Infrared Small Target Detection via Nonnegativity-Constrained Variational Mode Decomposition

Xiaoyang Wang, Zhenming Peng, *Member, IEEE*, Ping Zhang, and Yanmin He

Abstract—Infrared small target detection is one of the key techniques in the infrared search and track system. Frequency differences among target, background, and noise are often important information for target detection. In this letter, a nonnegativity-constrained variational mode decomposition (NVMD) method is proposed. Unlike the traditional frequency-domain methods, the proposed method can adaptively decompose the input signal into several separated band-limited subsignals, with the nonnegativity constraint. First, a bandpass filter is used as a preprocessing step. Second, by exploring the frequency and nonnegativity properties of the small target, the NVMD model is constructed. The potential target subsignal can be obtained by solving the NVMD model. By performing threshold segmentation on the potential target subsignal, we can obtain the detection result of the infrared small target. Experiments on six real infrared image sequences demonstrate that the proposed method has a good performance in target enhancement and background suppression. Additionally, the proposed method shows strong robustness under various backgrounds.

Index Terms—Infrared image, nonnegativity constraint, small target detection, variational mode decomposition (VMD).

I. INTRODUCTION

INFRARED small target detection is a widely used technique in the infrared search and track (IRST) system. Due to the long imaging distance of the IRST system, the target usually has a small size and is easily corrupted by the background and noise. For small target detection in a single frame, a basic idea is to use the different features of target, background, and noise to suppress the background and extract the target. Existing methods for target detection in a single frame are often performed in either of two domains: the spatial domain and the frequency domain. Spatial-domain methods include the morphological filtering method [1], [2], the human visual system (HVS) method [3], and the statistical regression method [4]. The morphological filtering method is designed to suppress the background using morphologic operators. The HVS-based methods enhance the target by using the visual saliency feature. Statistical regression methods build a specific

The authors are with the University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: xywang1211@outlook.com; zmpeng@uestc.edu.cn; pingzh@uestc.edu.cn; heyanmin@uestc.edu.cn).

model to estimate the background distribution or use the input data itself to estimate the regression function. Frequency-domain methods consider the frequency difference between the target and background. An analysis shows that the frequency of small targets is generally higher than that of the background [5]. Using this prior information in the frequency domain, the target and background can be well separated by high-pass filters [6]. Spectral saliency methods are also proposed to enhance and extract the target from the spectral domain [7].

As discussed earlier, infrared small target detection in the frequency domain has always been an important research trend. Most of the frequency-domains methods utilize the idea of frequency filtering. In practice, traditional high-pass filters rely on preset parameters and prior information, which could cause less self-adaption. Additionally, the high-pass filter methods can be easily interfered with by the changing background and noise. Recently, a 2-D variational mode decomposition (2-D-VMD) method was proposed, which can *adaptively* decompose the input signal into several separate narrow spectral bands with a theoretical guarantee [8]. By constructing a valid mathematical model, the input signal can be decomposed concurrently by solving an optimization problem. The 2-D-VMD method has been applied to various applications, including signal and image denoising [9], [10], classification [11], and statistical prediction [12].

In this letter, an infrared small target detection method based on nonnegativity-constrained VMD (NVMD) is proposed. NVMD is a modified form of 2-D VMD. The key idea is to exploit the frequency difference between the target and background clutter and to detect the infrared target by solving the NVMD model. By decomposing the infrared image into several separate band-limited signals, the target can be extracted from the corresponding target subsignal. The nonnegativity constraint is introduced to better describe the target component, which can further suppress the background and enhance the target.

The rest of this letter is organized as follows. Section II presents the infrared small target detection method using NVMD. Section III shows the experiments and results. Numerical experiments are presented to demonstrate the performance of the proposed method. Section IV concludes this letter.

