An Infrared Small Target Detecting Algorithm Based on Human Visual System

Jinhui Han, Yong Ma, Jun Huang, Xiaoguang Mei, and Jiayi Ma

***Abstract*—Infrared (IR) small target detection with high detection rate, low false alarm rate, and multiscale detection ability is a challenging task since raw IR images usually have low contrast and complex background. In recent years, robust human visual system (HVS) properties have been introduced into the IR small target detection field. However, existing algorithms based on HVS, such as difference of Gaussians (DoG) filters, are sensitive to not only real small targets but also background edges, which results in a high false alarm rate. In this letter, the difference of Gabor (DoGb) filters is proposed and improved (IDoGb), which is an extension of DoG but is sensitive to orientations and can better suppress the complex background edges, then achieves a lower false alarm rate. In addition, multiscale detection can be also achieved. Experimental results show that the IDoGb filter produces less false alarms at the same detection rate, while consuming only about 0.1 s for a single frame.**

***Index Terms*—Human visual system (HVS), improved difference of Gabors (IDoGb), infrared (IR) small target.**

I. Introduction

IN SOME practical applications of the Infrared Search and Track (IRST) system, due to that the target is at a far distance, its projected image is usually very small with low contrast and does not have adequate structure information for detecting or matching [1]–[3], which makes the detection of IR small targets difficult [4]–[6]. Meanwhile, complex background edges may be falsely detected as targets, which results in a high false alarm rate [7], [8]. In addition, multiple targets with different scales may emerge in one frame at the same time; hence, multiscale detection ability is usually needed for detecting algorithms [9], [10].

In recent years, robust human visual system (HVS) properties, such as *contrast mechanism*, *size-adaptation process*, and *attention shift mechanism* [11], have been introduced to the IR small target detection field [12]–[17], for they are helpful to achieve good detecting results. For example, based on HVS, Kim *et al.* [13] and Shao *et al.* [14] used the Laplacian of Gaussian (LoG) filter to improve image contrast; Wang *et al.* [15] proposed the difference of Gaussians (DoG) filter, which is similar to LoG but is easier to construct; Chen *et al.* [16] used the ratio of the center block and the surrounding blocks as the local contrast measure (LCM); Xie *et al.* [17] replaced blocks of LCM with active pixels to construct the accumulated center-surround difference measure (ACSDM). All these algorithms can improve IR image contrast and then improve the detection rate, while keeping multiscale detection ability. However, LoG and DoG filters are sensitive to not only real small targets but also background edges, which results in a high false alarm rate; LCM and ACSDM algorithms cost too much time and are not suitable for real-time detection [18].

In this letter, an IR small target detecting algorithm based on HVS is proposed. Considering that most of the background edges (such as roof, wall, cloud edge, etc.) tend to spread along a certain orientation in a local small area, while real small targets usually spread equally in all orientations, the difference of Gabor (DoGb) filters, which is based on HVS *contrast mechanism* but is sensitive to orientations, is proposed and improved. The improved DoGb (IDoGb) filters can improve image contrast and then achieve a high detection rate and can better suppress complex background edges and then achieve a lower false alarm rate than DoG. In addition, multiscale detection can be achieved by IDoGb filter too, which is consistent with HVS *size-adaptation process*. In the last stage, HVS *attention shift mechanism* is adopted to output targets fast.

This letter is organized as follows: In Section II, the IDoGb processing for IR small target detection is discussed. In Section III, the target searching mechanism for the processed image is given. Experimental results are presented in Section IV, and this letter concludes in Section V.

II. IDoGB Processing

A. Definition of DoGb and IDoGb Filter

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. The DoG filter [15], which is based on HVS *contrast mechanism* and has been used widely in the target detection field, is defined as

Manuscript received May 13, 2015; revised August 11, 2015 and September 9, 2015; accepted September 30, 2015. Date of publication February 4, 2016; date of current version February 24, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 61503288, Grant 61078062, Grant 61108074, and Grant 61275098; in part by the Research Fund for the Doctoral Program of Higher Education of China under Grant 20100142120012; and in part by the China Postdoctoral Science Foundation under Grant 2015M570665 and Grant 2015M572194. *(Corresponding author: Yong Ma.)*

J. Han and X. Mei are with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan 430074, China.

Y. Ma, J. Huang, and J. Ma are with the School of Electronic Information, Wuhan University, Wuhan 430079, China (e-mail: mayong@whu.edu.cn).

