Robust Infrared Small Target Detection Using Local Steering Kernel Reconstruction

Article *in* Pattern Recognition · December 2017

DOI: 10.1016/j.patcog.2017.12.012

2 authors:

Yansheng Li
Wuhan University
32 PUBLICATIONS **262** CITATIONS
SEE PROFILE

Yongjun Zhang
Wuhan University
112 PUBLICATIONS **665** CITATIONS
SEE PROFILE

Pattern Recognition 77 (2018) 113–125

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

Robust infrared small target detection using local steering kernel reconstruction

Yansheng Li, Yongjun Zhang*

School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

ARTICLE INFO

Article history:
Received 17 October 2016
Revised 3 December 2017
Accepted 17 December 2017
Available online 18 December 2017

Keywords:
Infrared small target detection
Local steering kernel (LSK)
Closed-form feature reconstruction

ABSTRACT

Because infrared small target detection plays a crucial role in infrared monitoring and early warning systems, it has been the subject of considerable research. Although many infrared small target detection approaches have been proposed, how to robustly detect small targets in poor quality infrared images remains a challenge. Since existing feature descriptors are often sensitive to the quality of infrared images, this paper advocates the use of a local steering kernel (LSK) to encode the infrared image patch because the LSK method can provide robust estimation of local intrinsic structure, even for poor quality images. Furthermore, this paper proposes a novel local adaptive contrast measure based on LSK reconstruction (LACM-LSK) for infrared small target detection. To demonstrate the effectiveness of the proposed approach, a diverse test dataset, including six infrared image sequences with different backgrounds, was collected. Extensive experiments on the test dataset confirm that the proposed infrared small target detection approach can achieve better detection performance than state-of-the-art approaches.

1. Introduction

Automatic infrared small target detection is one of the most significant aspects of an automatic target recognition (ATR) system [1,2]. For early warning applications, the incoming targets must be detected at a great distance. As the infrared sensor is distant from an object of interest, such objects occupy only a small number of pixels in the infrared image. Furthermore, the infrared radiant energy of the object of interest decays greatly over long distances, which causes the signal-to-noise ratio (SNR) of the target of interest generally to be very low [3]. Additionally, the targets are often buried in complex background clutter. In addition, the brightness of infrared images is often affected by the variability of thermal radiation sources in the field of view, and infrared images tend to be polluted by infrared sensor noise. As a consequence, infrared images are often of poor quality. Due to the poor quality images, traditional image descriptors are unable to provide infrared small target detection. Hence, robust small target detection from poor quality infrared images presents a number of challenges.

In the past several decades, numerous sequential image-based approaches for the infrared detection of small moving targets [4,5] have been proposed. However, temporal information extraction from infrared sequential images is still extremely challenging because common feature descriptors [6–8] and matching methods [9–10] are not often used directly when compensating for motion in infrared sequential images. To avoid this open problem, many researchers and engineers turn to infrared small target detection based on a single image. Generally, an infrared small target presents as a Gaussian light blob in the image, which is the crucial distinction between the infrared small target and background clutter. Methods based on this distinction include max-mean and max-median filtering [11], the mathematical morphological-based methods [12,13], line-based reconstruction [14], and the image entropy-based method [15]. Based on the theory of extrema, facet-based methods [16–18] and LS-SVM-based filtering [19] were used to identify objects of interest as well as suppress background clutter. From the visual attention perspective, several kinds of saliency-based infrared small target detection approaches were proposed [20–22]. Unfortunately, most saliency-based infrared small target detection approaches have high computational complexity and are difficult to optimize using parallelism. Subsequently, Chen et al. [23] proposed a local contrast measure to identify infrared small targets. These researchers subsequently proposed a refinement of the local contrast measure to improve dark target detection performance [24]. Deng et al. [25] proposed a weighted local contrast measure to suppress cloudy-sky backgrounds. To improve on the local contrast measure, Wei et al. [26] proposed a patch-based contrast measure. The patch-based processing style has also been adopted for infrared small target detection in [27] and our previous work [4]. In these patch-based infrared small target detection

* Corresponding author.
E-mail addresses: yansheng.li@whu.edu.cn (Y. Li), zhangyj@whu.edu.cn (Y. Zhang).

https://doi.org/10.1016/j.patcog.2017.12.012
0031-3203/

approaches, the feature descriptor of an infrared image patch was encoded by simply vectoring the raw pixel intensities. The feature descriptor encoded by raw pixel intensities is directly affected by the image quality, which negatively affects patch-based infrared small target detection approaches [4,26–28]. Hence, realizing robust patch-based infrared detection of small targets requires research into the robust representation of infrared image patches.

To address this aim, this paper investigates the use of a local steering kernel (LSK) [29] to encode the infrared image patch because the LSK has been specifically designed to address pixel-level image noise and uncertainty through estimation of the local intrinsic structure. To date, the LSK method has been successfully used in such applications as noise reduction [29], super-resolution [30], saliency modeling [31], and object detection [32]. In this paper, we explore the utility of LSK for the infrared small target detection task. Extensive experiments have shown that feature descriptors using LSKs are robust to brightness variation and noise interference. In addition, a patch containing a uniform region (flat region), another containing textural clutter and a third containing a structural region have dramatically different LSK descriptors. Unfortunately, the LSK descriptor of an image patch containing a small target highly resembles a texture clutter patch. Hence, use of the LSK alone cannot accomplish the infrared small target detection task. Considering that the small target is isolated but that texture clutter is repeated along at least one direction, this paper proposes a local adaptive contrast measure based on regularized LSK reconstruction (LACM-LSK) that can simultaneously identify the small target of interest and suppress the background clutter, including the texture clutter. Extraction of LSK descriptors and calculation of the contrast map using the LACM-LSK can be easily optimized by a parallel algorithm to meet the requirements of real-time applications. From the contrast map, the final determination of the small targets' locations can be achieved using a simple threshold segmentation method.

A test dataset composed of six infrared sequences containing several typical backgrounds with sea, ground, and cloud clutter is used to demonstrate the validity of the proposed approach. Compared with state-of-the-art approaches, the proposed approach shows very impressive results. As a whole, the contributions of this paper can be summarized as follows:

- Because the quality of infrared images is often poor (e.g., infrared images are often polluted by infrared sensor noise), we propose encoding the infrared image patch using the LSK method. More specifically, LSKs can robustly estimate the intrinsic image structure and address pixel-level image noise and ambiguity.
- This paper proposes a local adaptive contrast measure based on regularized LSK reconstruction that has an efficient closed-form solution. This proposed adaptive contrast measure is used to identify infrared small targets. Many experiments have confirmed that the proposed infrared small target detection approach can outperform existing approaches.

The remainder of this paper is organized as follows. Section 2 details the proposed infrared small target detection approach, including descriptions of the LSK and LACM-LSK. Section 3 presents the experimental results, which include a sensitivity analysis of the crucial parameters of the proposed infrared small target detection approach and a comparison of the proposed approach to state-of-the-art approaches. The study's conclusions are presented in Section 4.

2. Methodology

In this section, we first introduce the calculation of the local steering kernel (LSK) and its infrared image results. Next, we present the local adaptive contrast measure based on regularized LSK reconstruction (LACM-LSK). Finally, the proposed infrared small target detection approach is summarized.

2.1. Robust representation of an infrared image patch using a local steering kernel (LSK)

In this paper, the position of one pixel is represented by a coordinate vector $\mathbf{x}_i = [x_i^1, x_i^2]^T$, and the pixel intensity at $\mathbf{x}_i$ is denoted by $\mathbf{I}(\mathbf{x}_i)$. As shown in [29], the descriptive power of the LSK mainly derives from the steering matrix, which is also known as the gradient covariance matrix. Given an image patch $\mathrm{W}(\mathbf{x}_i) = \{\mathbf{x}_1, \cdots, \mathbf{x}_i, \cdots, \mathbf{x}_P\}$ centered at $\mathbf{x}_i$ where P denotes the number of pixels in the given patch, the corresponding LSK representation can be modeled as

$$K(\mathbf{x}_i, \mathbf{x}_j) = \frac{\sqrt{\det(\mathbf{C}_j)}}{h^2} \exp\left\{\frac{(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{C}_j (\mathbf{x}_i - \mathbf{x}_j)}{-2h^2}\right\} \tag{1}$$

where $\mathbf{x}_j \in \mathrm{W}(\mathbf{x}_i)$, and h is the global smoothing parameter and is set to 0.2 for all experiments, according to Seo and Milanfar [31]. In addition, the covariance matrix $\mathbf{C}_j$ is estimated from a collection of spatial gradient vectors within the neighborhood centered at position $\mathbf{x}_j = [x_j^1, x_j^2]^T$. The covariance matrix estimation is of great importance to LSK generation and is specified as follows.

Let $\Omega(\mathbf{x}_j) = \{\mathbf{x}_1, \cdots, \mathbf{x}_j, \cdots \mathbf{x}_M\}$ denote the coordinate set of neighboring pixels centered at pixel coordinate $\mathbf{x}_j = [x_j^1, x_j^2]^T$ and M denote the number of neighboring pixels. Natively, the covariance matrix $\mathbf{C}_j$ can be directly estimated by $\mathbf{G}_j^T \mathbf{G}_j$ where $\mathbf{G}_j$ can be expressed by

$$\mathbf{G}_j = \begin{bmatrix} \mathbf{I}_1(\mathbf{x}_1) & \mathbf{I}_2(\mathbf{x}_1) \\ \vdots & \vdots \\ \mathbf{I}_2(\mathbf{x}_j) & \mathbf{I}_2(\mathbf{x}_j) \\ \vdots & \vdots \\ \mathbf{I}_1(\mathbf{x}_M) & \mathbf{I}_2(\mathbf{x}_M) \end{bmatrix} \tag{2}$$

where $\mathbf{I}_1(\cdot)$ and $\mathbf{I}_2(\cdot)$ denote the first derivatives along the horizontal and vertical axes.