II. METHODOLOGY

A. Components of Infrared Images

Generally, an infrared image with a small target can be described by the following additive formula:

$$f(x, y) = f_T(x, y) + f_B(x, y) + f_N(x, y) \qquad (1)$$

where $f(x, y)$ is the grayscale data of the input infrared image, $f_T(x, y)$ and $f_B(x, y)$ are the grayscale data of the target and background, respectively, and $f_N(x, y)$ stands for the noise. In the frequency domain, the background always occupies most of the low-frequency bands, while the target always occupies a higher frequency band. The noise occupies the highest frequency band. The frequency difference feature has been widely studied [5]. Note that image parts with different frequency components may overlap with each other in space. The frequency differences among image components can be used to extract the small target from the infrared image.

B. Preprocessing Using a Bandpass Filter

To better take advantage of the frequency difference feature, a preprocessing step is necessary. Here, we introduce the bandpass filter. The bandpass filter is designed to eliminate the highest frequency component and the lowest frequency component and reserve the middle-frequency component. In the case of infrared image processing, it eliminates smooth background and the high-frequency component (including noise) while preserving the target component. Practically, we choose the difference of Gaussians (DoG) filter, with parameters $\sigma_1 = 0.5$ and $\sigma_2 = 2$ (σ_1 and σ_2 are the variance of Gaussians functions in the DoG filter). Other bandpass filters can also be utilized. Assuming the input infrared image $f(x, y) \in R^{m \times n}$, the result of the bandpass filter can be written as $f'(x, y) \in R^{m \times n}$.

C. Mathematical Model of NVMD

As discussed in Section II-A, the target, background, and noise are usually located in different frequency bands in infrared images. Since the target has a small size and no texture features, it can be viewed as a narrow-band component that occupies the middle-frequency band. The frequency characteristic can be utilized to extract the small target from the infrared image. In addition, the property $f_T(x, y) > 0$ holds in all infrared images. The small target can be detected by decomposing the image into nonnegative narrow-band components.

According to a newly proposed signal decomposition method [8], the input signal can be decomposed into several narrow-band subsignals by minimizing the sum of their bandwidths. An NVMD approach is proposed to extract the target from $f'(x, y)$. $f'(x, y)$ has finite bandwidth, which can be decomposed into K subsignals $u_k(x, y)$, $k = 1, \ldots, K$. The 2-D analytic signal of each subsignal $u_k(x, y)$ can be written as

$$u_{AS,k}(\boldsymbol{x}) = u_k(\boldsymbol{x}) * \left(\delta(\langle \boldsymbol{x}, \boldsymbol{e}_k \rangle) + \frac{j}{\pi \langle \boldsymbol{x}, \boldsymbol{e}_k \rangle} \right) \delta(\langle \boldsymbol{x}, \boldsymbol{e}_{k,\perp} \rangle)$$

$$(2)$$

where $\boldsymbol{x}$ is the simplified form of the spatial coordinate (x, y), $*$ is the convolution operator, δ is the 2-D impulse signal, and $\boldsymbol{e}_k$ is the unit vector, representing the direction of reference [13]. The analytic signal, which has only half-plane frequency, can remove the negative frequency component of $u_k(x, y)$, $k = 1, \ldots, K$.

After computing the corresponding analytic signal of $u_k(\boldsymbol{x})$, it is shifted to a pre-estimated baseband frequency
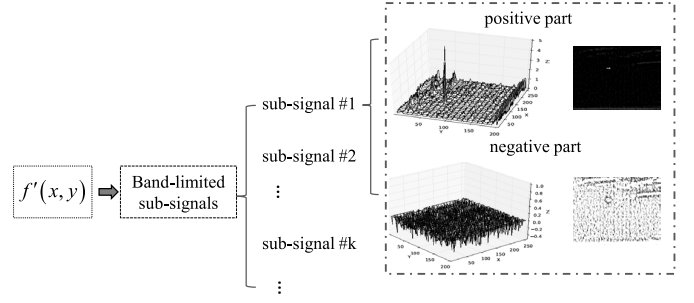

Fig. 1. Diagram shows the intermediate result of the band-limited subsignal decomposition for $f'(x, y)$ without additional constraints. For subsignal # 1, the 3-D projections of the positive part and negative part are given, as well as the grayscale images.