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

Digital Object Identifier 10.1109/LGRS.2016.2519144

the difference of two 2-D Gaussian functions with different standard deviations, i.e.,

$$\begin{aligned}\mathrm{DoG}(x,y) &= G_1(x,y) - G_2(x,y) \\ &= \frac{1}{2\pi\sigma_1^2} e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{2\pi\sigma_2^2} e^{-\frac{x^2+y^2}{2\sigma_2^2}} \end{aligned} \tag{1}$$

where $\sigma_1 < \sigma_2$.

Generally, the DoG filter is a circularly symmetrical weighted function with positive center and negative surrounding. It can enhance target and suppress background, simultaneously, and then improve detection rate. In addition, σ_1 and σ_2 decide the size of the positive center; hence, multiscale detection can be achieved with variable choices of σ_1 and σ_2. In the frequency domain, the DoG filter can be taken as a bandpass filter, with G_1 and G_2 in (1) specifying the high and low cutoff frequency, respectively. However, DoG is not sensitive to orientation; hence, it cannot distinguish between real targets and complex background edges, which results in a high false alarm rate.

In this letter, considering that small targets and background edges usually have different directional features in a local small area, the DoGb filter, which is sensitive to orientation, is proposed. It is defined as the difference of two 2-D real Gabor functions.

The 2-D real Gabor function [19], [20] is defined as

$$Gb(x,y) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{x'^2}{2\sigma_1^2} - \frac{y'^2}{2\sigma_2^2}} \tag{2}$$

$$\begin{cases} x' = +x\cos\theta + y\sin\theta \\ y' = -x\sin\theta + y\cos\theta \end{cases} \tag{3}$$

where σ_1 and σ_2 are the scaling parameters of the filter and determine the effective size of the neighborhood of a pixel, in which the weighted summation takes place. θ specifies the orientation of the Gabor function.

Then, the DoGb filter is defined as

$$\begin{aligned}\mathrm{DoGb}(x,y) &= Gb_1(x,y) - Gb_2(x,y) \\ &= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{x'^2}{2\sigma_1^2} - \frac{y'^2}{2\sigma_2^2}} - \frac{1}{2\pi\sigma_3\sigma_4} e^{-\frac{x'^2}{2\sigma_3^2} - \frac{y'^2}{2\sigma_4^2}} \end{aligned} \tag{4}$$

where $\sigma_1 \le \sigma_2$, $\sigma_1 < \sigma_3$, and $\sigma_1/\sigma_2 = \sigma_3/\sigma_4 = Q$. In this letter, Q is set to 0.5. θ specifies the orientation of the DoGb filter. Note that normalizations for Gb_1 and Gb_2 are needed before the subtraction, to make sure that the sum of all weighting coefficients in DoGb is zero.

Comparing (1) and (4), it can be easily seen that DoG is a special case of DoGb, when $Q = 1$ and $\theta = 0$. Similar to the DoG filter, the DoGb filter has positive center and negative surrounding. In addition, multiscale detection can be achieved by DoGb filter with variable choices of $\sigma_1 - \sigma_4$ as well. In addition, the DoGb filter is sensitive to orientations and will have the ability to distinguish between real targets and background edges.

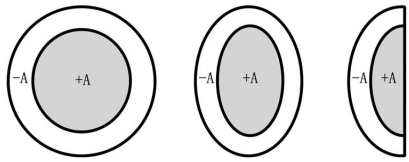

Fig. 1. Ichnography of (a) DoG, (b) DoGb ($\theta = 0$), and (c) IDoGb ($\theta = 0$).

However, the DoGb filter may still get nonzero processing results near background edges since it is symmetrical. To better suppress complex background edges, we make some improvements on the DoGb filter: taking the major axis of $\mathrm{DoGb}(x,y)$ as boundary, if (i,j) is at the right side of boundary, set $\mathrm{DoGb}(i,j)$ to 0 (see Fig. 1, where $+$A and $-$A are the sum of all positive and negative weighting coefficients, respectively).

B. IDoGb Processing for Edge Suppression

For a given raw input IR image I_{in}, the output image I_{out} after IDoGb processing will be calculated as

$$Itmp_i = |I_{\mathrm{in}} * \mathrm{IDoGb}_i|, \quad i = 1,2,\ldots,8. \tag{5}$$

$$I_{\mathrm{out}}(x,y) = \min\left(Itmp_i(x,y)\right), \quad i = 1,2,\ldots,8 \tag{6}$$

where IDoGb_i is the ith IDoGb template. Eight templates are used, corresponding $\theta = 0, \pi/4, \pi/2, \ldots, 7\pi/4$. $*$ denotes convolution. (x,y) is the pixel coordinate in the processed image.