To improve robustness and stability, based on the successful experience of Takeda et al. [29], our approach estimates the covariance matrix using a regularized parametric method. Based on the singular value decomposition (SVD) of $\mathbf{G}_j$, the stable covariance matrix $\mathbf{C}_j$ can be expressed by

$$\mathbf{C}_j = \left(\frac{s_1 s_2 + \lambda''}{N}\right)^{\alpha} \left(\left(\frac{s_1 + \lambda'}{s_2 + \lambda'}\right)^2 \mathbf{u}_1 \mathbf{u}_1^T + \left(\frac{s_2 + \lambda'}{s_1 + \lambda'}\right)^2 \mathbf{u}_2 \mathbf{u}_2^T\right) \tag{3}$$

where λ' and λ'' are parameters that are set to 1 and 10^{-7}, respectively, for all experiments, based on Seo and Milanfar [31]. The sensitivity analysis of the amplification factor α is detailed in Section 3. In addition, the singular values (s_1, s_2) and the singular vectors $(\mathbf{u}_1, \mathbf{u}_2)$ come from the SVD formula $\mathbf{G}_j = \mathbf{U}_j \mathbf{S}_j \mathbf{V}_j^T = \mathbf{U}_j \mathrm{diag}[s_1, s_2]_j [\mathbf{u}_1, \mathbf{u}_2]_j^T$.

In our implementation, before calculating the LSK representation, we first normalize the gradients in the global domain because the contrast in the infrared image is relatively smaller than in natural images. Fig. 1 illustrates the results of the LSK representation. As depicted in Fig. 1, LSK shows discrimination stability under conditions of brightness variation and noise interference. From intuitive comparison of the shapes, the LSK representations of the patch containing the flat region (i.e., the 2nd patch in Fig. 1), the patch containing textural clutter (i.e., the 3rd patch in Fig. 1), and the patch containing the structural region (i.e., the 4th patch in Fig. 1) are dramatically different from one another. Unfortunately, the LSK representation of the patch containing the small target of

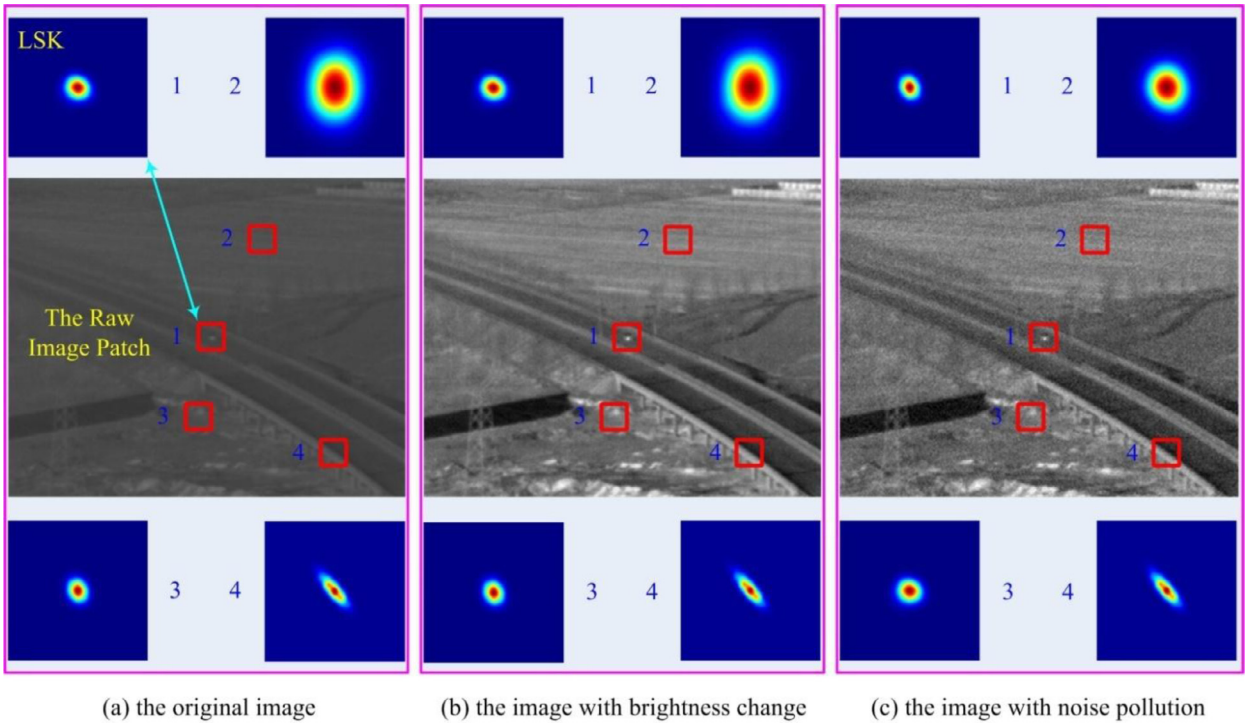


Fig. 1. Robust and invariant representation of infrared image patch using LSK. (a) shows the original image; (b) shows the image with brightness change; (c) shows the image polluted by Gaussian noise with a standard deviation = 10. In addition, the 1st patch contains the target of interest, the 2nd patch contains a flat region, the 3rd patch contains textural clutter, and the 4th patch contains an edge structure.

interest (i.e., the 1st patch in Fig. 1) is highly correlated with the patch containing textural clutter (i.e., the 3rd patch in Fig. 1). Accordingly, an infrared small target cannot be directly recognized by the existing LSK template matching method [32] or the metric learning methods [33,34] because a unique target feature signature is a prerequisite of these methods. How to make full use of the robust LSK representations to uniquely enhance small targets of interest deserves more exploration.

2.2. Local adaptive contrast measure based on regularized LSK reconstruction (LACM-LSK)

As mentioned in Section 2.1, LSK representations can robustly represent an infrared image even under conditions of brightness variation and noise interference. However, a single LSK representation cannot recognize an infrared small target because the LSK representation of the infrared small target region is indistinguishable from the LSK representation of the textural clutter region. In practice, human beings can readily distinguish infrared small targets from complex backgrounds, which results primarily from the selective attention mechanism of the human visual system (HVS). In the literature, various types of local contrast measures have been proposed to imitate the selective attention mechanism and have shown great potential in infrared small target detection. Specifically, the contrast measures a difference between the current unit and its neighboring units, which makes full use of spatial context and is capable of eliminating the confusion resulting from the examination of a single unit in infrared small target detection. Accordingly, defining the contrast is one of the most important tasks in infrared small target detection [26]. Hence, this paper proposes a local contrast measure based on LSK representations that simultaneously enhances infrared small targets and suppresses background clutters.

First, we provide examples of the infrared small target and background clutter images to show the appearance of the area of interest and the appearances of its neighbors to intuitively demonstrate why a contrast measure based on LSK representations would detect an infrared small target. As is widely recognized in the literature, an infrared small target has a signature discontinuous with its neighborhood [23,26]. As depicted in Fig. 2(a), the center patch (i.e., the 5th patch) corresponds to one infrared small target, and its neighboring patches greatly differ from it. In addition, textural clutter is often not isolated, which means that the neighboring patches also probably contain textural clutter if the center patch does. For example, as illustrated in Fig. 2(b), the 5th patch is the center patch and contains ground clutter, and some of its neighboring patches, i.e., the 1st, 2nd, 4th, 6th, and 9th patches, also contain ground clutter. Accordingly, an ideal local contrast measure should output a large contrast value when the center patch contains the infrared small target and a small contrast value when the center patch contains background clutter. As depicted in Fig. 2(a), the center patch (i.e., the 5th patch), containing the small target, differs from its neighbors in both the raw image representation domain and the LSK representation domain, thus the contrast between the center patch and its neighbors should be large in both domains. In contrast, as depicted in Fig. 2(b), the center patch (i.e., the 5th patch), containing ground clutter, differs from all of its neighbors in the raw image representation domain, but resembles some of its neighbors, such as the 1st, 2nd, 4th, 6th, and 9th patches, in the LSK representation domain. As a consequence, when the center patch contains only background clutter, the contrast value between the center patch and its neighbors would be large in the raw image representation domain, but the contrast value between the center patch and its neighbors may be small in the LSK representation domain using a qualified contrast measure. Therefore, this paper attempts to use an appropriate contrast measure based on the LSK representations to highlight the small target of interest and suppress various types of background clutter, including textural clutter, simultaneously.

