Pattern Recognition 61 (2017) 66–77

Contents lists available at ScienceDirect

Pattern Recognition

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

Entropy-based window selection for detecting dim and small infrared targets

He Deng, Xianping Sun, Maili Liu, Chaohui Ye, Xin Zhou*

State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, National Center for Magnetic Resonance in Wuhan, Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences, Wuhan 430071, China

A R T I C L E I N F O

Article history:
Received 10 August 2015
Received in revised form
8 June 2016
Accepted 23 July 2016
Available online 26 July 2016

Keywords:
Dim and small target detection
Infrared image
Local difference measure
Window selection

A B S T R A C T

Dim and small target detection in complex background is considered a difficult and challenging problem. Conventional algorithms using the local difference/mutation possibly produce high missed or mistaken detection rates. In this paper, we propose an effective algorithm for detecting dim and small infrared targets. In order to synchronously enhance targets and suppress complex background clutters, we adopt an adaptive entropy-based window selection technique to construct a novel local difference measure (LDM) map of an input image, which measures the dissimilarity between the current region and its neighboring ones. In this way, the window size can be adaptively regulated according to local statistical properties. Compared with the original image, the LDM map has less background clutters and noise residual. This guarantees the lower false alarm rates under the same probability of detection. Subsequently, a simple threshold is used to segment the target. More than 600 dim and small infrared target images against different complex and noisy backgrounds were utilized to validate the detection performance of the proposed approach. Extensive experimental results demonstrate that the proposed method not only works more stably for different target movements and signal-to-clutter ratio values, but also has a better performance compared with classical baseline methods. The evaluation results suggest that the proposed method is simple and effective with regard to detection accuracy.

1. Introduction

The performance of detection, identification and tracking of dim and small targets restricts the development of infrared search and tracking (IRST) systems, although some relevant techniques have been widely applied in military/civilian fields, such as precise guidance, early warning, geological analysis, and industrial flaw detection [1–12]. The main challenge is that there is scarcely any prior knowledge about the target shape, target size, and textural features that could be utilized in the detection, identification and tracking [3]. Against complicated backgrounds, those problems will be deteriorated because the target intensities are weak, partially obscured by jamming objects, or buried in clutters and noise. In these cases, it is difficult to separate targets from complex and noisy backgrounds.

There exist numerous methods for detecting dim and small infrared targets. These algorithms could be generally classified into two groups: track-before-detect (TBD) [8] and detect-before-track (DBT) methods [2]. Compared with TBD methods, DBT methods are more powerful because of shorter computation time, and fewer requirements of assumptions and prior knowledge. In general, two successive procedures are involved in most of DBT methods [1,9]: the pre-detection procedure in a single-frame image and the tracking procedure in multiple-frame images. The result of the former procedure impacts both the computation and accuracy of the latter one, especially for applications with fast moving sensor platforms (for example, airplane- or missile-based IRST systems) [3].

Generally, dim and small target detection methods are founded on many assumptions on the target, background, or both of them. The robustness of target detection methods is dependent on the appropriateness of assumptions in applications. There are two assumptions adopted in the majority of DBT methods [6]: the first is that the background has the correlation in spatial domain and the stability in time domain, and it occupies the low frequency portion of an infrared image in frequency domain. The second assumption is that the target is unrelated to the background and noise, and it dominates the high frequency portion of the image. Accordingly, some DBT methods based on these assumptions are widely used to eliminate background clutters, such as methods based on the finite or infinite impulse response filter, median filter,

* Corresponding author.
E-mail addresses: denghe@wipm.ac.cn (H. Deng), xpsun@wipm.ac.cn (X. Sun), ml.liu@wipm.ac.cn (M. Liu), ye@wipm.ac.cn (C. Ye), xinzhou@wipm.ac.cn (X. Zhou).

http://dx.doi.org/10.1016/j.patcog.2016.07.036

Top-hat filter [1], Max-mean filter, Max- median filter [13], or space-time maximum likelihood [14]. However, these methods are only suitable for those backgrounds whose statistical characteristics are constant or slowly varying. And then, some methods, such as the Butterworth high-pass filter [15] or wavelet transform [16], are utilized to deal with nonlinear, non-stationary or rapidly varying backgrounds. Recently, the local dependency histogram [17] and online dictionary learning [18] are applied to detect moving objects. Moreover, classification-based approaches are developed to efficiently eliminate various clutter points [19], including the nearest neighbor classifier [20], learning-based neural network [21], and manifold learning [22]. There are still many other algorithms for detecting dim and small targets, such as the algorithms based on the tri-feature-based detector [11], statistical regression [12], or biological vision [23].

Since a dim and small target in an image often occupies several pixels, the target birth causes remarkable changes of local rather than global textural characteristics. According to the local difference between the target and neighboring background clutters, some methods based on the probabilistic principal component analysis [24], empirical mode decomposition [25], or sparse ring representation [9] are developed to detect targets. In addition, some operators are presented to measure the local mutation/difference owing to the target occurrence, such as the local contrast map (LCM) [4], the local mutation weighted information entropy (LMWIE) [26], and the average gray absolute difference maximum map (AGADM) [27].

Since the choice of window (block, patch, sub-image, or neighborhood) size utilized in the above approaches is a difficult problem, the window size is usually predefined and is invariant in representing the local difference/mutation. However, if the window size is too large, those slight changes of textural characteristics may be undetected and the operation will consume more computation. Whereas in a window with too small size, some jamming pixels may be detected as target pixels, such as noise or pixels affected by illumination variation or other factors. Then, a proper window size is helpful to discriminate true targets from jamming objects and clutters. Furthermore, a target has a conspicuous discontinuity with its neighboring areas and concentrates in a relatively small, homogeneous and compact region. Where the discontinuity is essentially involved ascertaining the property of average gray difference founded on the neighboring pixels [22,27]. Hence, we utilize an adaptive entropy-based window selection technique to construct a measure that represents the local difference between the target and neighboring background clutters. In this way, the window size can be adaptively regulated according to the local statistical characteristics. After the local difference measure, the local region whose difference is larger than a given threshold in some scale may be a position where the target emerges. With these considerations in mind, we design a method based on the novel local difference measure (LDM) to detect dim and small targets submerged in complex and noisy backgrounds in this paper.

The contributions of this paper can be summarized as follows: 1) An adaptive entropy-based window selection technique is proposed in the construction of local difference measure. The window size can be adaptively adjusted according to the local statistical properties, rather than preset to some fixed value and invariant across the frames. This can lower missed and false detection rates. 2) A new local difference measure is utilized to enhance targets and suppress background clutters and noise synchronously. The LDM map can significantly improve SCR values of the image and have little clutters and noise residual. This guarantees low false alarm rates under the same probability of detection. 3) A LDM-based detection method is designed to process infrared images with low SCR values. By applying such method on more than 600 low SCR images with diverse complicated backgrounds, it demonstrates that the designed method not only works more robustly, but also has better detection performance in comparison to well-known baseline methods.

The remainder of this paper is organized as follows: In Section 2, we explain the LDM-based target detection method in detail. In Section 3, we give extensive experimental results and discussions. And Section 4 reaches a conclusion.

2. Dim and small target detection based on LDM

In this section, we introduce a new scheme for detecting dim and small targets embedded in intricate backgrounds. This scheme initially adopts an adaptive entropy-based window selection technique to found a LDM map of an input image that achieves target enhancement and background suppression at the same time, and subsequently utilizes a simple threshold to segment targets from the LDM map. In particular, the LDM scheme can improve SCR values of the image significantly.

2.1. Entropy-based window selection

Some evidences [4,27] have indicated that a dim and small target concentrates in a small, homogeneous and compact area. The target has brightness discontinuity with its surrounding background clutters from a natural scene. And this discontinuity is fundamentally involved determining the property of average gray difference according to neighboring pixels [27]. If a proper measure is designed to measure the discontinuity between the current region and its neighboring ones in the scene, the most dissimilar point will be considered as a target. Owing to the change of imaging distance, the target size varies within a range, which suggests that the window (neighborhood or patch) size with respect to the target should be adaptively adjusted in the design of the measure [3,4]. Moreover, the noise and clutter levels change across the infrared image frames, which also imply that the widow size should be adaptively regulated to accommodate the changes in the noise and clutter from frame- to-frame. However, the prior knowledge, such as the target size, target velocity, target shape, or noise and clutter levels, is impossibly acquired in IRST applications. In this case, we utilize an adaptive entropy-based window selection technique to appropriately regulate the size of the neighboring window surrounding the target. Subsequently, a multi-scale representation model is adopted to measure the discontinuity between the target region and neighboring background clutters.