ω_k by mixing with a complex exponential expression, i.e., $u_{AS,k}(\boldsymbol{x}) e^{-j \langle \omega_k, \boldsymbol{x} \rangle}$. The bandwidth B_k of the frequency-shifted subsignal can be written as

$$B_k = \| \nabla (u_{AS,k}(\boldsymbol{x}) e^{-j \langle \omega_k, \boldsymbol{x} \rangle}) \|_2^2. \tag{3}$$

Minimizing the sum of B_k is a direct approach to estimating the band-limited subsignals $u_k(\boldsymbol{x})$, $k = 1, \ldots, K$.

Note that in the task of target detection, the target subsignal is our primary concern. The bandpass filter in the first step is designed to eliminate smooth background and high-frequency noise while preserving the target component. Thus, the target component in $f'(\boldsymbol{x})$ is nonnegative. Another motivation for the nonnegativity constraint comes from the subsignal decomposition procedure. Assume that $f'(\boldsymbol{x})$ is decomposed into several subsignals directly without any additional constraint; Fig. 1 shows the intermediate result. Take subsignal #1 as an example. The positive part of subsignal #1 contains the target signal and a small amount of clutter. The negative part, on the contrary, contains meaningless clutter and oscillations. This observation demonstrates the necessity of adding the nonnegativity constraint to the processed decomposition procedure. The nonnegativity constraint removes the clutter in subsignals, forcing the decomposition result to be clean and target only.

According to the above-mentioned analysis, the objective function of NVMD can be written as

$$\min_{u_k, \omega_k} \left\{ \sum_k \| \nabla [u_{AS,k}(\boldsymbol{x}) e^{-j \langle \omega_k, \boldsymbol{x} \rangle}] \|_2^2 \right\}$$

$$\text{s.t. } \forall \boldsymbol{x} : u_k(\boldsymbol{x}) \geq 0 \left\| f'(\boldsymbol{x}) - \sum_k u_k(\boldsymbol{x}) \right\|_F^2 \leq \varepsilon \tag{4}$$

where ε controls the reconstruction error. In practice, we choose $\varepsilon = 5e - 4$ or $\varepsilon = 1e - 3$.

D. Optimization Algorithm for NVMD

The objective function of NVMD can be effectively solved by the alternating direction method of multipliers (ADMM) [14]. The algorithm is detailed in Algorithm 1. Care needs to be taken when applying the nonnegativity constraint to $u_k(\boldsymbol{x})$. The ADMM is a loop calculation method. The nonnegativity constraint is introduced by setting the negative part of the current decomposed subsignal to zero in each iteration. In this way, the negative part in the current iteration
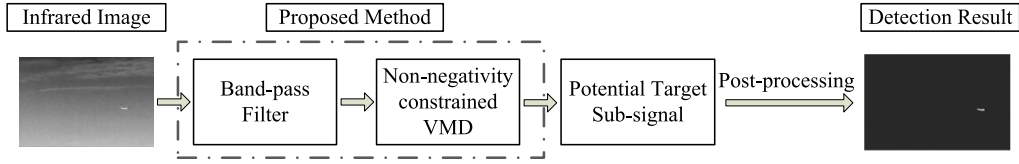
Fig. 2.   Flowchart of the proposed small target detection method.