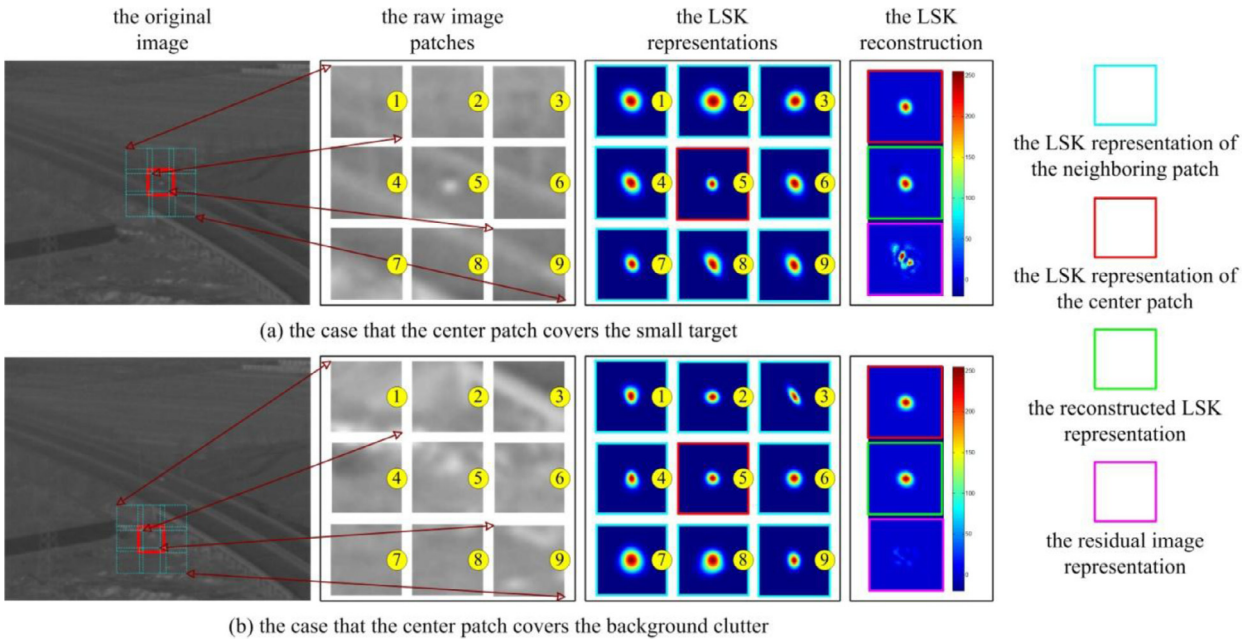


Fig. 2. Illustration of raw image patches and LSK representations in a local neighborhood, and the intuitive reconstruction residual at different regions. The blue rectangles denote the LSK representations of the neighboring patches. The red rectangle delineates the LSK representation of the center patch. The green rectangle denotes the reconstructed LSK representation when we use the neighboring LSK representations to approximate the center LSK representation. In addition, the magenta rectangle shows the difference between the reconstructed LSK representation and the original LSK representation of the central patch. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

This paper proposes a local adaptive contrast measure based on regularized LSK reconstruction (LACM-LSK) to identify infrared small targets. The intuitive sense behind this proposal is illustrated by Fig. 2. As depicted in Fig. 2(a), the LSK representation of the center patch (i.e., the 5th patch) containing the small target is significantly different from its neighbors' and cannot be linearly reconstructed from the neighbors' LSK representations, which is verified by the large reconstruction residual shown in Fig. 2(a). However, as depicted in Fig. 2(b), the linear combination of the neighboring patches' LSK representations can approximately reconstruct the LSK representation of the central patch containing only textural clutter, which is verified by the small reconstruction residual shown in Fig. 2(b). Hence, the reconstruction residual of the center patch's LSK representation can be used to define an adaptive contrast measure that can robustly indicate whether the given center patch contains a small target.

In the following, the proposed LACM-LSK is described. Let $\mathbf{f}^c \in R^{P\times 1}$ denote the vectorization of the center's LSK representation, where $P = s \times s$ and s is the patch size. In addition, $\mathbf{F} = [\mathbf{f}^1, \cdots, \mathbf{f}^N]$ denotes the union of the neighboring feature vectors, each of which is the vectorization of the corresponding neighboring LSK representation, where N is the number of neighboring patches. Let $\mathbf{w} \in R^{N\times 1}$ be the linear combination coefficients. Based on the neighboring LSK representations, the LSK representation of the center patch can be reconstructed by optimizing the following objective function:

$$\min_{\mathbf{w}} \| \mathbf{f}^c - \mathbf{F}\cdot\mathbf{w} \| + \lambda \| \mathbf{w}\|^2 \tag{4}$$

where λ is the regularization term that controls the trade-off between the estimation bias and the variance of the fitting model. Benefiting from the discriminative LSK representation, the reconstruction process is less sensitive to the regularization term λ than when the raw image feature of Li et al. [4] is used. Hence, the regularization term λ is empirically set to 1.0.

Optimizing the objective function in Eq. (4), the coefficient vector $\mathbf{w}$ has a closed-form solution expressed as $\mathbf{w} = (\mathbf{F}^T\cdot\mathbf{F} + \lambda\cdot\mathbf{I})^{-1}\cdot\mathbf{F}^T\cdot\mathbf{f}^c$. Furthermore, based on the neighboring feature vectors $\mathbf{F} = [\mathbf{f}^1, \cdots, \mathbf{f}^N]$ and the coefficient vector $\mathbf{w}$, the reconstruction feature vector $\mathbf{f}^r$ of the central feature vector $\mathbf{f}^c$ can be efficiently expressed by

$$\mathbf{f}^r = \mathbf{F}\cdot\mathbf{w} = \mathbf{F}\cdot(\mathbf{F}^T\cdot\mathbf{F} + \lambda\cdot\mathbf{I})^{-1}\cdot\mathbf{F}^T\cdot\mathbf{f}^c \tag{5}$$

The reconstruction residual of the center's LSK representation is further used to measure the local contrast between the center patch and its neighboring patches. More specifically, the proposed local adaptive contrast measure based on LSK reconstruction (LACM-LSK) can be expressed by

$$\begin{aligned}\mathbf{ALCM-LSK}(\mathbf{f}^c, \mathbf{F}) &= \| \mathbf{f}^c - \mathbf{f}^r \|^2 \\ &= \| \mathbf{f}^c - \mathbf{F}\cdot(\mathbf{F}^T\cdot\mathbf{F} + \lambda\cdot\mathbf{I})^{-1}\cdot\mathbf{F}^T\cdot\mathbf{f}^c \|^2\end{aligned} \tag{6}$$

In the following section, we demonstrate the use of the proposed LACM-LSK to enhance the representation of potential infrared small targets.

2.3. Robust infrared small target detection using LACM-LSK

Fig. 3 presents the flowchart of the proposed infrared small target detection approach. As depicted in Fig. 3, the overall infrared small target detection approach can be divided into two main steps: infrared small target enhancement and infrared small target segmentation. Infrared small target enhancement works to highlight the infrared small target of interest and suppress background clutter, and infrared small target segmentation explicitly identifies the targets of interest.

As an example, in the infrared small target enhancement step, the 1st image in Fig. 3 is the input, the patch size is set to s, and the size of the overlap between adjoining patches is set to t. The LSK representation centered at each pixel is first encoded according to Eq. (1). Based on the proposed LACM-LSK defined in Eq. (6), the confidence value $\mathbf{E}(i, j)$ can be estimated using the LSK representation centered at (i, j) and its neighboring LSK representations, where $\mathbf{E}(i, j)$ stands for the likelihood value of the pixel belonging

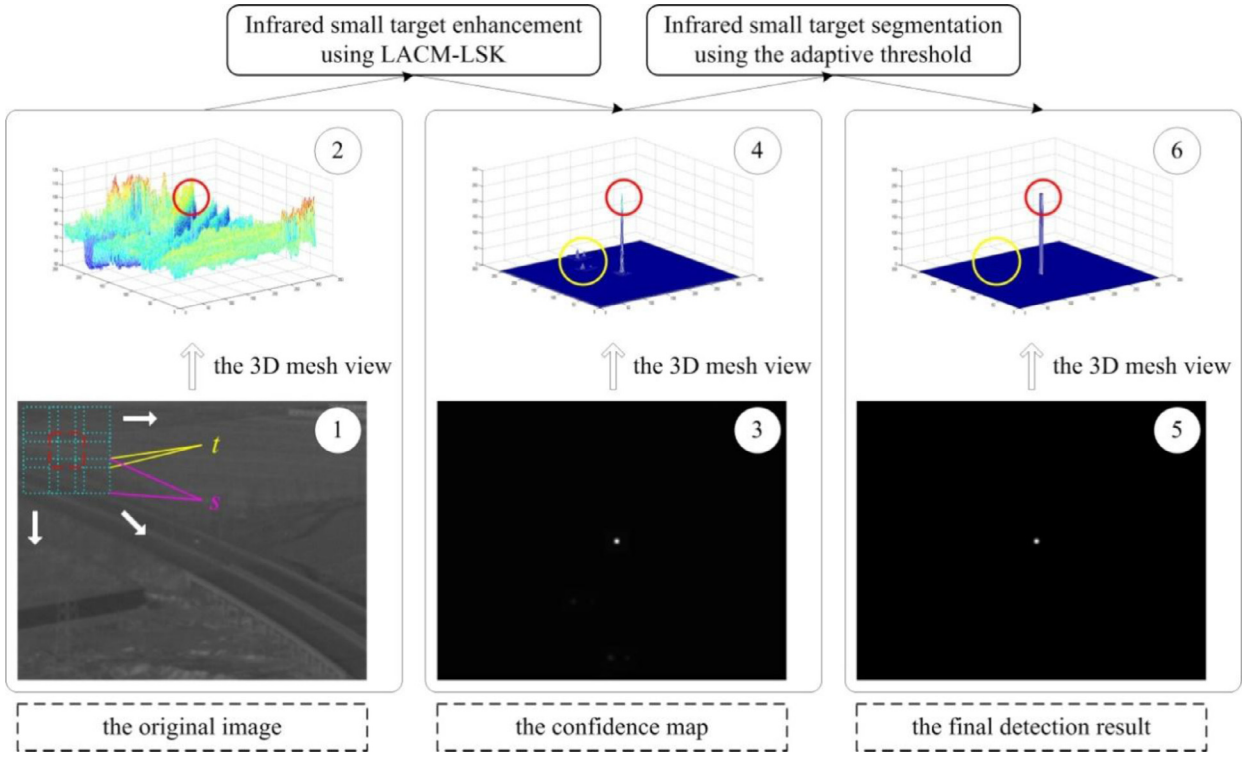


Fig. 3. The flowchart of the proposed infrared small target detection approach. *s* and *t* denote the size of the patch and the overlap size between adjoining patches, respectively. The red circle contains the infrared small target, and the yellow circle shows the background clutter. The 1st, 3rd, and 5th images show the original image, the confidence map, and the final detection result, respectively. The 2nd, 4th, and 6th images present the 3D mesh views of the 1st, 3rd, and 5th images, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Algorithm 1 Proposed infrared small target detection approach.