Noticed that the information content of a heterogeneous window in an image is proportional to the fraction of entropy of the window, patch or sub image [29]. Accordingly, an idea of the window selection (viz., window size) primarily depends on the proportion of information content contained in the window and whole image [29]. In this section, the concept of local entropy is adopted to represent the information content contained in a window because it tells how much information there is in an information source. For a neighboring window, its local entropy is defined as,

$$E_w = -\sum_{i=0}^{S-1} p_i \log_2 p_i, \tag{1}$$

where pi is the probability density function of the i-th gray level in the window and S is the maximum gray level. The local entropy is involved in the variance of gray values in the window [29]. It is large for a heterogeneous area but small for a homogeneous one. The appearance of dim and small target enriches the gray value information in a local area, which means that the target region has more local entropy than that in background areas.

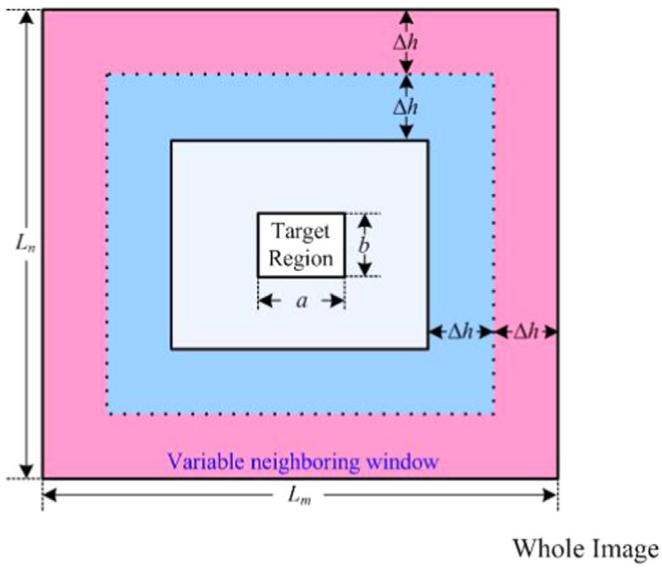

Fig. 1. Variable neighboring window.

After that, if the local entropy of a neighboring window (see Fig. 1) is comparable to some fraction of the entropy of the whole image, that is,

$$E_w \geq \lambda \cdot E_D, \tag{2}$$

the window size will be fixed. Where λ is a constant in an interval [0,1], E_w is the local entropy of the neighboring window, and E_D is the entropy of the whole image. Otherwise, the window size is incremented by Δh until the above condition is satisfied. The window growing process is displayed in Fig. 1, where the initial size of target region is assumed to be $a \times b$, the final neighboring window size is $L_m \times L_n$, and the neighboring window is enlarged by an increment Δh, respectively.

According to Eq. (2), the size of a homogeneous neighboring window is larger than that of a heterogeneous one. Then the window surrounding a target is with small size because of the target birth. Besides, the neighboring window can move in the image from top to bottom and form left to right. As a result, we determine the window size centered at each pixel point in the image once the parameter lambda in Eq. (2) is preset.

In summary, the entropy-based window selection scheme is described in Algorithm 1, where E_w is the local entropy of the neighboring window, E_D is the entropy of the whole image, L_m and L_n are positive integers, λ is a constant in [0,1], and Δh is an increment (a small integer), respectively. In general, the user can manually/experimentally select the coefficient lambda for practical design requirements, which will be further discussed in detail in Section 2.5.

Algorithm 1. Entropy-based window selection scheme.

Input: Given a pixel point (x,y) in an image.
Output: A neighboring window centered at the point.

1. Enough scales (L_m and L_n) of the neighboring window (E_w) are given.
2. Enough scale of the parameter (λ) is given.
3. Compute the whole entropy of the image (E_D).
4. Compute the local entropy of the neighboring window (E_w) according to Eq. (1).
5. **if** $E_w \geq \lambda \cdot E_D$, then
Fix the window size as $L_m \times L_n$.
6. **else**
7. Enlarge the window size by Δh.
8. $L_m \times L_n \leftarrow (L_m + 2 \cdot \Delta h) \times (L_n + 2 \cdot \Delta h)$.
9. **end if**

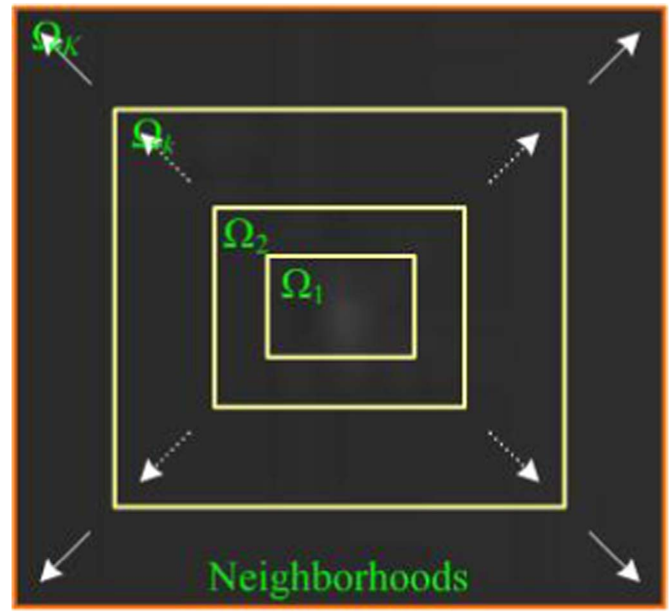

Fig. 2. Dilation of target region.

2.2. LDM

Based on the above, the determination of neighboring window size depends upon the information content contained in the window. Once the window size is assigned, denoted as $L_m \times L_n$, the target region can dilate in the window from $a \times b$ to $L_m \times L_n$, where $a \times b$ denotes the initial target size (see Fig. 2). Hence, we can acquire a series of neighborhoods surrounding the target center, Ω_k, $k = 1,2,\ldots,K$, where K is the maximum scale of the fixed neighboring window.

As a matter of fact, the discontinuity between the target region and surrounding background clutters is relevant to the average gray difference based on neighboring pixels [27]. The average gray value of the k-th neighborhood is defined as,

$$C_k = \frac{1}{N_k} \sum_{j=1}^{N_k} g_j^k, \text{ where } k = 1, 2, \ldots, K, \tag{3}$$

where N_k is the number of pixels contained in the k-th neighborhood Ω_k and g_j^k is the gray level of the j-th pixel in Ω_k.

The local difference between the target region and the k-th neighborhood is expressed as the following formula.

$$\hat{E}_k = |C_1 - C_k|^2 / |C_{\max} - C_{\min}|^2$$
$$\text{where } C_{\max} = \max_{k=1,\ldots,K} C_k,\ C_{\min} = \min_{k=1,\ldots,K} C_k, \tag{4}$$

Then, the LDM is defined as,

$$\hat{E} = \max\left\{0, \hat{E}_2, \cdots, \hat{E}_K\right\}, \tag{5}$$

The LDM represents the maximum discontinuity between the current region and its neighboring ones. When the current region moves in the image from top to bottom and form left to right, a

LDM map is constructed. It suggests that the larger LDM is, the more likely a target appears. Based on this fact, the method to compute the LDM is described in Algorithm 2, where the size of an input image is $M \times N$, $\hat{E}$ is the LDM, and K is the largest scale, respectively.

Algorithm 2. Local difference measure.
Input: Given a frame.
Output: LDM map.

1. **for** $x = 1:M$ **do**
2. **for** $y = 1:N$ **do**

Determine the neighboring window according to Algorithm 1.

3. Acquire a set of neighbourhoods, Ω_k, $k=1,2,\dots,K$.
4. **for** $k=1,2,\dots,K$, **do**

Compute the k-th local difference measure according to Eqs. (3,4).

5. **end for**
6. $\hat{E}(x,y) = \max\{0, \hat{E}_2(x,y),\dots, \hat{E}_K(x,y)\}$.
7. Replace the value of central pixel with $\hat{E}(x,y)$.
8. **end for**
9. **end for**

2.3. LDM-based dim and small target detection

We know that a dim and small target is discontinuous with its neighboring areas and concentrates in a small, homogeneous and compact region. While the background is consistent with its neighborhoods [4,27]. The LDM broadens the discontinuity between the target region and surrounding background clutters. As a result, the target is well enhanced and the background clutters and noise are effectively suppressed (see Fig. 3). Accordingly, we conceive that, if the final LDM map is achieved, the most salient point in the scene is probably a target. The LDM-based dim and small target detection method is described in Algorithm 3, where t_1 and t_2 denote the mean and maximum values of the final LDM map, and ε is a constant. In order to intuitively show the proposed method, a target detection system is given in Fig. 3.

Algorithm 3. LDM-based target detection method.
Input: One frame.
Output: Target position.