Then, we will discuss the IDoGb processing result at the location (x,y), when it is pure background, background edge, and small target center, respectively.

1) (x,y) and its neighboring pixels are pure background with gray value X [see Fig. 2(a)]. The processing result will be independent to orientations, and it will be

$$I_{\mathrm{out}}(x,y) = |AX - AX| = 0. \tag{7}$$

2) (x,y) is near a background edge with smaller gray value X and larger gray value Y at each side [see Fig. 2(b)]. The minimum result of (6) will be

$$I_{\mathrm{out}}(x,y) = |AX - AX| = 0. \tag{8}$$

The processing result using the DoG filter is also shown in Fig. 2(b). It can be easily seen that the DoG processing result will be

$$\begin{aligned} I_{\mathrm{out}}(x,y) &= |AX - \alpha AX - (1-\alpha)AY| \\ &= (1-\alpha)A|X - Y| > 0 \end{aligned} \tag{9}$$

where α is a coefficient between 0 and 1.

Hence, it can be easily seen that IDoGb has a better suppression on background edges than DoG.

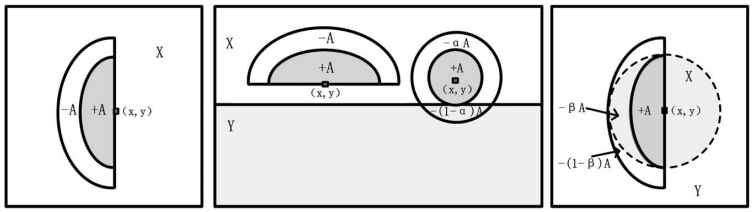

Fig. 2. Processing of (a) pure background, (b) background edge, and (c) target center using IDoGb filter. The processing of background edge using DoG filter is shown in (b) as well.

3) (x, y) is a small target center [see Fig. 2(c)]. The processing result will be

$$I_{\text{out}}(x,y) = |AX - \beta AX - (1-\beta)AY| = (1-\beta)A|X - Y| > 0 \quad (10)$$

where β is a coefficient between 0 and 1.

From the aforementioned discussions, we can get the conclusion that the IDoGb filter can simultaneously enhance target and suppress background and then improve the image contrast, which is consistent with HVS *contrast mechanism* and will be helpful to achieve a high detection rate. In addition, comparing with the DoG filter, the IDoGb filter can better suppress complex background edges, which is helpful to reduce the false alarm rate.

C. IDoGb for Multiscale Target Detection

Multiple targets with different scales may emerge in one IR image at the same time. The HVS *size-adaptation process* can locate all the targets at the scales adaptive to each target's size. For the IDoGb filter, multiscale target detection can be solved with variable choices of $\sigma_1 - \sigma_4$. Supposing that p scales are used, which means $4 \times p$ scaling parameters (σ_{s1}, σ_{s2}, σ_{s3}, σ_{s4}, $s = 1, 2, \ldots, p$) are needed, the procedure of multiscale detection using IDoGb filter can be described as follows.

1) For the sth scale, generate IDoGb templates IDoGb$_{si}$, $i = 1, 2, \ldots, 8$, using standard deviation $\sigma_{s1} - \sigma_{s4}$.
2) To better compare the processing results between different scales, normalize the sum of positive/negative weighting coefficients of IDoGb$_{si}$ from $+A_s/-A_s$ to $+1/-1$.
3) For the sth scale, calculate output image I_{outs} using the normalized IDoGb$_{si}$, $i = 1, 2, \ldots, 8$.
4) After all scales are calculated, using a maximum pooling between different scales for each pixel of the output image, we have

$$I_{\text{mout}}(x,y) = \max\left(I_{\text{outs}}(x,y)\right),\ s = 1, 2, \ldots, p \quad (11)$$

where (x, y) is the pixel coordinate, I_{outs} is the output image for the sth scale, and I_{mout} is the final output image for multiscale detection.

It is obvious that the maximum result of (11) will be given only when the scale of the IDoGb filter's positive center is approximate to (half of) the small target size.

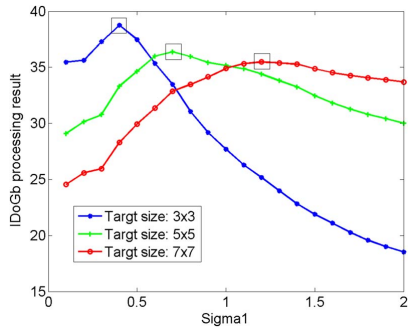

Fig. 3. Multiscale detection using IDoGb filter.

