DETECTION RATE OF DIFFERENT METHODS FOR SEQ1–3 (IN PERCENT)

Sequence	High-pass	Max-mean	Top-hat	LCM	Ours
1	50.57	30.27	57.21	89.42	97.54
2	65.43	45.55	67.53	90.31	100
3	45.91	32.54	50.48	85.66	96.83

TABLE II
FALSE ALARM RATE OF DIFFERENT METHODS FOR SEQ1–3 (IN PERCENT)

Sequence	High-pass	Max-mean	Top-hat	LCM	Ours
1	77.58	90.50	79.53	15.92	2.74
2	50.25	84.34	68.35	11.77	0.57
3	74.23	80.83	89.45	10.80	1.13

threshold because a larger threshold leads to a low detection rate and a smaller one leads to a high false alarm rate. LCM can get a low false alarm rate whenever the background is complex or clean; however, it costs much time. The proposed method has low computational complexity and good performance simultaneously among these methods.

TABLE III
COMPUTING TIME COMPARISON AMONG THE PROPOSED METHOD AND OTHER FOUR METHODS FOR A SINGLE IMAGE (IN SECONDS)

Sequence	High-pass	Max-mean	Top-hat	LCM	Ours
1	0.4726	1.0627	0.1082	22.9614	0.1788
2	0.4035	1.0433	0.1054	21.8535	0.1734
3	0.4789	1.1453	0.1136	22.9348	0.2003

V. CONCLUSION

In this letter, an efficient and robust IR small target detection method based on the human visual contrast mechanism has been proposed. The original image is preprocessed by a morphological filter and an adaptive threshold operation based on CFAR. In this way, the suspicious target regions are obtained, and the computational cost is reduced. Then, an efficient local contrast with weight, which can improve the detection rate and decrease the false alarm rate, is adopted to calculate an LSM, and a simple threshold operation is used to get the true targets. Experimental results show that the proposed method achieves satisfying performance on the aspects of detection rate, false alarm rate, and computing time. It is an efficient method to the IR dim and small target detection in the complex background.

REFERENCES

[1] 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, Jan. 2012.

[2] N. Thành, H. Sahli, and D. Hao, "Infrared thermography for buried landmine detection: Inverse problem setting," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 3987–4004, Dec. 2008.

[3] P. Wang, J. W. Tian, and C. Q. Gao, "Infrared small target detection using directional highpass filters based on LS-SVM," *Electron. Lett.*, vol. 45, no. 3, pp. 156–158, Jan. 2009.

[4] H. Ai, W. S. Cao, and R. F. Zong, "The algorithm research of infrared small target detection and recognition," in *Proc. Int. Conf. Meas., Inf. Control*, Aug. 2013, vol. 2, pp. 936–940.

[5] F. Zhang, C. F. Li, and L. Shi, "Detecting and tracking dim moving point target in IR image sequence," *Infrared. Phys. Technol.*, vol. 46, no. 4, pp. 323–328, Apr. 2005.

[6] 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.

[7] R. Liu, Y. Lu, C. Gong, and Y. Liu, "Infrared point target detection with improved template matching," *Infrared Phys. Technol.*, vol. 55, no. 4, pp. 380–387, Jul. 2012.

[8] F. A. Sadjadi, "Infrared target detection with probability density functions of wavelet transform subbands," *Appl. Opt.*, vol. 43, no. 2, pp. 315–323, 2004.

[9] S. D. Deshpande, M. H. Er, V. Ronda, and P. Chan, "Max-mean and max-median filters for detection of small-targets," *Proc. SPIE*, vol. 3809, pp. 74–83. Oct. 1999.

[10] K. Melendez and J. Modestino, "Spatiotemporal multiscan adaptive matched filtering," *Proc. SPIE*, Sep. 1995, vol. 2561, pp. 51–65.

[11] Y. Xu, "Small moving target detection in infrared image sequences," *Infrared Technol.*, vol. 24, no. 6, pp. 27–30, Jan. 2004.

[12] K. E. Matthews and N. M. Namazi, "A Bayes decision test for detection uncovered background and moving pixels in image sequences," *IEEE Trans. Image Process.*, vol. 7, no. 5, pp. 720–728, May 1998.

[13] J. H. Han *et al.*, "A robust infrared small target detection algorithm based on human visual system," *IEEE. Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2168–2172, Dec. 2014.

[14] E. Rahtu and J. Heikkila, "A simple and efficient saliency detector for background subtraction," in *Proc. 12th ICCV Workshops*, 2009, pp. 1137–1144.

[15] N. D. B. Bruce, X. Shi, and J. K. Tsotsos, "Recurrent refinement for visual saliency estimation in surveillance scenarios," in *Proc. IEEE 9th Conf. Comput. Robot Vis.*, 2012, pp. 117–124.

[16] S. Razakarivony and F. Jurie, "Discriminative autoencoders for small targets detection," in *Proc. IEEE 22nd Int. Conf. Pattern Recognit.*, 2014, pp. 3528–3533.

[17] H. J. Seo and P. Milanfar, "Visual saliency for automatic target detection, boundary detection, and image quality assessment," in *Proc. ICASSP*, 2010, pp. 5578–5581.

[18] 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.

[19] W. Li, C. H. P, and L. X. Liu, "Saliency-based automatic target detection in forward looking infrared images," in *Proc. IEEE Conf. Image Process.*, Nov. 2009, pp. 957–960.

[20] 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.

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

[22] 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.

[23] K. Chatfield *et al.*, "Efficient retrieval of deformable shape classes using local self-similarities," in *Proc. IEEE ICCV Workshops*, 2009, pp. 264–271.

[24] E. Shechtman and M. Ira, "Matching local self-similarities across images and videos," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2007, pp. 1–8.

[25] 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.

[26] M. Zeng, J. Li, and Z. Peng, "The design of top-hat morphological filter and application to infrared target detection," *Infrared Phys. Technol.*, vol. 48, no. 1, pp. 67–76, Apr. 2006.

[27] J. C. Hou and S. P. Dong, "Research of Adaptive Threshold of Constant False Alarm Detection," Syst. Eng. Electron., Res. Inst. Harbin Inst. Technol., Beijing, China, no. 7, pp. 8–12, 1994.

[28] L. Sevg, "Hypothesis testing and decision making: Constant-false-alarm-rate detection," *IEEE Antennas Propag. Mag.*, vol. 51, no. 3, pp. 218–224, Jun. 2009.

[29] S. Nilufar, N. Ray, and H. Zhang, "Object detection with DoG scalespace: A multiple kernel learning approach," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3744–3756, Aug. 2012.