1. Obtain LDM map according to Algorithm 2.
2. Compute the threshold according to

$$T = \varepsilon \cdot t_1 + (1-\varepsilon)\cdot t_2. \tag{6}$$

3. Segment targets from the LDM map according to T.

2.4. Detection ability analysis

From the definition, it can be found that the LDM is able to enhance the dim and small target and suppress the background clutters and noise simultaneously (See Fig. 3). Let (x_0,y_0) be the center pixel point of the target, its LDM can be expressed as,

$$\begin{aligned}
&\hat{E}^t(x_0,y_0) = \max\left\{0, \hat{E}_2^t(x_0,y_0), \cdots, \hat{E}_K^t(x_0,y_0)\right\}\\
&\text{where } \hat{E}_k^t(x_0,y_0) = \left|C_1^t(x_0,y_0) - C_k^t(x_0,y_0)\right|^2 / \left|C_{\max}^t(x_0,y_0) - C_{\min}^t(x_0,y_0)\right|^2\\
&\text{and } C_{\max}^t(x_0,y_0) = \max_{k=1,\dots,K} C_k^t(x_0,y_0),\ C_{\min}^t(x_0,y_0) = \min_{k=1,\dots,K} C_k^t(x_0,y_0)\\
&\text{and } C_k^t(x_0,y_0) = \frac{1}{N_k^t}\sum_{j\in\Omega_k^t}\left(g_j^k\right)^t,
\end{aligned} \tag{7}$$

where K is the maximum neighboring window scale according to Algorithm 1, Ω_k^t is the k-th neighborhood, $k=1,2,\dots,K$, N_k^t is the number of pixels contained in Ω_k^t, and $(g_j^k)^t$ is the gray level of the j-th pixel in Ω_k^t, respectively.

In a dim and small infrared target image, the target is unrelated to background in spatial domain. There exists brightness difference between the target and surrounding background clutters despite that the discrimination is small. Then, for a bright target, it can be found that,

$$C_k^t(x_0,y_0) \ge C_{k+1}^t(x_0,y_0), \text{ where } k = 1, 2, \dots, K-1, \tag{8}$$

Thus,

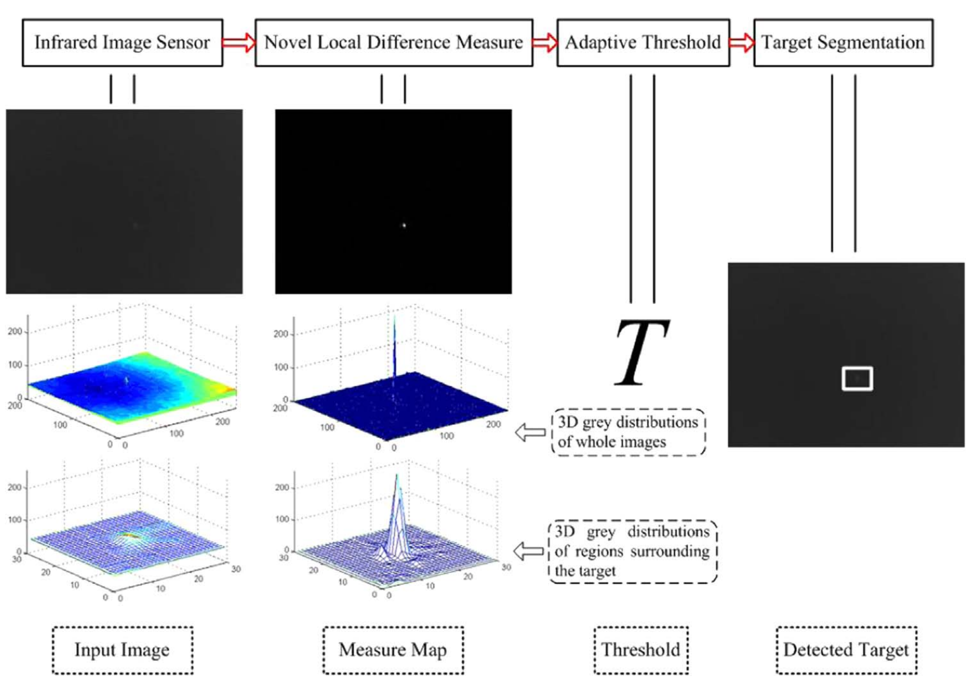

Fig. 3. Proposed target detection system.

$$C^t_{\max}(x_0, y_0) = C^t_1(x_0, y_0), \text{ and } C^t_{\min}(x_0, y_0) = C^t_K(x_0, y_0), \tag{9}$$

Therefore,

$$\hat{E}^t(x_0, y_0) \cong 1, \tag{10}$$

For a dark target, we can find that,

$$\tilde{C}^t_k(x_0, y_0) \le \tilde{C}^t_{k+1}(x_0, y_0), \text{ where } k = 1, 2, \dots, K-1, \tag{11}$$

Likewise, the following expression holds for this case.

$$\hat{E}^t(x_0, y_0) \cong 1, \tag{12}$$

Accordingly, the local difference measure is distinct in the bright/dark target region. If the neighboring window size (viz., the maximal scale of neighborhoods) is selected appropriately, the LDM value will be near to 1 in the target region.

On the other hand, if the current location is a pixel point in the background, there exists minor local difference among its neighborhoods because the background has the correlation with the neighboring areas in spatial domain. Then, the following relationship may be tenable.

$$C^b_k(x_0, y_0) \cong C^b_{k+1}(x_0, y_0), \text{ where } k = 1, 2, \dots, K-1, \tag{13}$$

After that,

$$\hat{E}^b(x_0, y_0) \cong 0, \tag{14}$$

In this way, the LDM value in a target region is greater than that in a background region. Consequently, the target can be enhanced and the background can be suppressed effectively. This reveals that the LDM map takes into account the problems of target enhancement and background suppression simultaneously. The LDM-based dim and small target detection method can work well for infrared images with diverse complicated backgrounds.

2.5. Discussion

There are two coefficients (viz., the parameter lambda in Eq. (2) and the parameter epsilon in Algorithm 3) involved in the LDM-based target detection scheme, which provides design simplicity and flexibility to accommodate specific requirements in applications. These coefficients can be manually/experimentally chosen. Nevertheless, this is a time-consuming approach and difficult to achieve the best results owing to the criterion lack for quantitative evaluation. Alternatively, the coefficients can be simply selected according to some reasonable assumptions.

The maximum neighboring window scale depends on the parameter lambda. If the window size is too large, those slight change of textural features may be undetected. And the operation consumes more computation. On the other hand, if the window size is too small, the potential target is possibly covered by the background. Owing to the long imaging distance, the target size is generally small in an image. A dim and small target defined by Society of Photo-Optical Instrumentation Engineers (SPIE) has a total spatial extent of less than 80 pixels [4]. This category consists of point source targets, small extended targets, as well as clusters of point source and small extended targets [4,30]. As a result, a dim and small target occupies less than 0.15% of an input image with size of 256×256. This criterion is usually independent of the image size [4]. Hence, the appropriate window size in the target region is no larger than 81 pixels. This is our assumption in experiments, that is, the neighboring window size in the target region is set to be 9×9. Our assumption fits the SPIE definition well, and is valid for most cases.

Moreover, the target birth enriches the gray information in a local area. Because the local entropy is related to the variance of gray values in the window/sub image [28], the local entropy value in target region is higher than that in background area. If the local entropy map of an image is achieved (In our experiments, the window size in computation of local entropy is set to be 9×9), it is likely that the most salient point in this case is a target point. This offers a strategy to determine the lambda in Eq. (2). Across the frames in an image sequence which may have varying noise and clutter levels, the parameter lambda can be adjusted, in an adaptive fashion, to accommodate the changes in the noise and clutter. We have done a large number of data experiments. The evaluation results suggest that the lambda selection based on the above strategy can bring satisfying detection performance.

In comparison to the lambda, the parameter epsilon in Algorithm 3 has less influence on the performance of LDM-based target detection scheme. The reason may be that the LDM map arouses little clutters and noise residual, especially improves SCR values of the image significantly (See Section 3 for the detailed discussion). Hence, the target is easily segmented in the LDM map through a simple threshold. In the target segmentation, we usually select the parameter epsilon in an interval [0.5, 0.65], which produces satisfactory results.

3. Experimental results

In this section, we firstly introduce the evaluation metrics, baseline methods and data for comparison. After that, we utilize several infrared image sequences against diverse backgrounds to demonstrate the effectiveness and feasibility of our method.

3.1. Metrics, baseline methods and data

The major task in detecting a dim and small target embedded in a complicated background is to availably suppress noise and background clutters, and then significantly enhance the target [2]. Less clutter and noise residual is a crux to keep lower false alarm rates under the same probability of detection [3]. Thus, if our method can eliminate more background clutters and noise, the target will be detected more readily.