III. Small Target Extraction

HVS *attention shift mechanism* means that the region with the most saliency will gain the highest priority (also called *pop-out phenomena*) and regions which have been processed will not be processed repetitively (also called *inhibition-of-return mechanism*), which is helpful to achieve a fast detection speed. In this letter, we adopted HVS *attention shift mechanism* to the threshold operation for I_{out} to detect targets quickly, as follows.

1) Calculate the threshold Th of I_{out}, i.e.,

$$Th = \frac{k \max_{\text{out}} + \min_{\text{out}}}{k+1} \quad (12)$$

where $\max_{\text{out}}$ and $\min_{\text{out}}$ is the maximum and minimum of I_{out}, respectively, and k is a given parameter. Our experiments show that the optimal range of k is from 10 to 20 for single target detection.
2) Find the maximum value max and its location (i, j) in I_{out}.
3) If max is larger than Th, output (i, j) as a small target center.
4) Inhibit (i, j) and its neighbor small area with 0, i.e.,

$$I_{\text{out}}(i-r : i+r, j-r : j+r) = 0 \quad (13)$$

where r is the inhabitation radius. Note that we inhibit a small area around the target center because a small target usually has a small surface. It is suggested to set r to 4 since small targets are usually not larger than 9×9.
5) Find the next maximum value max and its location (i, j) in I_{out}; return to step 3), until max is smaller than Th.

In this stage, step 2) and step 5) simulate *pop-out phenomena*, step 4) simulates *inhibition-of-return mechanism*, and both of them are helpful to improve the detection speed.

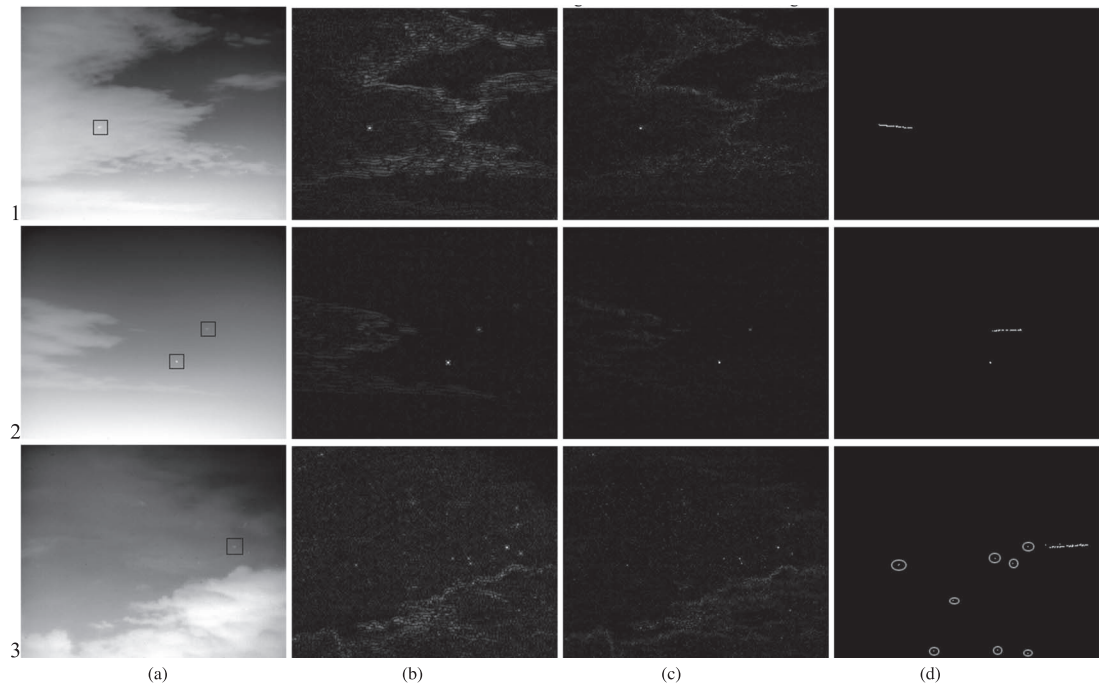

Fig. 4. IDoGb versus DoG processing result for three real IR sequences. (a) First frame of Seq. 1, Seq. 2, and Seq. 3. (b) Processing result using DoG. (c) Processing result using IDoGb. (d) Final detected target traces for the whole sequences using IDoGb. False alarms are labeled with circles.