Algorithm 1 ADMM Optimization for NVMD

Require:
   Preprocessed image $f'(x)$, number of sub-signals K;

Ensure:
   Decomposed sub-signals $u_k(x)$, $k = 1, \ldots, K$;

1: Initialize: $\{\hat{u}_k^0\}$, $\{\hat{\omega}_k^0\}$, $n = 0$;
2: **while** $n < N$ **do**
3:   **for** $k = 1 : K$ **do**
4:     $H_k^{n+1}(\omega) = 1 + sgn(\langle \omega_k^n, \omega \rangle)$
5:     Update $\hat{u}_{AS,k}$:

$$\hat{u}_{AS,k}^{n+1}(\omega) = H_k^{n+1}(\omega)$$

$$\times \left[\frac{\hat{f}(\omega) - \sum_{i<k} \hat{u}_i^{n+1}(\omega) - \sum_{i>k} \hat{u}_i^n(\omega) + \frac{\hat{\lambda}^n(\omega)}{2}}{1 + 2\alpha|\omega - \omega_k^n|^2} \right]$$

6:     Update ω_k:

$$\omega_k^{n+1} = \frac{\int_{R^2} \omega |\hat{u}_{AS,k}^{n+1}(\omega)|^2 \, d\omega}{\int_{R^2} |\hat{u}_{AS,k}^{n+1}(\omega)|^2 \, d\omega}$$

7:     Solve u'_k in the spatial domain:

$$u_k'^{n+1}(x) = \mathcal{R}\{\mathcal{F}^{-1}\hat{u}_{AS,k}^{n+1}(\omega)\}$$

8:     Apply the non-negativity-constraint:

$$\forall x, u_k^{n+1}(x) = \begin{cases} u_k'^{n+1}(x) & \text{if } u_k'^{n+1}(x) \geq 0 \\ 0 & \text{if } u_k'^{n+1}(x) < 0 \end{cases}$$

9:   **end for**
10:

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \tau\left(\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)\right)$$

11:   $n = n + 1$
12: **end while**
13: $u_k(x) = u_k^n(x), k = 1, \ldots, K$

---

is removed, and the positive part continues to be updated in the next iteration, which can further narrow the bandwidth of positive subsignals. The maximum number of iterations N is a preset empirical value. In our experience, $N < 10$ (usually $N = 8$) can yield good results.

E. Infrared Small Target Detection Using NVMD

Fig. 2 shows the proposed method of small target detection using NVMD. The detailed steps are listed as follows.
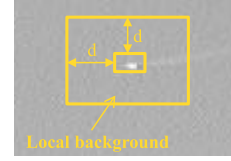

Fig. 3.   Infrared target and local background area.

1) Preprocess the input image to $f'(x, y) \in R^{m \times n}$.
2) Decompose $f'(x, y)$ into K subsignals $u_1(x, y), \ldots,$ $u_K(x, y)$ using NVMD. Extract the potential target $u_{\text{target}}(x, y)$ from the K subsignals. Normally, $u_{\text{target}}(x, y) = u_1(x, y)$. Subsignals with k other than 1 always show the characteristic of oscillation. The direction of oscillation changes in different test sequences.
3) Use threshold segmentation as a postprocessing method. The final detection result is derived from $u_{\text{target}}(x, y)$ by segmentation. The threshold can be decided adaptively.

III. EXPERIMENTS AND ANALYSIS

In this section, the proposed NVMD method is validated through experiments on several infrared images. First, we introduce the evaluation metrics. Then, we show the representative test images and corresponding detection results. Numerical analyses are also presented in this section.

A. Evaluation Metrics

Here, we introduce the evaluation metrics. Fig. 3 shows the target region and the local background area. The signal-to-clutter ratio (SCR), the SCR gain, and the background suppression factor (BSF) are defined as

$$\text{SCR} = \frac{|\mu_t - \mu_b|}{\sigma_b}, G_{\text{SCR}} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}}}, \text{BSF} = \frac{C_{\text{in}}}{C_{\text{out}}} \quad (5)$$

where μ_t and μ_b are the average gray levels of the target region and the local background area, respectively. σ_b is the local background standard deviation. d is a parameter determining the size of the neighborhood area. In our experiment, $d = 20$. SCR_{in} and SCR_{out} are the SCRs of the original image and processed image. C_{in} and C_{out} are the standard deviations of the backgrounds in the original image and processed image, respectively.