Input: The Original Infrared Image
Output: The Target Detection Result
1: Compute the local steering kernel (LSK) representations centered at each pixel according to Eq. (1).
2: Compute the infrared small target enhancement map **E** using LACM-LSK.
E(*i, j*) is calculated using the LSK representation centered at the pixel coordinate (*i, j*) and its neighboring LSK representations. The specific calculation process is based on Eq. (6).
3: Segment the infrared small target from the confidence map **E** based on the adaptive threshold given in Eq. (7).

to a small target and (*i, j*) denotes the pixel coordinates. In this process, the sliding patch moves on the infrared image top down and left to right, pixel by pixel. At each position, **E**(*i, j*) is calculated based on the LSK representation of the current center patch and the LSK representations of its neighbors. The confidence map **E** is the output of this step, and it can effectively identify the small target, as depicted in the 3rd image in Fig. 3. As depicted in the 4th image of Fig. 3, which is the 3D mesh view of the confidence map shown in the 3rd image of Fig. 3, the confidence map still contains a number of false positives, which are marked by the yellow circle. Fortunately, the false positives are easily distinguished from the true positive (the infrared small target).

In the small target segmentation step, the confidence map is the input, and the infrared small targets can be segmented from confidence map **E** based on the widely adopted adaptive threshold *Th* [4,23,27] defined in Eq. (7). The final infrared small target detection result is depicted in the 5th image and the 6th image in Fig. 3. By visually comparing the 4th and 6th images of Fig. 3, we can see that the false positives have been removed and the infrared small target is explicitly identified in the segmentation step.

$$Th = u + k \cdot \sigma \tag{7}$$

where u is the mean value of the confidence map **E**, σ is the standard deviation value of the confidence map **E**, and k is an empirical constant that is set to 10 in our implementation.

Finally, we summarize the specific calculation process of the proposed infrared small target detection approach in Algorithm 1.

3. Experimental results

In this section, we first introduce the test dataset and evaluation metrics for infrared small target detection. Next, the sensitivity of the crucial parameters of the proposed infrared small target detection approach is quantitatively analyzed in Section 3.2. Finally, a comprehensive comparison with state-of-the-art approaches is given in Section 3.3.

3.1. Test dataset and evaluation metrics

All experiments are implemented on a laptop with a 2.5 GHz Intel Core i5 CPU, 8 GB memory and using MATLAB. To fairly evaluate the performance of infrared small target detection approaches, a representative dataset comprised of six infrared sequences was constructed. With the field of view set to $4^\circ \times 3^\circ$, these infrared sequences were shot under various background conditions (e.g.,

Table 1
Test dataset.

	Sequence 1	Sequence 2	Sequence 3	Sequence 4	Sequence 5	Sequence 6
Number of Frames	100	63	100	45	125	100
Number of Key Frames	20	13	20	9	25	20
Target Category	Ship	Vehicle	Vehicle	Airplane	Airplane	Airplane and Bait
Image Resolution	228×280	236×345	160×220	256×256	452×561	150×200
Target Size	4×4	8×8	5×5	3×3	4×6	$5 \times 53 \times 3$
Meteorological Conditions	Rainy	Sunny	Cloudy	Sunny	Cloudy	Sunny
Infrared Sensor Position	Airborne Platform	High Building	Airborne Platform	Ground	Airborne Platform	Airborne Platform
Background Type	Sky-Sea	Ground	Ground	Sky	Sky	Sky
Primary Interference Type	Sensor Noise	Ground Clutter	Ground Clutter	Structure Clutter	Cloud Clutter	Sensor Noise

sea-sky, ground, and sky). In addition, the operating wavelength interval of the infrared sensor is in the longwave infrared range, 8-14*um*. For details of the constructed dataset, please refer to Table 1, where the key frames are comprised of every fifth image frame sequentially selected from the original sequences and the infrared small targets in the key frames have been manually annotated at the pixel-level. From Table 1, we can see that the six sample sequences include a variety of situations encountered in infrared small target detection. Hence, the test dataset is sufficient to evaluate the performance of infrared small target detection approaches.

Various quantitative evaluation metrics, including the SCR Gain, the background suppression factor (BSF), and the receiver operating characteristic (ROC) curve, are commonly used for infrared small target detection [4,20,27]. More specifically, the SCR Gain reflects the amplification of target signals relative to their backgrounds before and after processing, and the BSF mirrors the level of background suppression. In addition, the ROC curve represents the varying relationship between the true positive rate (TPR) and the false positive rate (FPR). Hence, SCR Gain, BSF, and ROC curve are chosen as the quantitative evaluation metrics in this paper.

3.2. Sensitivity analysis of the crucial parameters

In this section, we analyze the sensitivity of the crucial parameters, i.e., the amplification factor and patch size, and discuss the superiority of the local steering kernel (LSK) in comparison to traditional local features, such as the local binary pattern (LBP) [6] and the histogram of gradients (HOG) [7].

In our implementation, the overlap size t is set to 2, and the number of neighboring patches N is set to 8. Holding these parameters fixed, we quantitatively analyze the change in infrared small target detection performance as we vary the amplification factor α and the patch size s. The sensitivity analysis of the amplification factor is given in Section 3.2.1, the sensitivity analysis of the patch size is reported in Section 3.2.2, and the superiority of the LSK is verified in Section 3.2.3.

3.2.1. Sensitivity analysis of the amplification factor

Table 2 and Fig. 4 summarize the quantitative evaluation results on the six test sequences for different amplification factor values. From the SCR Gain and BSF results in Table 2, we can see that if α is a member of {3, 5, 7}, the proposed approach provides effective target enhancement and background suppression under various background conditions. Furthermore, the ROC curve results show that the proposed approach achieves the best performance on the six test sequences when $\alpha = 3$. Hence, in our proposed approach, the amplification factor α is set to 3.

3.2.2. Sensitivity analysis of the patch size

With amplification factor α set to 3, the quantitative evaluation results on the six test sequences for different patch sizes are as reported in Table 3 and Fig. 5. From the SCR Gain and BSF results in Table 3, we can see that for a patch size s of 3 or 5,the proposed approach is effective for target enhancement and background suppression under various background conditions. As $s = 5$ provides the proposed approach with better ROC performance than does $s = 3$, the patch size s is set to 5.Our proposed approach represents the image patch by an LSK that describes the local image structure, instead of using the raw image. As a benefit of using the LSK, the detection performance of the proposed approach is only moderately sensitive to the patch size. To a certain extent, the proper choice of a fixed patch size makes the proposed method suitable for the detection of small targets of different sizes.

Table 2
Average evaluation scores of the proposed infrared small target detection approach with variation of the amplification factor α.