We adopt the SCR [3] and BSF (background suppression factor) [26] as two metrics in comparison of target enhancement performance, because they can measure the residual degree of clutters and noise after different detection methods. Normally, the higher SCR and BSF values are, the easier a target can be detected. The definitions of SCR and BSF are expressed as,

$$SCR = |m_t - m_b|/\sigma_b, \; BSF = s_1/s_2, \tag{15}$$

where m_t denotes the mean gray value in the target region, m_b and σ_b denote the mean and standard deviation of gray values in a neighboring region surrounding the target, and s_1 and s_2 denote the standard deviations of gray values in the original and filtered images, respectively.

The probability of detection (Pd) and the false alarm rate (Fa) are generally utilized to evaluate the detection performance of diverse methods [3,31]. The definitions of Pd and Fa are described as,

$$Pd = n_t/n_c, \; Fa = n_f/n, \tag{16}$$

where n_t, n_c, n_f, and n denote the number of true detections, the number of actual targets, the number of false detections, and the number of images, respectively.

Because our method is to represent the local mutation in an image causing by the appearance of a dim and small target, we select some local descriptors that have a similar goal to ours, as

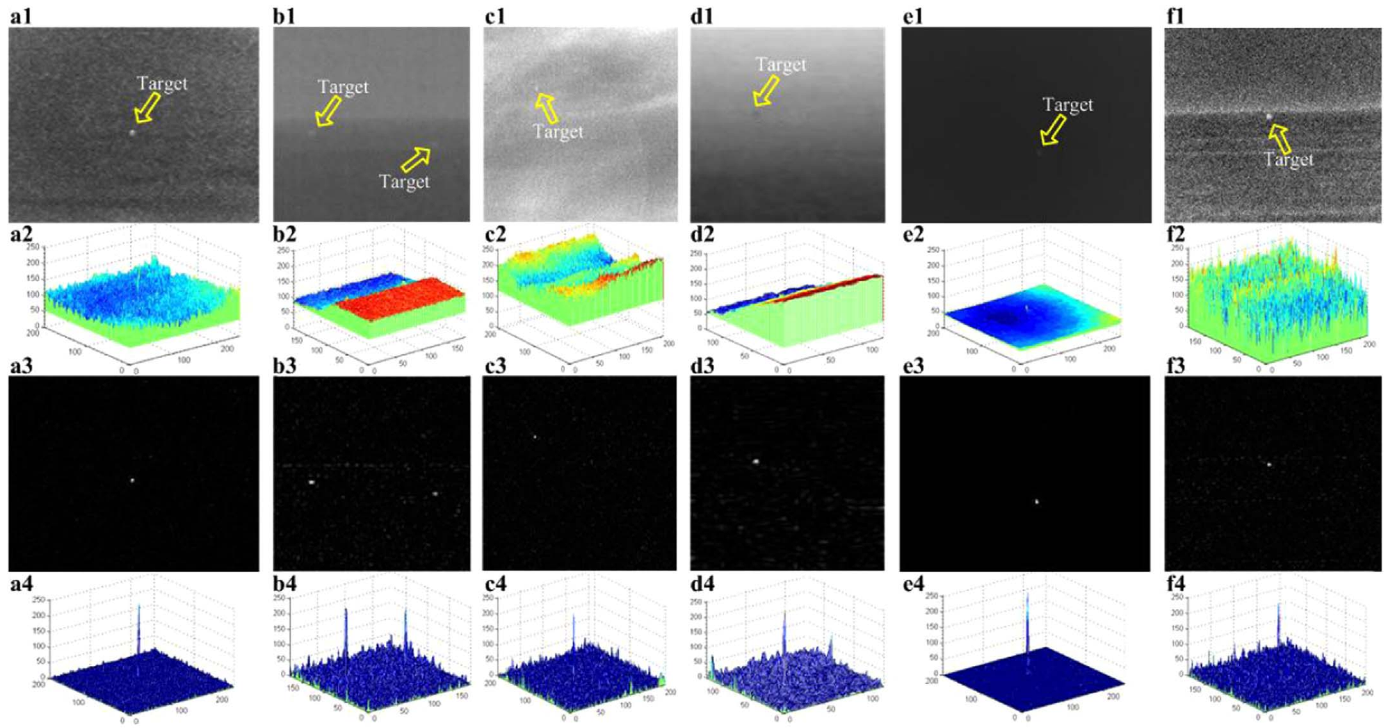

Fig. 4. Dim and small target enhancement: (a1)–(f1) Original images against different complex and noisy backgrounds. (a2)–(f2) 3D gray distribution of (a1)–(f1). (a3)–(f3) Enhanced results of (a1)–(f1) obtained by using our method. (a4)–(f4) 3D gray distributions of (a3)–(f3).

Table 1
Details of six dim and small target infrared image sequences.

	# Frame	Image size	Target shape	Target details	Background details
Sequence 1	120	200 × 256	Circular	1. A long imaging distance. 2. Keeping motionless.	1. Heavy noise. 2. Changing backgrounds.
Sequence 2	110	183 × 185	Rectangular	1. Two targets. 2. Keeping motionless.	1. Heavy noise. 2. Heavy sea-sky background clutters.
Sequence 3	100	208 × 208	Circular	1. Low SCR values. 2. Keeping motionless.	1. Heavy noise.
Sequence 4	100	128 × 128	Rectangular	1. Keeping motion. 2. A changing size within a small range.	1. Heavy noise. 2. Heavy sea-sky background clutters.
Sequence 5	65	200 × 256	Circular or Rectangular	1. Keeping motion. 2. A changing size within a small range.	1. Relatively homogeneous backgrounds.
Sequence 6	130	196 × 218	Circular	1. Low SCR values. 2. Keeping motionless.	1. Heavy noise. 2. Heavy sea-sky background clutters.

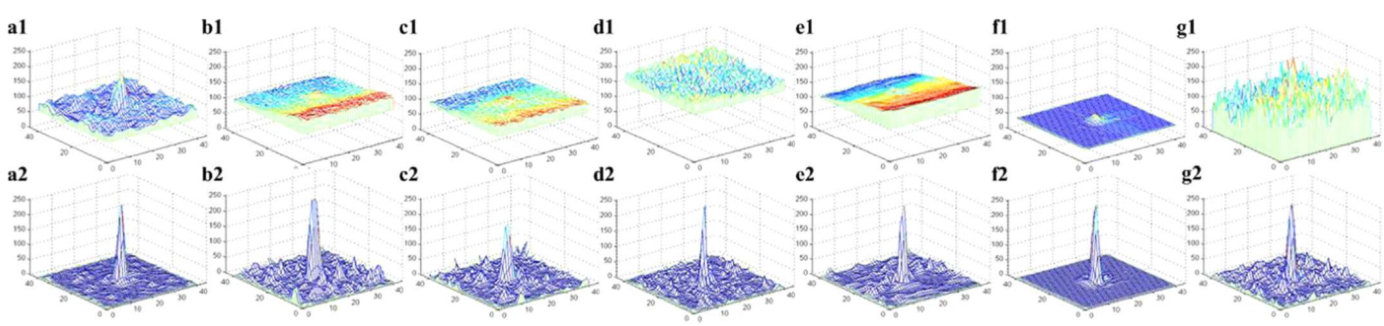

Fig. 5. 3D gray distributions of regions surrounding the target: (a1)–(g1) 3D gray distributions of original regions. (a2)–(g2) 3D gray distributions of enhanced regions obtained by using our method.

baseline methods for comparison, including the LMWIE [26], AGADM [27], and LCM [4]. Moreover, we also choose other detection methods as baseline methods, such as methods based on the Top-hat transform (THT) [2], maximum background prediction model (MBPM) [32], Max-mean filter (MME), or Max-median filter (MED) [13].

Six dim and small target image sequences with low SCR values (more that 600 images) are used to compare the proposed methods with baseline methods, denoted as Sequence 1–6, respectively. The upper row in Fig. 4 is representative images of six sequences against different backgrounds. The details about targets and backgrounds are listed in Table 1.