The probability of detection P_d and the false-alarm rate F_a are also introduced

$$P_d = \frac{\text{number of true detections}}{\text{number of actual targets}} \quad (6)$$

$$F_a = \frac{\text{number of pixels of all false detections}}{\text{number of pixels of all test images}}. \quad (7)$$
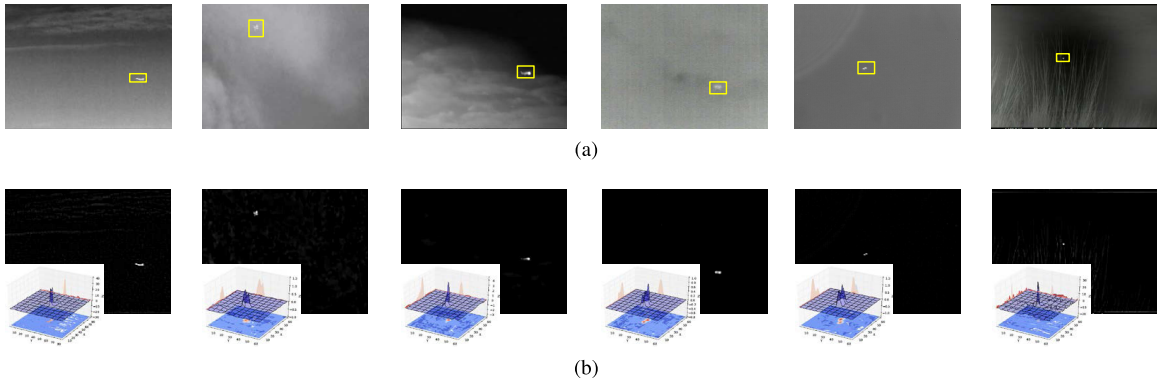
(a)

(b)

Fig. 4. Original test images and detection results.

TABLE I

$\overline{G_{SCR}}$ AND $\overline{BSF}$ COMPARISONS OF TEST IMAGES

Detective Methods	Evaluation indicators	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6
Tophat	$\overline{G_{SCR}}$	4.0907	2.0979	2.9106	1.5703	1.0074	1.3594
	$\overline{BSF}$	9.7581	12.0404	13.1325	2.6486	1.8060	2.9762
Adaptive BHPF	$\overline{G_{SCR}}$	3.3776	2.0440	0.9925	2.0937	1.0506	1.7690
	$\overline{BSF}$	6.7916	10.7980	1.0044	3.2007	2.1207	3.8166
LS-SVM	$\overline{G_{SCR}}$	1.9978	1.1265	5.2317	1.0403	1.2013	2.7101
	$\overline{BSF}$	134.6170	74.4708	9.0771	31.0763	9.1319	3.1882
KR-CFAR	$\overline{G_{SCR}}$	0.9974	0.8334	1.0536	0.2930	0.2088	0.6292
	$\overline{BSF}$	9.2509	10.8794	6.6287	3.0979	1.9865	3.4700
PQFT	$\overline{G_{SCR}}$	4.5409	2.0524	3.0275	1.7337	1.3048	1.0775
	$\overline{BSF}$	51.5555	25.2831	32.8696	23.4065	3.7970	14.5046
NVMD	$\overline{G_{SCR}}$	4.6144	3.6152	5.9209	4.9755	1.3134	2.5353
	$\overline{BSF}$	56.6139	737.8575	311.6139	47.0275	40.8632	24.4592

B. Experimental Results

Six real infrared image sequences are chosen to test the effectiveness of the proposed method. Fig. 4(a) shows the representative images of test sequences 1–6, in which the targets are marked by yellow rectangles. Fig. 4(b) shows the processed results of the proposed method, as well as the 3-D surface in the target area. It can be concluded from Fig. 4 that the proposed method can achieve good detection performance with dim targets (see test images 2 and 4) and extremely small targets (see test image 6). Additionally, the background clutter in six test image sequences is well suppressed by the proposed method.