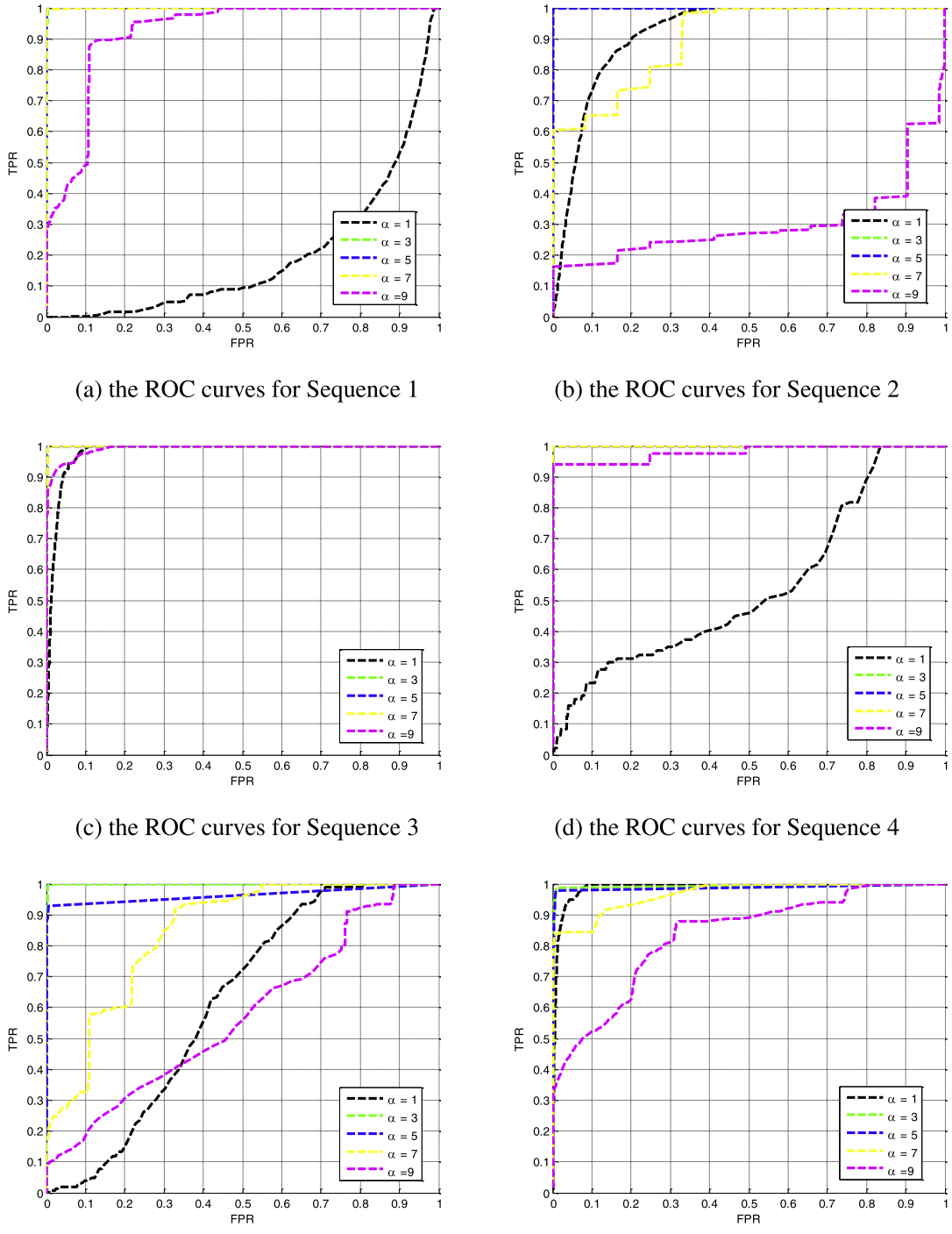
		$\alpha = 1$	$\alpha = 3$	$\alpha = 5$	$\alpha = 7$	$\alpha = 9$
Sequence 1	$\overline{\text{SCRGain}}$	0.02	158.18	207.06	216.74	209.43
	$\overline{\text{BSF}}$	3.06	21.72	28.34	30.52	27.77
Sequence 2	$\overline{\text{SCRGain}}$	1.05	26.19	26.34	17.54	0.01
	$\overline{\text{BSF}}$	2.49	17.10	20.82	18.29	3.07
Sequence 3	$\overline{\text{SCRGain}}$	4.70	27.54	22.19	20.99	17.71
	$\overline{\text{BSF}}$	0.58	2.52	2.27	2.33	2.08
Sequence 4	$\overline{\text{SCRGain}}$	0.26	30.15	26.56	25.18	13.16
	$\overline{\text{BSF}}$	1.26	7.14	7.45	7.88	6.27
Sequence 5	$\overline{\text{SCRGain}}$	0.28	29.88	30.48	20.30	2.42
	$\overline{\text{BSF}}$	2.55	23.90	28.98	22.58	7.30
Sequence 6	$\overline{\text{SCRGain}}$	7.96	43.52	40.77	37.41	11.18
	$\overline{\text{BSF}}$	2.59	13.46	13.73	12.86	4.46

Table 3
Average evaluation scores of the proposed infrared small target detection approach with variation of the patch size s.

		$s = 3$	$s = 5$	$s = 7$	$s = 9$	$s = 11$
Sequence 1	$\overline{\text{SCRGain}}$	206.31	158.18	135.27	107.17	114.50
	$\overline{\text{BSF}}$	52.07	21.72	20.20	18.79	13.89
Sequence 2	$\overline{\text{SCRGain}}$	26.02	26.19	24.11	19.97	13.37
	$\overline{\text{BSF}}$	27.04	17.10	14.10	8.25	5.70
Sequence 3	$\overline{\text{SCRGain}}$	15.58	27.54	24.57	22.84	17.27
	$\overline{\text{BSF}}$	3.63	2.52	1.63	1.27	0.99
Sequence 4	$\overline{\text{SCRGain}}$	39.27	30.15	19.55	11.88	7.46
	$\overline{\text{BSF}}$	14.25	7.13	4.94	3.63	2.66
Sequence 5	$\overline{\text{SCRGain}}$	12.59	29.88	23.18	21.37	17.31
	$\overline{\text{BSF}}$	45.62	23.90	18.97	15.77	12.98
Sequence 6	$\overline{\text{SCRGain}}$	88.82	43.52	27.22	16.38	10.90
	$\overline{\text{BSF}}$	41.34	13.46	6.63	4.86	3.93

3.2.3. Detection performances using different local features

To show the superiority of the LSK representation we test two other local features, specifically the LBP presented in [6] and the HOG presented in [7], keeping all other details of the proposed approach the same. For the LBP, the radius is set to 1 and the pattern number is set to 8. For the HOG, the number of histogram bins is set to 8 and the angle is set to 360.Thequantitative evaluation results for our approach using these parameters are summarized in Table 4 and Fig. 6. From Table 4 and Fig. 6, we can see that



(a) the ROC curves for Sequence 1

(b) the ROC curves for Sequence 2

(c) the ROC curves for Sequence 3

(d) the ROC curves for Sequence 4

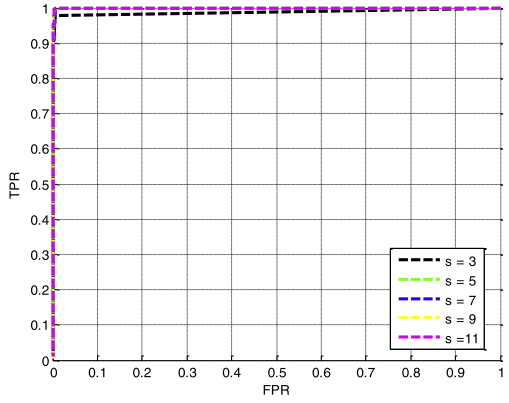
(e) the ROC curves for Sequence 5

(f) the ROC curves for Sequence 6

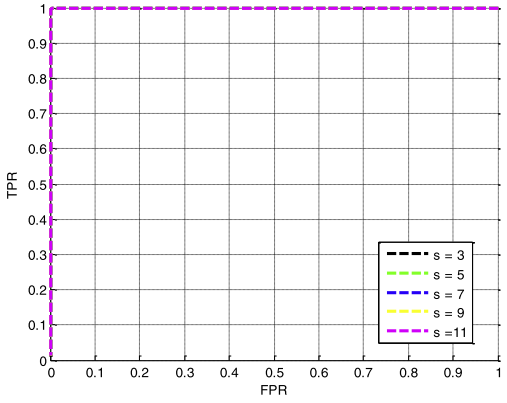
Fig. 4. The ROC curves for the six sequences with variation of the amplification factor α.

Table 4
Average evaluation scores using different local features.

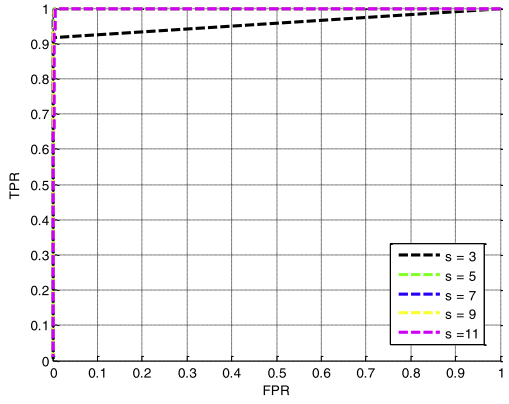
		Sequence 1	Sequence 2	Sequence 3	Sequence 4	Sequence 5	Sequence 6
$\overline{\text{SCRGain}}$	LBP	0.45	0.02	0.14	0.20	0.02	0.19
	HOG	0.07	0.01	0.40	0.01	0.08	0.01
	LSK	158.18	26.19	27.54	30.15	29.88	43.52
$\overline{\text{BSF}}$	LBP	24.63	11.92	2.70	7.20	36.43	7.84
	HOG	3.77	1.84	0.66	1.38	4.06	2.06
	LSK	21.72	17.10	2.52	7.14	23.90	13.46