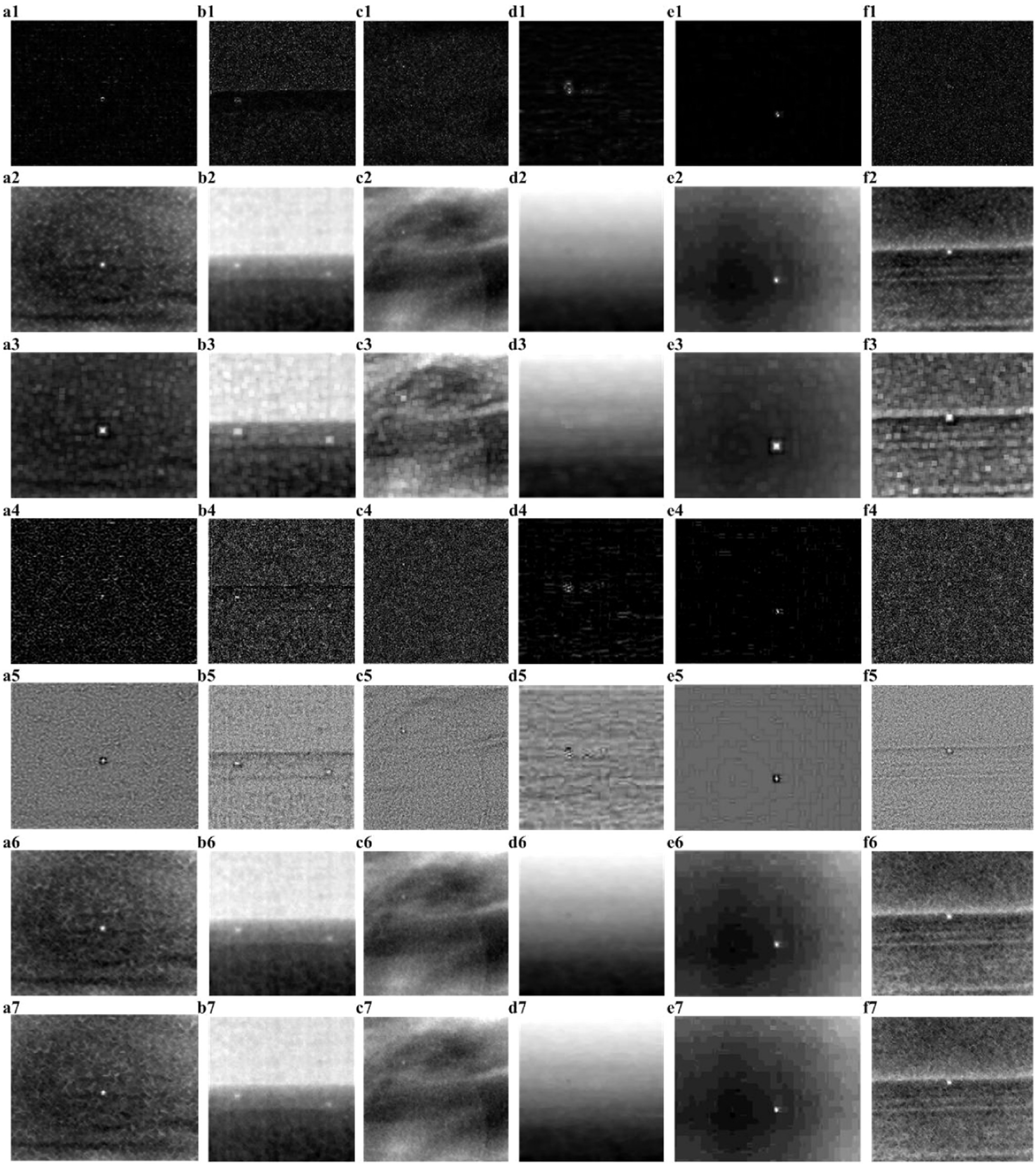

Fig. 6. Enhanced results of Fig. 4(a1)–(f1) obtained by using baseline methods: (a1)–(f1), (a2)–(f2), (a3)–(f3), (a4)–(f4), (a5)–(f5), (a6)–(f6), and (a7)–(f7) Enhanced results obtained by using the LMWIE, AGADM, LCM, THT, MBPM, MME, and MED methods, respectively.

3.2. Target enhancement

If our method can enhance a dim and small target and suppress background clutters and noise better, the target is detected more easily. In order to demonstrate the target enhancement performance of the proposed method, extensive images with different complicated backgrounds are used in this experiment. Some results are shown in Figs. 4 and 5.

Fig. 4(a1)–(f1) show six original images randomly chosen from Sequence 1 to 6, where the targets are labeled by arrows. The corresponding 3D gray distributions are shown in Fig. 4(a2)–(f2), respectively. The enhanced results obtained by using the proposed method are shown in Fig. 4(a3)–(f3), and their 3D gray distributions are shown in Fig. 4(a4)–(f4). The gray value ranges of original

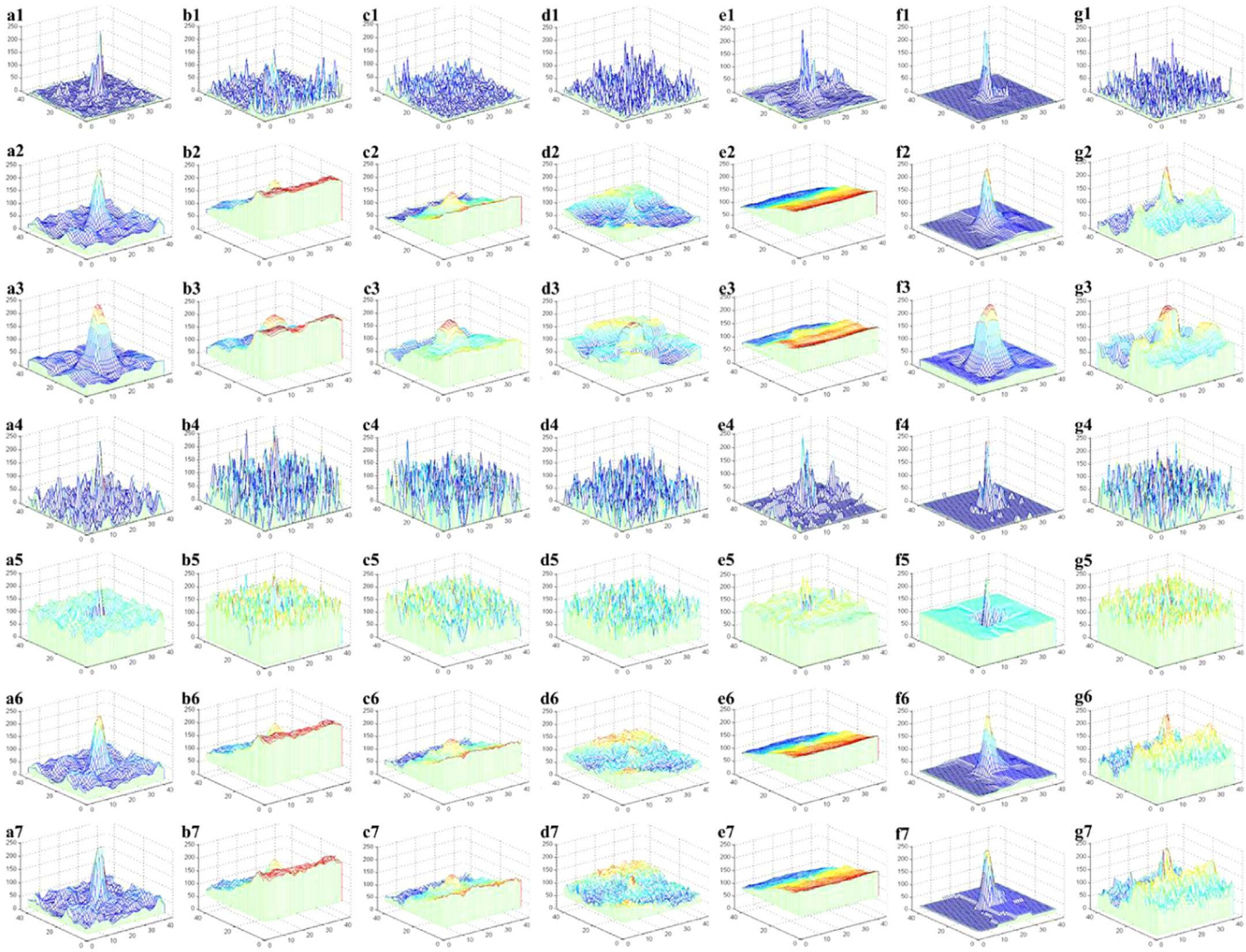

Fig. 7. 3D gray distributions of enhanced regions surrounding the target obtained by using the baseline methods: (a1)–(f1), (a2)–(f2), (a3)–(f3), (a4)–(f4), (a5)–(f5), (a6)–(f6), and (a7)–(f7) 3D gray distributions of enhanced results obtained by using the LMWIE, AGADM, LCM, THT, MBPM, MME, and MED methods, respectively.

Table 2
Comparison results on six sequences by using SCR.