IV. Experimental Results and Discussions

A. IDoGb for Multiscale Target Detection

Fig. 3 gives the simulating result for multiscale detection using the IDoGb filter, totally three scales are tested. For each scale, σ_1 is set from 0.1 to 2, σ_2 is set to $2 \times \sigma_1$ (i.e., Q is set to 0.5), while σ_3 and σ_4 are both set to $+\infty$ (in fact, similar to DoG, the IDoGb filter can be taken as a bandpass filter too, with Gb_1 and Gb_2 in (4) specifying the high and low cutoff frequency, respectively. To highlight the target regions fully, σ_3 and σ_4 are suggested to set to $+\infty$. Detailed introductions for frequency domain can be found in [15]), and the template size of IDoGb filter is set to 9×9, for all the three scales.

It is shown in Fig. 3 that, for a smaller target, the maximum processing result will be given when σ_1 is smaller, and vice versa. In particular, the simulating results in Fig. 3 also show that the optimal range of σ_1 is between 0.4 and 1.2 for single-scale detection when the target size is unknown.

B. IDoGb Processing for Real IR Sequences

Fig. 4 gives the processing results of IDoGb filter for three real IR image sequences. The sequences are obtained around Wuhan Tianhe International Airport using a cooled HgCdTe IR detector with a resolution of 320×256, and all the small targets are flying airplanes. Seq. 1 (180 frames) contains one bright target in heavy cloud; Seq. 2 (150 frames) contains two targets in sunny sky, i.e., target 1 is moving and weak while target 2 is stationary and bright; Seq. 3 (100 frames) contains one weak target with heavy cloud. The template size of IDoGb filter is set to 5×5, σ_1 and σ_2 are set to 0.85 and 1.7, respectively, while σ_3 and σ_4 are both set to $+\infty$. r is set to 4, and k is set to 15 for Seq. 1/Seq. 3 and 0.24 for Seq. 2 since the moving target in Seq. 2 is too much weaker than the stationary one. The DoG processing results are shown in Fig. 4 as well, where σ_1 and σ_2 are set to 1.7 and $+\infty$, respectively. It is shown in Fig. 4 that the IDoGb filter can suppress background edges better than the DoG filter, while improving real targets, which will be helpful to achieve a high detection rate and a low false alarm rate.

TABLE I
Image SCR and BSF Comparison Between DoG and IDoGb

Sequence	1	2		3
Target	only 1	target 1	target 2	only 1
Original SCR	0.5977	0.4218	1.1008	0.8347
SCR after DoG	19.4023	28.5683	72.0821	20.2087
SCR after IDoGb	36.2874	43.7736	117.8996	34.8339
SCR_g	1.8703	1.5322	1.6356	1.7237
BSF	1.4203	1.1374		1.1427

To better illustrate the efficiency of the IDoGb filter, Table I gives the image SCR and background suppress factor (BSF) before and after DoG/IDoGb processing for the first frame of the three sequences. Here, SCR, SCR_g, and BSF are defined as

$$\text{SCR} = \frac{|I_t - \mu_{\text{nb}}|}{\sigma_c} \tag{14}$$

$$\text{SCR}_g = \frac{\text{SCR after IDoGb}}{\text{SCR after DoG}} \tag{15}$$

$$\text{BSF} = \frac{\sigma_c \text{ after DoG}}{\sigma_c \text{ after IDoGb}} \tag{16}$$

where I_t is the maximum gray value of target, μ_{nb} is the average gray value of the target's neighboring background, and σ_c is the background standard deviation in the whole image.

In Fig. 4 and Table I, it is shown that, in the raw IR images, signal-to clutter ratio (SCR) is usually very low. DoG filter can improve SCR to some extent, but background edges are also enhanced. IDoGb filter can improve SCR more significant than DoG, while achieving a better suppression on background edges.

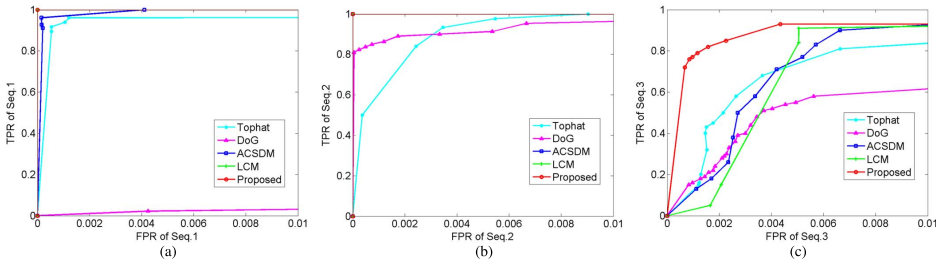

Fig. 5. Comparison ROC curve with other four algorithms for (a) Seq.1, (b) Seq. 2, and (c) Seq. 3.