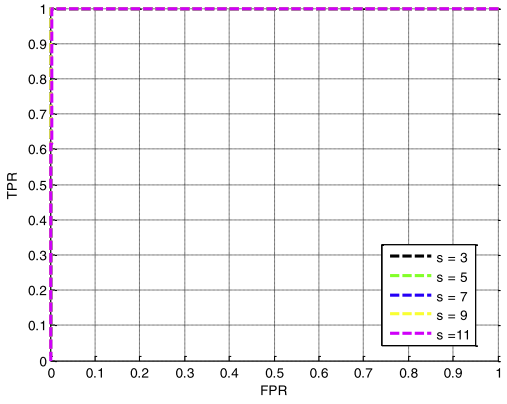
(a) the ROC curves for Sequence 1



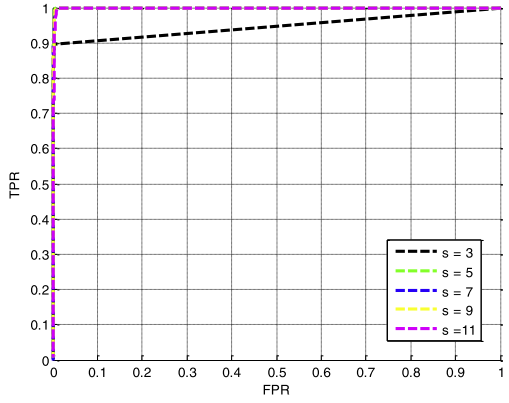
(b) the ROC curves for Sequence 2



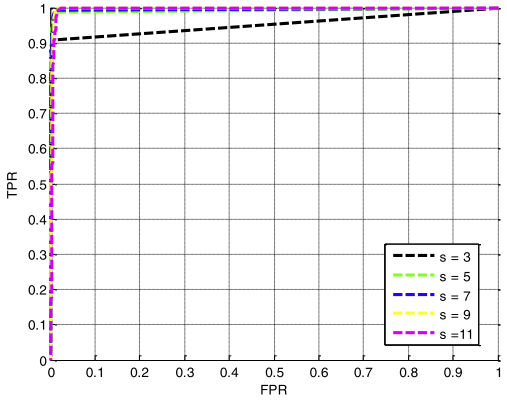
(c) the ROC curves for Sequence 3



(d) the ROC curves for Sequence 4



(e) the ROC curves for Sequence 5



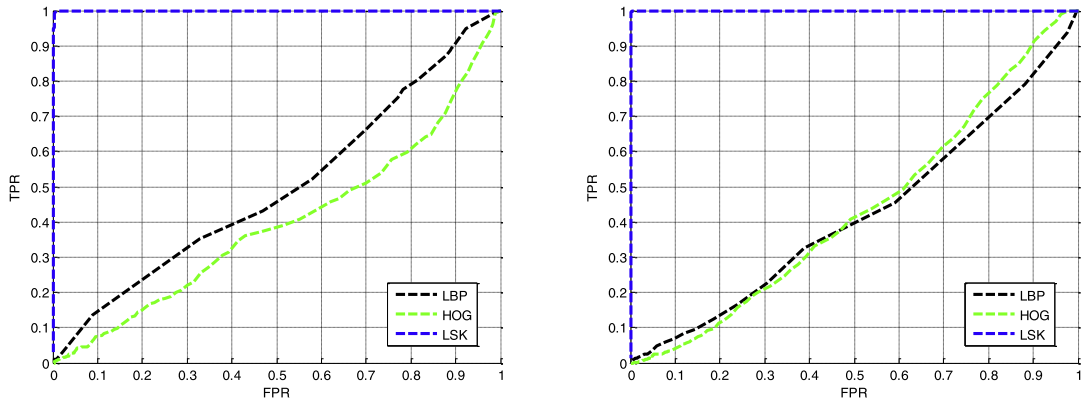
(f) the ROC curves for Sequence 6

Fig. 5. ROC curves for the six sequences with variation of the patch size *s*.

the proposed LSK noticeably outperforms the other local features in the infrared small target detection task. Hence, this comparison demonstrates the effectiveness of the LSK representation.

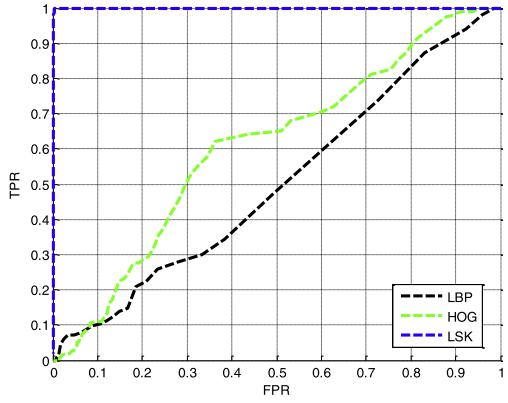
3.3. Comparison with other existing approaches

To show the effectiveness of the proposed infrared small target detection approach based on a local adaptive contrast measure using LSK reconstruction (LACM-LSK), several representative infrared small target detection methods are used as baselines for comparison. More specifically, the facet-based infrared small target detection approach (hereafter referred to as Facet) of Wang et al. [16], the LS-SVM filter-based infrared small target detection approach (LS-SVM) of Wang et al. [19], the new top-hat operator-based infrared small target detection approach (New Top-Hat) of Bai and Zhou [13], the local contrast measure-based method (LCM) of Chen et al. [23], the local adaptive contrast measure based on regularized feature reconstruction (LACRFR) of [4], and the multi-scale patch-based contrast measure (MPCM) of Wei et al. [26] are considered in this comparison.

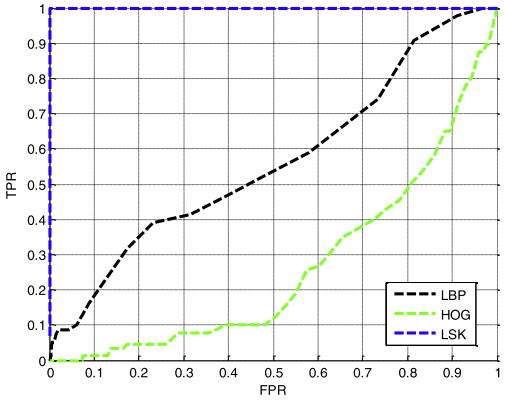


(a) the ROC curves for Sequence 1

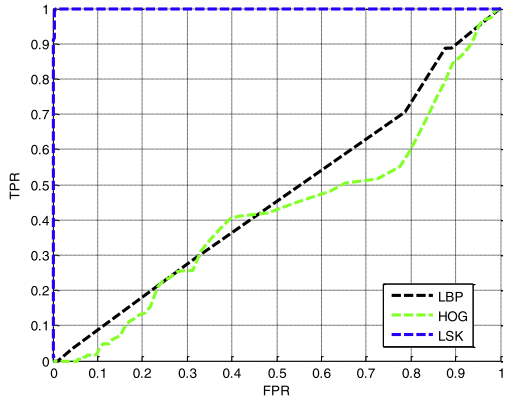
(b) the ROC curves for Sequence 2



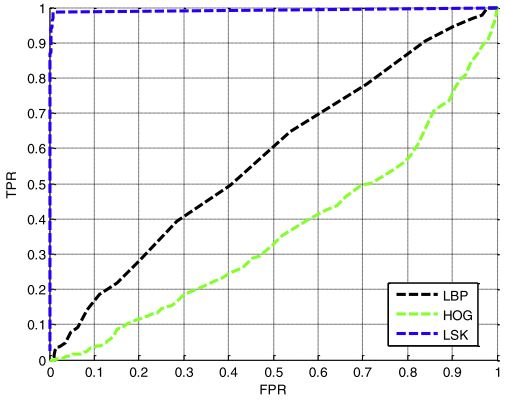
(c) the ROC curves for Sequence 3



(d) the ROC curves for Sequence 4



(e) the ROC curves for Sequence 5



(f) the ROC curves for Sequence 6

Fig. 6. ROC curves on six sequences using different local features.

Sample visual results for the various methods are shown in Fig. 7. As depicted in Fig. 7, the proposed approach can obviously outperform the baseline methods.

We also perform quantitative comparison of the proposed method and these state-of-the-art approaches. The evaluation results for the SCR Gain indicator and BSF indicator are summarized in Table 5 and the ROC curves are reported in Fig. 8.

We can see from Table 5 that LS-SVM [19], New Top-Hat [13], and MPCM [26] can occasionally achieve the best target enhancement or background suppression performance. In most situations, the proposed LACM-LSK achieves the best performance on the SCR Gain and BSF indicators. In addition, with respect to the ROC metric, the proposed approach achieves the best performance on all six sequences compared to every other existing approach, which means that the proposed approach outputs a more intact target segmentation compared to the other, state-of-the-art approaches.

To intuitively show the complexity of the proposed approach, the average per-image running times of the proposed approach