	Original	LMWIE	AGADM	LCM	THT	MBPM	MME	MED	Our method
Sequence 1	5.7613	7.3036	7.6472	5.9712	3.2314	1.1991	6.8118	7.0124	19.7931
Sequence 2 (left)	1.1512	1.5861	1.0569	1.8309	1.0497	1.2324	1.0862	1.0863	10.9545
Sequence 2 (right)	1.8969	1.0048	1.8833	3.0403	0.8331	1.1209	1.8678	1.8969	11.7004
Sequence 3	1.3966	1.4504	2.5949	3.8462	0.6773	0.6233	2.3266	2.0299	7.0448
Sequence 4	0.7013	3.1909	0.2283	0.2636	0.7976	1.7501	0.2812	0.2294	6.4974
Sequence 5	6.8032	8.2707	6.3913	3.4550	6.4692	1.2944	5.7373	6.9622	20.9301
Sequence 6	2.2053	0.1820	0.5462	1.1107	0.1851	0.2934	0.5283	0.5111	11.8493
Average	**2.8451**	**3.2841**	**2.9069**	**2.7883**	**1.8919**	**1.0734**	**2.6627**	**2.8451**	**12.6814**

Table 3
Comparison results of six sequences by using BSF.

	LMWIE	AGADM	LCM	THT	MBPM	MME	MED	Our method
Sequence 1	1.3380	0.4178	0.5790	0.5340	0.6999	0.4203	0.4230	2.8469
Sequence 2	1.1664	0.2273	0.2476	0.4743	0.7049	0.2356	0.2388	2.4746
Sequence 3	1.1357	0.4571	0.4591	0.4741	0.6739	0.4725	0.4740	2.6556
Sequence 4	6.1857	0.6002	0.6244	3.9112	2.7125	0.6029	0.6052	6.4557
Sequence 5	0.9148	0.0729	0.1216	0.4614	0.5096	0.0750	0.0773	1.1950
Sequence 6	1.6249	1.0528	1.1058	0.7386	1.0432	1.0460	1.0435	4.9400
Average	**2.0609**	**0.4714**	**0.5229**	**1.0989**	**1.0573**	**0.4754**	**0.4770**	**3.4280**

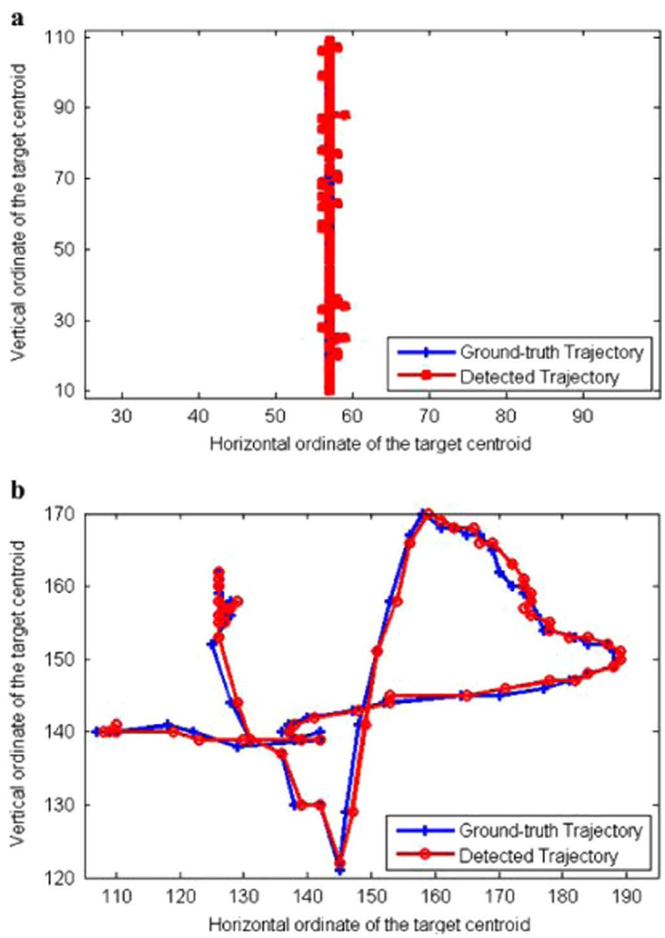

Fig. 8. Ground-truth and detected trajectories obtained by using our method: (a) and (b) Ground-truth and detected trajectories of Sequence 4 and 5.

and result images are normalized to [0,255]. It can be seen that there are heavy background clutters and noise existed in original images. But, after our method, the clutters and noise are well suppressed, and the targets are effectively enhanced. Therefore, the targets in the enhanced results can be easily detected.

The emergence of a dim and small target produces weak mutation of texture characteristics in the whole image plane, but causes great changes of texture features in an area surrounding the target. For Fig. 4(a1)–(f1), 3D gray distributions of areas surrounding the target are shown in Fig. 5(a1)–(g1) (there are two dim and small targets in Fig. 4(b1)). The corresponding 3D gray distributions of enhanced results obtained by using our method are shown in Fig. 5(a2)–(g2). Compared with Fig. 5(a1)–(g1), the local texture mutation in Fig. 5(a2)–(g2) is more significant. This demonstrates that our method can enlarge the difference between the target and surrounding background clutters. Then, the target is well enhanced.

3.3. Enhancement performance

In order to further prove the enhancement performance of our method, seven baseline methods are used in the comparison, including the LMWIE, AGADM, LCM, THT, MBPM, MME, and MED methods. Some comparison results are displayed in Figs. 6 and 7, and Tables 2 and 3.

For Fig. 4(a1)–(f1), Fig. 6 shows the enhanced results obtained through different baseline methods, and Fig. 7 displays the corresponding 3D gray distributions of enhanced regions surrounding the target. From Figs. 4–7, we can find that our method has less clutters and noise residual for different complicated backgrounds, compared with baseline methods. This suggests that the proposed method can maintain lower false alarm rates under the same probability of detection. It can draw a conclusion that our method can efficiently detect targets and perform better than baseline methods.

Both SCR and BSF are appropriate to demonstrate the detection performance of different methods and are utilized here for comparison. For each individual measure, a higher score means the better performance. For the six image sequences, Table 2 lists the average SCR values of the original images and enhanced results obtained by using the LMWIE, AGADM, LCM, THT, MBPM, MME, MED and our method. Table 3 displays the average BSF values obtained by using different methods. Sequence 2 (left) and Sequence 2 (right) in Table 2 denote the respective average SCR values of the left and right targets in Sequence 2. More than 600 images are used to calculate the ensemble average SCR/BSF values, listed in the bottom rows of Tables 2 and 3. Tables 2 and 3 suggest that our method is superior to baseline methods with respect to target enhancement.

The above experimental results demonstrate that the proposed method is high-performance for enhancing dim and small targets and suppressing intricate backgrounds simultaneously. Accordingly, the proposed method is robust and effective to detect dim and small infrared targets submerged in heavy and noisy background clutters.

3.4. Detection performance

The detected result is considered correct if the pixel distance between centers of the ground-truth and the result is less than a threshold (5 pixels [2], 4 pixels [4], and so on). The smaller threshold chosen implies less distance errors and workload in IRST applications. We choose the threshold as 3 pixels in this paper.

The targets in Sequence 4 and 5 keep motion in each frame, while the targets in Sequence 1, 2, 3, and 6 have little motion (see Table 1). For Sequence 4 and 5, Fig. 8 displays the ground-truth target movement trajectories and detected trajectories obtained by using our method. The corresponding histograms of horizontal and vertical error distributions are shown in Fig. 9. For Sequence 4, the ground-truth trajectory is linear, and the detected trajectory does almost match that of the target movement, as shown in Fig. 8 (a). The most horizontal and vertical errors are zeros (a small number of horizontal errors are less than 2 pixels, and a small number of vertical errors are 1 pixel), as shown in Fig. 9(a) and (b). For Sequence 5, the ground-truth trajectory is curvilinear, but the detected trajectory does almost match that of the target movement as well (see Fig. 8(b)). The horizontal and vertical errors are shown in Fig. 9(c) and (d), which suggests that the most both horizontal and vertical errors are less than 1 pixel (a small number of horizontal errors are 2 pixels). It draws a conclusion from Figs. 8 and 9 that the proposed method can achieve high probabilities of detection as well as low false alarm rates for different target movements, SCR values and background clutters.