TABLE II
TIME CONSUMING COMPARISON

	Tophat	DoG	ACSDM	LCM	Proposed
Time(s)	0.0062	0.0025	0.5981	1.8904	0.1072

C. Comparison With Other Algorithms

Fig. 5 gives the receiver operating characteristic (ROC) curve comparisons with the other four IR small target detection algorithms, including the Tophat, DoG, LCM, and ACSDM. The ROC curve represents the varying relationship of the false positive rate [FPR, defined as (17)] and true positive rate [TPR, defined as (18)]. A better algorithm produces less false alarms at the same detection rate. Thus

$$\text{FPR} = \frac{\text{number of detected false targets}}{\text{number of total pixels in an image}} \tag{17}$$

$$\text{TPR} = \frac{\text{number of detected true targets}}{\text{number of total true targets}}. \tag{18}$$

Table II gives the comparison of the average time consuming for a single frame. All experiments are implemented by MATLAB software on a PC with 8-GB memory and 4-GHz Intel i7 processor.

In Fig. 5 and Table II, it is shown that Tophat and DoG can achieve fast detection speed, but their performances in TPR and FPR may be not satisfied. ACSDM and LCM consume too much time, although they can achieve better performances in TPR and FPR. The proposed IDoGb filter can achieve the best performance in the five algorithms, while consuming only about 0.1 s for a single frame, which is far less than ACSDM and LCM.

V. CONCLUSION

In this letter, based on HVS, the DoGb filter and its improvement, the IDoGb filter, have been proposed for IR small target detection. The IDoGb filter can improve image contrast, while suppressing background edges. In addition, multiscale detection can be achieved by different choices of parameters. Experimental results show that the IDoGb filter produces less false alarms at the same detection rate, while consuming only about 0.1 s for a single frame.

REFERENCES

[1] J. Ma *et al.*, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, Dec. 2015.

[2] X. Dong, X. Huang, Y. Zheng, S. Bai, and W. Xu, "A novel infrared small moving target detection method based on tracking interest points under complicated background," *Infrared Phys. Technol.*, vol. 65, pp. 36–42, Jul. 2014.

[3] J. Ma, J. Zhao, and A. L. Yuille. "Non-rigid point set registration by preserving global and local structures," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 53–64, Jan. 2016.

[4] H. Leung and A. Young, "Small target detection in clutter using recursive nonlinear prediction," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 2, pp. 713–718, Apr. 2000.

[5] H. Qi, B. Mo, F. Liu, Y. He, and S. Liu, "Small infrared target detection utilizing local region similarity difference map," *Infrared Phys. Technol.*, vol. 71, pp. 131–139, Jul. 2015.

[6] Y. Zhao, H. Pan, C. Du, and Y. Zheng, "Principal curvature for infrared small target detection," *Infrared Phys. Technol.*, vol. 69, pp. 36–43, Mar. 2015.

[7] C. Yang, J. Ma, M. Zhang, S. Zheng, and X. Tian, "Multiscale facet model for infrared small target detection," *Infrared Phys. Technol.*, vol. 67, pp. 202–209, Nov. 2014.

[8] C. Yang *et al.*, "Directional support value of Gaussian transformation for infrared small target detection," *Appl. Opt.*, vol. 54, no. 9, pp. 2255–2265, Mar. 2015.

[9] 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.

[10] 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.

[11] 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.

[12] J. Ma, J. Zhao, Y. Ma, and J. Tian, "Non-rigid visible and infrared face registration via regularized Gaussian fields criterion," *Pattern Recognit.*, vol. 48, no. 3, pp. 772–784, Mar. 2015.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] K. Xie *et al.*, "Small target detection based on accumulated center-surround difference measure," *Infrared Phys. Technol.*, vol. 67, pp. 229–236, Nov. 2014.

[18] J. 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.

[19] M. Liu and Z. L. Xiang, "Contour detection based on Gabor filter and directional DoG filter," in *Proc. IEEE Mechatron. Mach. Vis. Pract.*, pp. 185–190, Dec. 2007.

[20] J. Geusebroek, A. W. M. Smeulders, and J. Weijer, "Fast anisotropic Gauss filtering," *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 938–943, Aug. 2003.