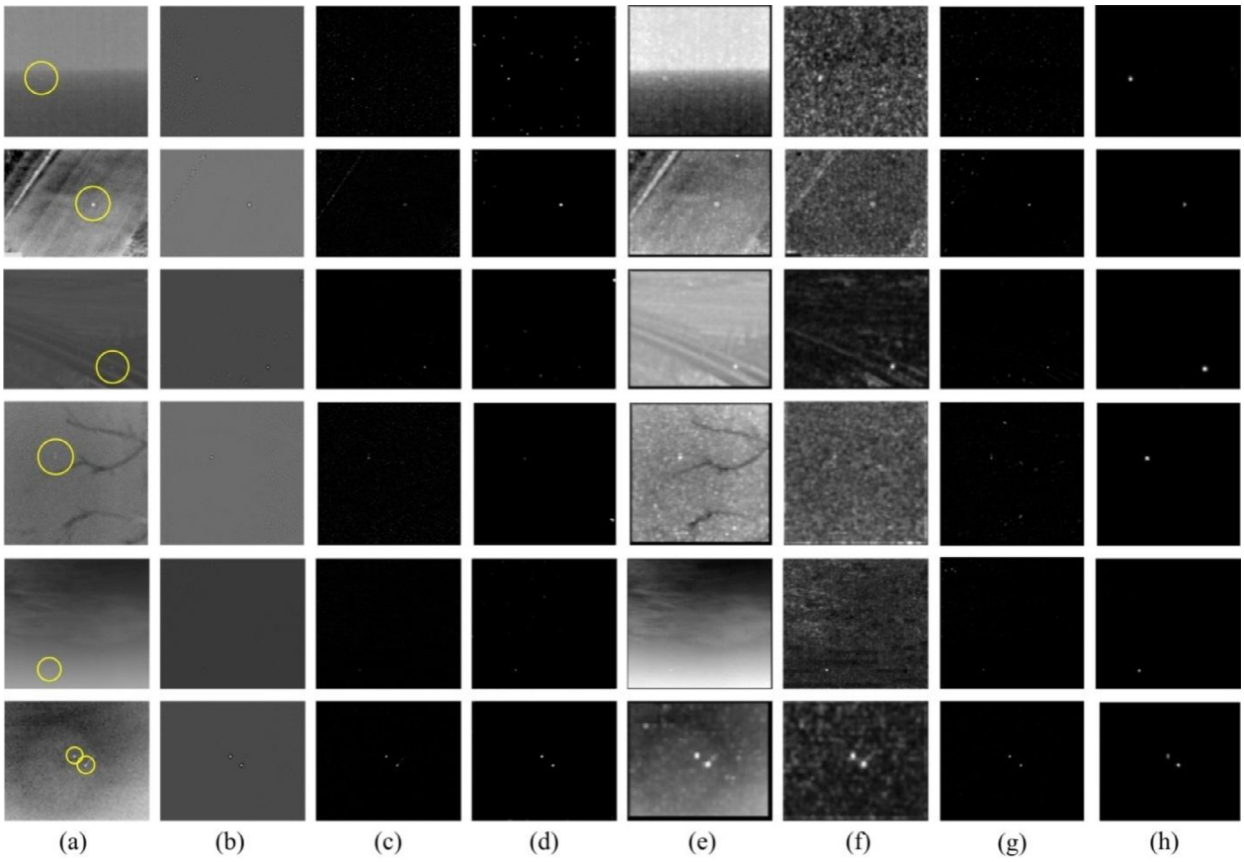
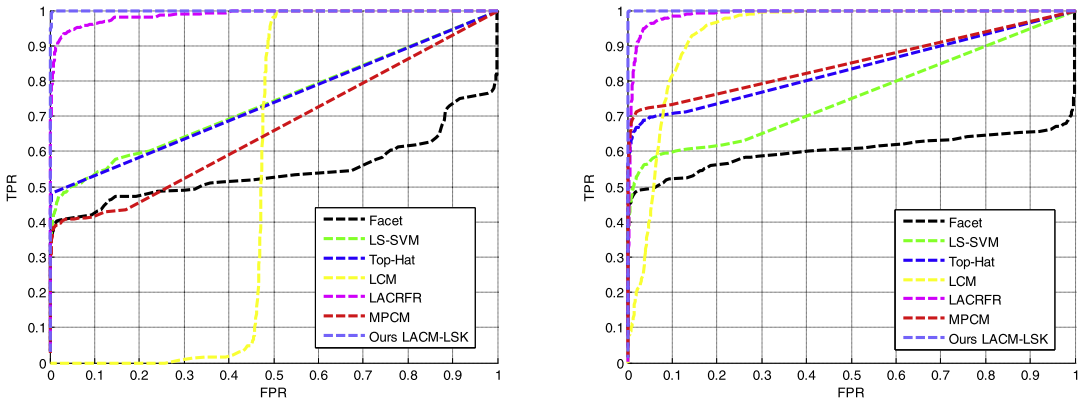


Fig. 7. The visual confidence maps for the proposed infrared small target detection approach and other existing approaches. From the 1st row to the 6th row, one representative image from Sequence 1 - Sequence 6 is used as an example for illustration. In addition, (a) - (h) present the original images, the Facet confidence maps [16], the LS-SVM confidence maps [19], the New Top-Hat confidence maps [13], the LCM confidence maps [23], the LACRFR confidence maps [4], the MPCM confidence maps [26], and the proposed LACM-LSK confidence maps, respectively. The yellow circle indicates the location of the infrared small target(s) of interest. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

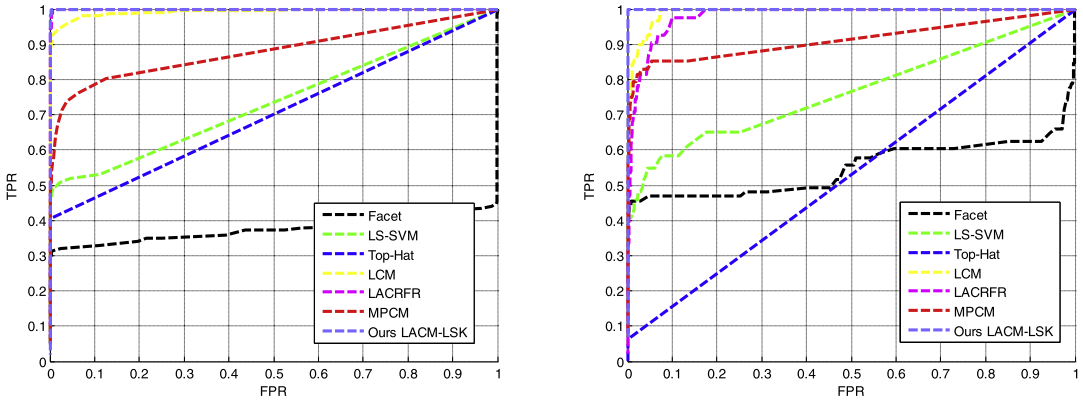
Table 5
Average evaluation scores and the average running time(s) of the proposed infrared small target detection approach and other existing infrared small target detection approaches.

		Facet in [16]	LS-SVM in [19]	Top-Hat in [13]	LCM in [23]	LACRFR in [4]	MPCM in [26]	Ours LACM-LSK
Sequence 1	$\overline{SCRGain}$	26.09	43.16	101.92	2.11	23.14	59.71	158.18
	$\overline{BSF}$	20.72	16.76	24.91	1.01	3.39	24.90	21.72
	$\overline{Time}$	0.012	0.013	0.074	3.95	4.03	0.662	4.57
Sequence 2	$\overline{SCRGain}$	1.80	3.43	22.41	1.04	2.25	10.15	26.19
	$\overline{BSF}$	8.03	5.32	14.37	1.34	1.88	13.17	17.10
	$\overline{Time}$	0.017	0.018	0.091	5.12	5.31	0.853	5.81
Sequence 3	$\overline{SCRGain}$	6.55	58.16	32.35	1.82	15.16	41.19	27.53
	$\overline{BSF}$	8.79	14.00	4.40	0.30	0.98	8.07	2.52
	$\overline{Time}$	0.009	0.010	0.045	1.87	2.01	0.423	2.26
Sequence 4	$\overline{SCRGain}$	3.69	6.94	1.54	1.43	2.36	11.30	30.15
	$\overline{BSF}$	6.19	4.91	80.99	0.75	1.37	6.97	7.14
	$\overline{Time}$	0.015	0.017	0.080	4.03	4.19	0.695	4.59
Sequence 5	$\overline{SCRGain}$	8.39	15.90	25.19	1.11	6.95	13.26	29.88
	$\overline{BSF}$	19.31	21.06	18.97	0.99	3.84	21.13	23.90
	$\overline{Time}$	0.049	0.051	0.318	8.18	8.58	2.642	9.16
Sequence 6	$\overline{SCRGain}$	5.40	53.15	85.46	4.40	12.97	84.89	43.52
	$\overline{BSF}$	47.66	63.91	51.91	1.37	3.26	91.57	13.46
	$\overline{Time}$	0.006	0.007	0.044	1.55	1.67	0.307	1.86

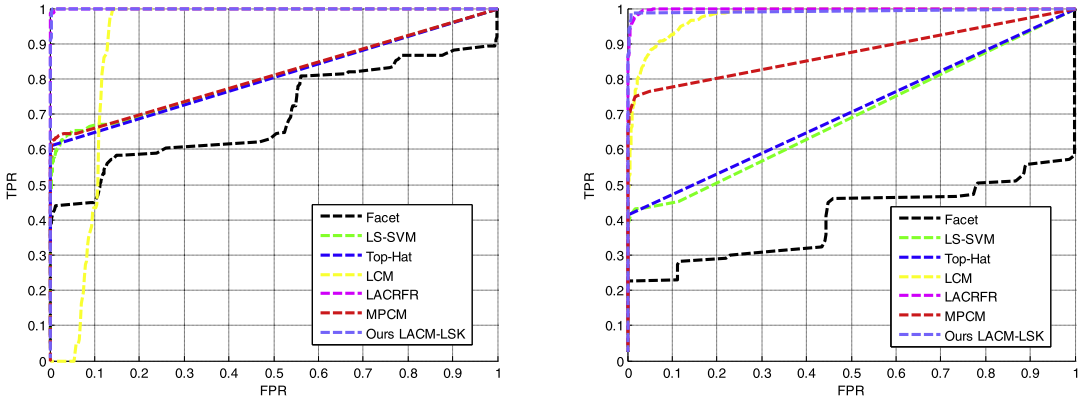
and other existing approaches are reported in Table 5. Among all methods, Facet [16] shows the most efficient performance. The complexity of the proposed approach is approximately equal to those of LCM [23] and LACRFR [4]. Fortunately, as described by Algorithm 1, the proposed approach is highly suited to parallel implementation. Hence, the proposed approach can be accelerated by a GPU or FPGA to meet the requirements of real-time applications.



(a) the ROC curves for Sequence 1 (b) the ROC curves for Sequence 2



(c) the ROC curves for Sequence 3 (d) the ROC curves for Sequence 4



(e) the ROC curves for Sequence 5 (f) the ROC curves for Sequence 6

Fig. 8. ROC curves of the proposed infrared small target detection approach and other existing infrared small target detection approaches for the six sequences.

4. Conclusions

To address the issue of often poor infrared image quality, this paper advocates the use of a local steering kernel (LSK) to encode the infrared image patch because the LSK method enjoys a good reputation for robustly capturing image structure when an image patch suffers from noise pollution or brightness variation. In fact, such situations are quite common in infrared images. Based on the robustness of LSK representation, this paper further proposes a novel local adaptive contrast measure based on LSK reconstruction (LACM-LSK) that can simultaneously enhance an infrared small target of interest and suppress any background clutter. To demonstrate the effectiveness of the proposed approach, the proposed LACM-LSK is visually and quantitatively compared to state-of-the-art approaches on six infrared image sequences shot under different background conditions. Extensive experiments show that the proposed infrared small target detection approach can significantly outperform these existing infrared small target detection approaches.