To further verify the validity of the proposed method, we adopt five other methods for comparison. These methods are the conventional Tophat method [15], the adaptive Butterworth high-pass filter method [6], the least-square support vector machine (LS-SVM)-based directional high-pass filter method [16], the kernel-based nonparametric regression method [4], and the phase spectrum of quaternion Fourier transform method [7]. Experimental results of these methods with the indicators given in Section III-A are shown in Table I. The highest value of each evaluation indicator in each column is marked in red, while the second highest value is marked in blue.

TABLE II

AUC VALUES OF TEST IMAGES ($\times 10^{-3}$)

	Tophat	Adaptive BHPF	LS-SVM	KR-CFAR	PQFT	NVMD
Image 1	999.94	999.30	730.71	999.92	999.91	999.99
Image 2	999.70	970.02	631.64	860.88	999.83	999.87
Image 3	1000.00	1000.00	1000.00	1000.00	1000.00	1000.00
Image 4	999.87	999.84	387.95	691.81	997.17	999.99
Image 5	998.38	983.85	619.44	876.58	999.85	999.98
Image 6	999.47	999.74	999.73	999.67	999.56	999.93

From Table I, we can see that in test image 1, the proposed method achieves the highest G_{SCR} value, while the LS-SVM method achieves the highest BSF value. In test images 2–5, the proposed method excels in both target enhancement and background suppression. In test image 6, the proposed method achieves the second highest G_{SCR}, while G_{SCR} of the LS-SVM method is slightly higher. This is because the LS-SVM method includes a strategy of correcting the filtering result by setting some of the values to zero. Overall, the proposed method can obtain better results than the other methods.
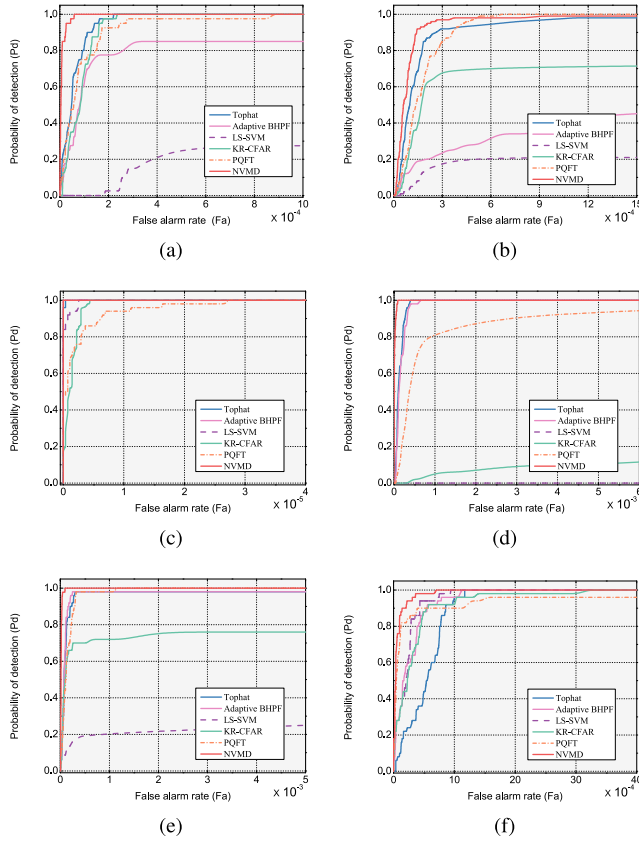
Fig. 5. ROC curves of six detection methods. (a)–(f) Results of test images 1–6. The closer a curve is to the top-left corner, the better the corresponding method.
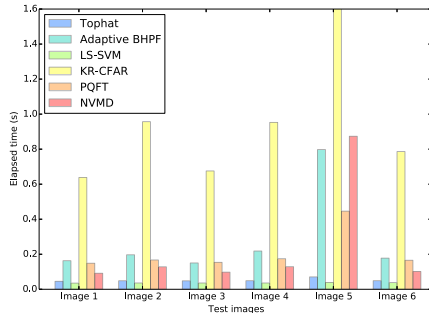

Fig. 6. Time-consumption comparison of six detection methods on six test image sequences.