The receiver operating characteristic (ROC) curve is a graphical plot of the probabilities of detection (a fraction of true positives over the positives) versus the false alarm rates (a fraction of false positives over the negatives). We provide ROC curves obtained by using the baseline methods and proposed method for the six image sequences in Fig. 10. It suggests that our method has better detection performance than baseline methods. Especially for Sequence 1, 2, 4, and 5, our method owns the highest probabilities of detection but the lowest false-alarm rates in the comparison. For Sequence 3 or 6, the LMWIE method has a little better detection performance than our method when $Fa \le 0.6$ or $Fa \le 0.5$, but our method can reach 1 faster when $Fa > 0.7$ or $Fa > 0.5$. For Sequence

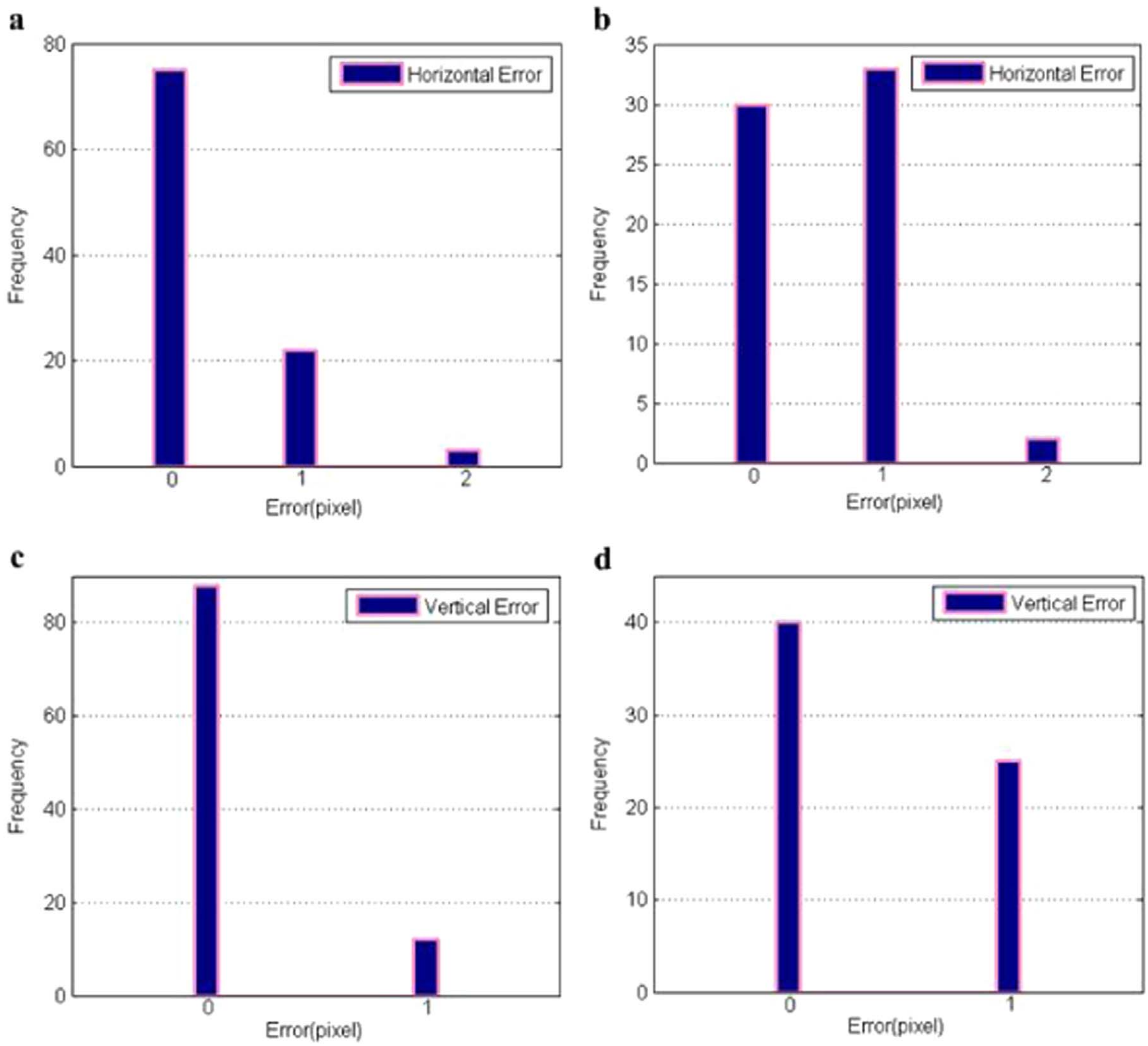

Fig. 9. Histograms of detected errors obtained by using our method: (a) and (b) Histograms of horizontal detected errors of Sequence 4 and 5. (c) and (d) Histograms of vertical detected errors of Sequence 4 and 5.

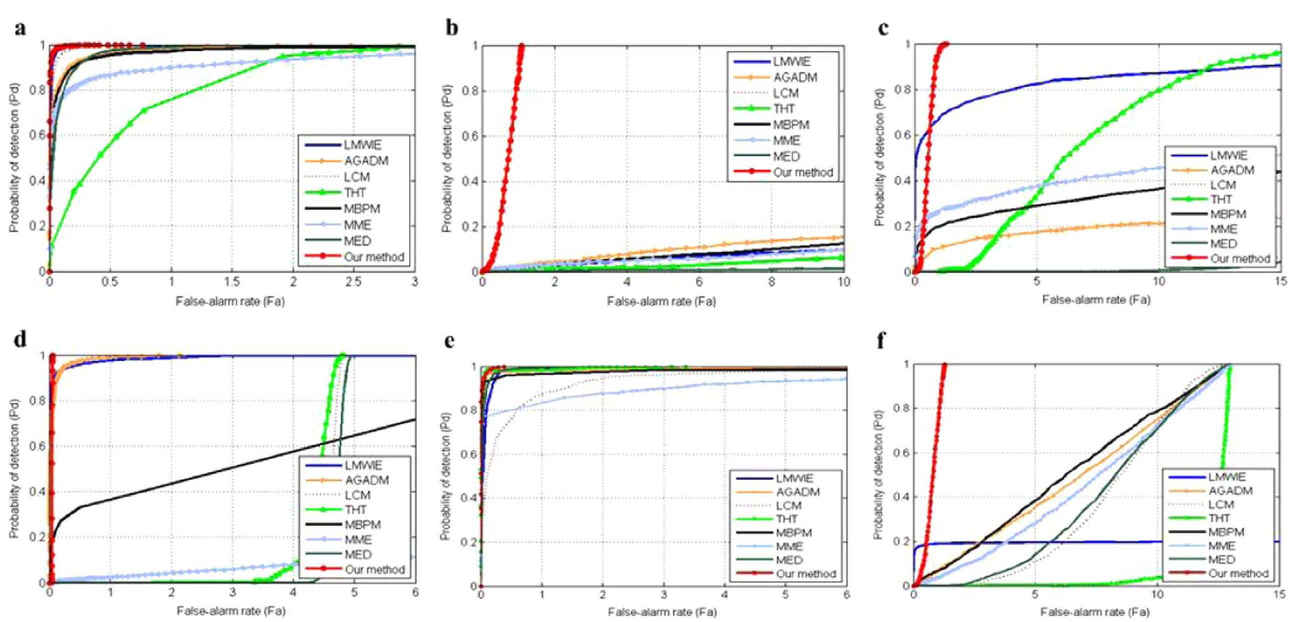

Fig. 10. ROC curves: (a)–(f) ROC curves of Sequence 1–6 obtained by using the baseline methods and our method.

Table 4
Comparison of AUC values based on different detection methods.

	LMWIE	AGADM	LCM	THT	MBPM	MME	MED	Our method
Sequence 1	0.9946	0.9717	0.9932	0.7847	0.9641	0.9035	0.9702	0.9979
Sequence 2	0.0965	0.1069	0.0094	0.0358	0.0914	0.0858	0.0098	0.9296
Sequence 3	0.8299	0.1945	0.0418	0.5334	0.3325	0.4144	0.0178	0.9617
Sequence 4	0.9862	0.9882	0.2292	0.2680	0.8504	0.0714	0.2136	0.9922
Sequence 5	0.9871	0.9868	0.9200	0.9920	0.9830	0.8870	0.9925	0.9982
Sequence 6	0.2196	0.5487	0.4455	0.1796	0.5698	0.5093	0.4564	0.9466
Average	**0.6857**	**0.6328**	**0.4398**	**0.4656**	**0.6319**	**0.4786**	**0.4434**	**0.9710**

3 and 5, the LMWIE method has better performance than the AGADM, LCM, MBPM, MME, and MED methods. Fig. 10 also shows that the LCM method is superior to other baseline methods for Sequence 1, but has low probabilities of detection for other image sequences. The comparison results derived from Fig. 10 suggests that the proposed method is suitable for detecting dim and small target against diverse complex and noisy backgrounds.

The area under the ROC curve (AUC) is widely utilized to quantitatively evaluate the classification performance of true or false targets [33,34]. The AUC value is within the range [0,1]. A higher AUC value obtained by using a method is, the better performance with respect to detection accuracy is. For the above six image sequences, Table 4 lists the AUC values obtained by using the baseline methods and the proposed method. The average AUC values acquired by using the eight methods are listed in the bottom rows of Table 4. It can be seen that the proposed method is superior to the baseline methods in the AUC comparison. From Figs. 4–10 and Tables 2–4, it can be found that the proposed method achieves the best performance, which implies that our method works more robustly for different target movements and backgrounds with low SCR values.