In future research, we will explore the utility of the proposed LACM-LSK in the image salient object detection task [35,36] and the video salient object detection task [37]. We will also work on

using LSK to cope with such issues as the infrared image fusion problem [38] and the infrared matching problem [39]. In addition, we will explore the utility of deep networks [40,41] in infrared small target detection.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under grants 41601352, 41322010, and 41571434; the China Postdoctoral Science Foundation under grants 2016M590716 and 2017T100581; and the Fundamental Research Funds for the Central Universitiesunder grant 2042016KF0054.

References

[1] S. Kim, J. Lee, Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track, Pattern Recognit. 43 (1) (2012) 393–406.
[2] H. Deng, Y. Wei, M. Tong, Small target detection based on weighted self-information map, Infrared Phys. Technol. 60 (2013) 197–206.
[3] Z. Li, J. Chen, Q. Hou, H. Fu, Z. Dai, G. Jin, R. Li, C. Liu, Sparse representation for infrared dim target detection via a discriminative over-complete dictionary learned online, Sensors 14 (6) (2014) 9451–9470.
[4] Y. Li, Y. Zhang, J. Yu, Y. Tan, J. Tian, J. Ma, A novel spatio-temporal saliency approach for robust dim moving target detection from airborne infrared image sequences, Inf. Sci. 369 (2016) 548–563.
[5] Y. Li, Y. Tan, H. Li, T. Li, J. Tian, Biologically inspired multilevel approach for multiple moving targets detection from airborne forward-looking infrared sequences, J. Opt. Soc. Am. A 31 (4) (2014) 734–744.
[6] T. Ahonen, A. Hadid, M. Pietikainen, Face description with local binary patterns: application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (12) (2006) 2037–2041.
[7] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: IEEE Conference on Computer Vision and Pattern Recognition, 1, 2005, pp. 886–893.
[8] Y. Li, C. Tao, Y. Tan, K. Shang, J. Tian, Unsupervised multilayer feature learning for satellite image scene classification, IEEE Geosci. Remote Sens. Lett. 13 (2) (2016) 157–161.
[9] J. Ma, J. Zhao, A. Yuille, Non-rigid point set registration by preserving global and local structures, IEEE Trans. Image Process. 25 (1) (2016) 53–64.
[10] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, J. Tian, Robust feature matching for remote sensing image registration via locally linear transforming, IEEE Trans. Geosci. Remote Sens. 53 (12) (2015) 6469–6481.
[11] S. Deshpande, M. Er, R. Venkateswarlu, P. Chan, Max-mean and max-median filters for detection of small-targets, Proc. SPIE 3809 (1999) 74–83.
[12] V. Tom, T. Peli, M. Leung, J. Bondaryk, Morphology-based algorithm for point target detection in infrared backgrounds, Proc. SPIE 1993 (1954) 2–11.
[13] X. Bai, F. Zhou, Analysis of new top-hat transformation and the application for infrared dim small target detection, Pattern Recognit. 43 (2010) 2145–2156.
[14] K. Shang, X. Sun, J. Tian, Y. Li, J. Ma, Infrared small target detection via line-based reconstruction and entropy-induced suppression, Infrared Phys. Technol. 76 (2016) 75–81.
[15] H. Deng, X. Sun, M. Liu, C. Ye, X. Zhou, Infrared small-target detection using multiscale gray difference weighted image entropy, IEEE Trans. Aerosp. Electron. Syst. 52 (1) (2016) 60–72.
[16] G. Wang, C. Chen, X. Shen, Faced-based infrared small target detection method, Electron. Lett. 41 (2005) 1244–1246.
[17] C. Yang, J. Ma, M. Zhang, S. Zheng, X. Tian, Multiscale facet model for infrared small target detection, Infrared Phys. Technol. 67 (2014) 202–209.
[18] S. Qi, G. Xu, Z. Mou, D. Huang, X. Zheng, A fast-saliency method for real-time infrared small target detection, Infrared Phys. Technol. 77 (2016) 440–450.
[19] P. Wang, J. Tian, C. Gao, Infrared small target detection using directional high-pass filters based on LS-SVM, Electron. Lett. 45 (2009) 156–158.
[20] S. Qi, J. Ma, C. Tao, C. Yang, J. Tian, A robust directional saliency-based method for infrared small-target detection under various complex backgrounds, IEEE Geosci. Remote Sens. Lett. 10 (3) (2013) 495–499.
[21] S. Qi, J. Ma, H. Li, S. Zhang, J. Tian, Infrared small target enhancement via phase spectrum of quaternion fourier transform, Infrared Phys. Technol. 62 (2014) 50–58.
[22] W. Wang, C. Li, J. Shi, A robust dim target detection based on template filtering and saliency extraction, Infrared Phys. Technol. 73 (2015) 19–28.
[23] C.P. Chen, H. Li, Y. Wei, T. Xia, Y. Tang, A local contrast method for small infrared target detection, IEEE Trans. Geosci. Remote Sens. 52 (1) (2014) 574–581.
[24] J. Han, Y. Ma, B. Zhou, F. Fan, K. Liang, Y. Fang, A robust infrared small target detection algorithm based on human visual system, IEEE Geosci. Remote Sens. Lett. 11 (12) (2014) 2168–2172.
[25] H. Deng, X. Sun, M. Liu, C. Ye, X. Zhou, Small infrared target detection based on weighted local difference measure, IEEE Trans. Geosci. Remote Sens. 54 (7) (2016) 4204–4214.
[26] Y. Wei, X. You, H. Li, Multiscale patch-based contrast measure for small infrared target detection, Pattern Recognit. 58 (2016) 216–226.
[27] C. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, A. Hauptmann, Infrared patch-image model for small target detection in a single image, IEEE Trans. Image Process. 22 (12) (2013) 4996–5009.
[28] Y. Wei, X. You, H. Deng, Small infrared target detection based on image patch ordering, Int. J. Wavelets Multiresolution Inf. Process. 14 (2) (2016) 1640007–1640020.
[29] H. Takeda, S. Farsiu, P. Milanfar, Kernel regression for image processing and reconstruction, IEEE Trans. Image Process. 16 (2) (2007) 349–366.
[30] H. Takeda, P. Milanfar, M. Protter, M. Elad, Super-resolution without explicit subpixel motion estimation, IEEE Trans. Image Process. 18 (9) (2009) 1958–1975.
[31] H.J. Seo, P. Milanfar, Static and space-time visual saliency detection by self-resemblance, J. Vision 9 (12) (2009) 1–27.
[32] H.J. Seo, P. Milanfar, Training-free, generic object detection using locally adaptive regression kernels, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2010) 1688–1704.
[33] Xinge You, Qiang Li, Dacheng Tao, Weihua Ou, Mingming Gong, Local metric learning for exemplar-based object detection, IEEE Trans. Circuits Syst. Video Technol. 24 (8) (2014) 1265–1276.
[34] Yanni Dong, Bo Du, Liangpei Zhang, Target detection based on random forest metric learning, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 8 (4) (2015) 1830–1838.
[35] Y. Tan, Y. Li, C. Chen, J. Yu, J. Tian, Cauchy graph embedding based diffusion model for salient object detection, J. Opt. Soc. Am. A 33 (5) (2016) 887–897.
[36] Q. Peng, Y. Cheung, X. You, Y. Tang, A hybrid of local and global saliencies for detecting image salient region and appearance, IEEE Trans. Syst. Man Cybern. 47 (1) (2017) 86–97.
[37] Y. Li, Y. Tan, J. Yu, S. Qi, J. Tian, Kernel regression in mixed feature spaces for spatio-temporal saliency detection, Comput. Vision Image Understanding 135 (2015) 126–140.
[38] J. Ma, C. Chen, C. Li, J. Huang, Infrared and visible image fusion via gradient transfer and total variation minimization, Inf. Fusion 31 (2016) 100–109.
[39] C. Yang, H. Zhou, S. Sun, R. Liu, J. Zhao, J. Ma, Good match exploration for infrared face recognition, Infrared Phys. Technol. 67 (2014) 111–115.
[40] Y. Li, Y. Zhang, X. Huang, H. Zhu, J. Ma, Large-scale remote sensing image retrieval by deep hashing neural networks, IEEE Trans. Geosci. Remote Sens. PP (99) (2017) 1–16, doi:10.1109/TGRS.2017.2756911.
[41] Y. Li, X. Huang, H. Liu, Unsupervised deep feature learning for urban village detection from high-resolution remote sensing images, Photogramm. Eng. Remote Sens. 83 (8) (2017) 567–579.

Yansheng Li received the B.S. degree from the School of Mathematics and Statistics, Shandong University, Weihai, China, in 2010, and the Ph.D. degree from the School of Automation, Huazhong University of Science and Technology, Wuhan, China, in 2015. Since 2015, he has been an Assistant Professor with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan. Currently, he is a Visiting Assistant Professor with the Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA; he will hold the position till 2018. He has authored more than 20 peer-reviewed articles in international journals from multiple domains such as remote sensing and computer vision. His research interests include computer vision, machine learning, deep learning, and their applications in remote sensing.

Yongjun Zhang received the B.S., M.S., and Ph.D. degrees from Wuhan University, Wuhan, China, in 1997, 2000, and 2002, respectively. He is currently a Professor of photogrammetry and remote sensing with the School of Remote Sensing and Information Engineering, Wuhan University. His research interests include space, aerial, and low-attitude photogrammetry; image matching; combined bundle adjustment with multisource data sets; 3-D city reconstruction; and industrial inspection.