To demonstrate the robustness of the proposed method, the receiver operating characteristic (ROC) curves are given in Fig. 5. The proposed method always has higher P_d than the other methods under identical F_a. Additionally, we compute the area under the curve (AUC) values of each method, which are shown in Table II. The proposed method has the largest AUC values for the six test images, which also proves its effectiveness.

The time-consumption comparison of six methods is given in Fig. 6. Bars with the same color represent the average time required to process an image with a specific method. All experiments are implemented by MATLAB software on a PC with 8-GB memory and an Intel i5 processor. We can see

that the proposed method is a little slower than the Tophat and LS-SVM methods, and faster than other methods.

IV. CONCLUSION

In this letter, an NVMD method is proposed to detect a small target from infrared images. This method is performed in the frequency domain, from which the target subsignal is extracted adaptively. The nonnegativity constraint is introduced to remove the background clutter and further enhance the target during the decomposition process. Various evaluation indicators, including SCR gain and BSF, demonstrate the good performance of the NVMD method in target enhancement and background suppression. The robustness of NVMD is validated by ROC curves of six real infrared image sequences. Additionally, the efficiency of the proposed method is acceptable. In the future, we will focus on the detection of an infrared target with bigger size by means of the NVMD method.

REFERENCES

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

[2] W. Wang, N. Yang, Y. Zhang, F. Wang, T. Cao, and P. Eklund, "A review of road extraction from remote sensing images," *J. Traffic Transp. Eng.*, vol. 3, no. 3, pp. 271–282, 2016.

[3] J. Han, Y. Ma, B. Zhou, F. Fan, K. Liang, and Y. Fang, "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.

[4] Y. Gu, C. Wang, B. Liu, and Y. Zhang, "A kernel-based nonparametric regression method for clutter removal in infrared small-target detection applications," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 3, pp. 469–473, Jul. 2010.

[5] S. Qi, D. Ming, J. Ma, X. Sun, and J. Tian, "Robust method for infrared small-target detection based on Boolean map visual theory," *Appl. Opt.*, vol. 53, no. 18, pp. 3929–3940, Jun. 2014.

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

[7] S. Qi, J. Ma, H. Li, S. Zhang, and J. Tian, "Infrared small target enhancement via phase spectrum of quaternion Fourier transform," *Infr. Phys. Technol.*, vol. 62, pp. 50–58, Jan. 2013.

[8] K. Dragomiretskiy and D. Zosso, "Two-dimensional variational mode decomposition," in *Energy Minimization Methods in Computer Vision and Pattern Recognition* (Lecture Notes in Computer Science), vol. 8932, E. Bae, T. F. Chan, and M. Lysaker, Eds. Hong Kong: Springer, 2015, pp. 197–208.

[9] S. Lahmiri and M. Boukadoum, "Physiological signal denoising with variational mode decomposition and weighted reconstruction after DWT thresholding," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2015, pp. 806–809.

[10] S. Lahmiri, "Denoising techniques in adaptive multi-resolution domains with applications to biomedical images," *Healthcare Technol. Lett.*, vol. 4, no. 1, pp. 25–29, 2017.

[11] S. Lahmiri, "High-frequency-based features for low and high retina haemorrhage classification," *Healthcare Technol. Lett.*, vol. 4, no. 1, pp. 20–24, 2017.

[12] S. Lahmiri, "Comparing variational and empirical mode decomposition in forecasting day-ahead energy prices," *IEEE Syst. J.*, to be published.

[13] T. Bülow and G. Sommer, "A novel approach to the 2D analytic signal," in *Proc. Int. Conf. Comput. Anal. Images Patterns*, Sep. 1999, pp. 25–32.

[14] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[15] 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.

[16] 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.