4. Conclusion

This paper presents an effective method based on the LDM to detect dim and small infrared targets embedded in different background clutters. The key idea of the presented method is to adopt an adaptive entropy-based window selection technique to construct a LDM map of an input image, which effectively enhances targets and suppresses background clutters and noise simultaneously. In this way, the LDM map can significantly improve SCR values of the image, and have little clutters and noise residual. This ensures the presented method arouses low false alarm rates under the same probability of detection. The experiments have been implemented on more than 600 dim and small target images against diverse complicated backgrounds, which demonstrate that the presented method outperforms conventional baseline methods, such as the LMWIE, AGADM, LCM, THT, MBPM, MME, and MED methods. The experimental results also justify that the presented method works more robustly for different target movements and complex and noisy backgrounds.

Although the experimental results justify the robustness of the presented method and provide empirical evidence, we can even improve it further more from different perspectives in the future work. For example, we will investigate a faster version of the current algorithm, as well as extend the definition of the local difference measure by applying different operations.

Acknowledgment

This work was supported by the Natural Science Foundation of China (61471355, 81227902).

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. 45 (1) (2012) 393–406.
[2] X.Z. Bai, F.G. Zhou, Analysis of new top-hat transformation and the application for infrared dim small target detection, Pattern Recognit. 43 (6) (2010) 2145–2156.
[3] C.Q. Gao, D.Y. Meng, Y. Yang, Y.T. Wang, X.F. Zhou, A.G. Hauptmann, Infrared patch-image model for small target detection in a single image, IEEE Trans. Image Process. 22 (12) (2013) 4996–5009.
[4] C.L. Philip, H. Li, Y.T. Wei, T. Xia, Y.Y. Tang, A local contrast method for small infrared target detection, IEEE Trans. Geosci. Remote Sens. 51 (1) (2014) 574–581.
[5] M. Malanowski, K. Kulpa, Detection of moving targets with continuous-wave noise radar: theory and measurements, IEEE Trans. Geosci. Remote Sens. 50 (9) (2012) 3502–3509.
[6] H. Deng, Y.T. Wei, M.W. Tong, Small target detection based on weighted self-information map, Infrared Phys. Technol. 60 (9) (2013) 197–206.
[7] N. Thanh, H. Sahli, D. Hao, Infrared thermography for buried landmine detection: inverse problem setting, IEEE Trans. Geosci. Remote Sens. 46 (12) (2008) 3987–4504.
[8] B. Porat, B. Friendlander, A frequency domain algorithm for multi-frame detection and estimation of dim targets, IEEE Trans. Pattern Anal. Mach. Intell. 12 (4) (1990) 398–401.
[9] C.Q. Gao, T.Q. Zhang, Q. Li, Small infrared target detection using sparse ring representation, IEEE Trans. Aerosp. Electron. Syst. 27 (3) (2012) 21–30.
[10] H. Deng, J.G. Liu, H. Li, EMD based infrared image target detection method, J. Infrared Millim. Terahertz Waves 30 (11) (2009) 1205–1215.
[11] P.L. Shui, D.C. Li, S.W. Xu, Tri-feature-based detection of floating small targets in sea clutter, IEEE Trans. Aerosp. Electron. Syst. 50 (2) (2014) 1416–1630.
[12] Y.F. Gu, C. Wang, B.X. Liu, Y. Zhang, A kernel-based nonparametric regression method for clutter removal in infrared small-target detection application, IEEE Geosci. Remote Sens. Lett. 7 (3) (2010) 469–473.
[13] S. Deshpande, M. Er, R. Venkateswarlu, Max-mean and max-median filters for detection of small-targets, Proc. SPIE 3809 (1999) 74–83.
[14] S.C. Pohlig, Spatial-temporal detection of electro-optic moving targets, IEEE Trans. Aerosp. Electron. Syst. 31 (2) (1995) 608–616.
[15] L. Yang, J. Yang, K. Yang, Adaptive detection for infrared small target under sea-sky complex background, Electron. Lett. 40 (17) (2004) 1083–1085.
[16] R.N. Strickland, H.I. Hahn, Wavelet transform methods for objects detection and recovery, IEEE Trans. Image Process. 6 (5) (1997) 724–735.
[17] S.P. Zhang, H.X. Yao, S.H. Liu, Dynamic background subtraction based on local dependency histogram, Int. J. Pattern Recognit. Artif. Intell. 23 (7) (2009) 1397–1419.
[18] S.P. Zhang, S. Kasiviswanathan, P.C. Yuen, M. Harandi, Online dictionary learning on symmetric positive definite manifolds with vision applications, in: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, 2015, pp. 3165–3173.
[19] S. Kim, J. Lee, Small Infrared target detection by region-adaptive clutter rejection for sea-based infrared search and track, Sensors 14 (7) (2014) 13210–13242.
[20] J.F. Khan, M.S. Alam, Target detection in cluttered forward-looking infrared imagery, Opt. Eng. 44 (7) (2005), 076404-1-076404-8.
[21] M. Shirvaikar, M. Trivedi, A neural network filter to detection small targets in high background clutters, IEEE Trans. Neural Netw. 6 (1) (1995) 252–257.
[22] H. Li, Y.T. Wei, L. Li, Y.Y. Tang, Infrared moving target detection and tracking based on tensor locality preserving projection, Infrared Phys. Technol. 53 (2) (2010) 77–83.
[23] D. Song, D. Tao, Biologically inspired feature manifold for scene classification, IEEE Trans. Image Process. 19 (1) (2010) 174–184.
[24] Y. Gao, R. Liu, J. Yang, Small target detection using two-dimensional least mean square (TDLMS) filter based on neighborhood analysis, J. Infrared Millim. Terahertz Waves 29 (2) (2008) 188–200.
[25] H. Deng, J.G. Liu, Z. Chen, Infrared small target detection based on modified local entropy and EMD, Chin. Opt. Lett. 8 (1) (2010) 24–28.
[26] X.J. Qu, H. Chen, G.H. Peng, Novel detection method for infrared small targets using weighted information entropy, J. Syst. Eng. Electron. 23 (6) (2012) 838–842.
[27] G.Y. Wang, T.X. Zhang, L.G. Wei, N. Sang, Efficient method for multiscale small target detection from a natural scene, Opt. Eng. 35 (3) (1996) 761–768.

[28] N.R. Pal, S.K. Pal, Entropy: a new definition and its application, IEEE Trans. Syst. Man Cybern. 21 (5) (1991) 1260–1270.
[29] B.N. Subudhi, P.K. Nanda, A. Ghosh, Entropy based region selection for moving object detection, Pattern Recognit. Lett. 32 (15) (2011) 2097–2108.
[30] W. Zhang, M. Cong, L. Wang, Algorithm for optical weak small targets detection and tracking: review, in: Proceedings of IEEE International Conference on Neural Network Signal Processing, 2003, pp. 643–647.
[31] J.F. Rivest, R. Fortin, Detection of dim targets in digital infrared imagery by morphological image processing, Opt. Eng. 35 (7) (1996) 1886–1893.
[32] K. Huang, X. Mao, Detectability of infrared small targets, Infrared Phys. Technol. 53 (3) (2010) 208–217.
[33] S. Kim, K.T. Kim, S. Kim, Infrared small target discrimination using sequential forward feature selection with AUC metric, Proc. Ind. Technol. (2014) 641–644.
[34] T. Fawcett, An introduction to ROC analysis, Pattern Recognit. Lett. 27 (8) (2006) 861–874.

He Deng was born in 1977. He received his Ph.D. degree in Control Science and Engineering from Huazhong University of Science and Technology in 2011. Now, he is an associate professor in Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences. His current research interests are target detection and image processing.E-mail: ml.liu@wipm.ac.cn

Xianping Sun received his B.S. degree from Peking University. Now, he is a professor in Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences. His current research interests include signal acquisition and processing.E-mail: xpsun@wipm.ac.cn

Maili Liu received his Ph.D. degree from University of London. Now. he is a professor in Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences. His current research interests include NMR spectroscopy, imaging acquisition and applications.E-mail: ml.liu@wipm.ac.cn

Chaohui Ye received his B.S. degree from Peking University. Now, he is a professor in Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences. His current research interests include novel imaging techniques and applications in biomedicine and physics.E-mail: ye@wipm.ac.cn

Xin Zhou received his Ph.D. degree from the Chinese Academy of Sciences. Now, he is a professor in Wuhan Institute of Physics and Mathematics, Chinese Academy of Sciences. His current research interests are scientific developments, imaging acquisition and data processing.E-mail: xinzhou@wipm.ac.cn