Infrared LSS-Target Detection Via Adaptive TCAIE-LGM Smoothing and Pixel-Based Background Subtraction

Yanfeng WU[1,2], Yanjie WANG[1], Peixun LIU[1], Huiyuan LUO[1,2], Boyang CHENG[1,2], and Haijiang SUN[1*]

[1]*Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China*

[2]*University of Chinese Academy of Sciences, Beijing 100049, China*

*Corresponding author: Haijiang SUN   E-mail: sunhaijiang@126.com

Abstract: Infrared small target detection is a significant and challenging topic for daily security. This paper proposes a novel model to detect LSS-target (low altitude, slow speed, and small target) under the complicated background. Firstly, the fundamental constituents of an infrared image including the complexity and entropy are calculated, which are invoked as adaptive control parameters of smoothness. Secondly, the adaptive L0 gradient minimization smoothing based on texture complexity and information entropy (TCAIE-LGM) is proposed in order to remove noises and suppress low-amplitude details in infrared image abstraction. Finally, difference of Gaussian (DoG) map is incorporated into the pixel-based adaptive segmentation (PBAS) background modeling algorithm, which can differ LSS-target from the sophisticated background. Experimental results demonstrate that the proposed novel model has a high detection rate and produces fewer false alarms, which outperforms most state-of-the-art methods.

Keywords: Small target detection; L0 smoothing; texture complexity; information entropy; pixel-based adaptive segmentation

1. Introduction

Owing to the dominant position of the thermal infrared imaging system to run in dark and low light environments, infrared cameras have gained popularity for missile guidance, military night vision, airborne early warning, etc. Accordingly, LSS-targets (low altitude, slow speed, and small target) can be captured due to their hot temperature. Nevertheless, infrared images are frequently of poor quality, as a result of salt-and-pepper noises, less texture, and non-uniformity noise, which render it rather difficult to detect LSS-targets. To make things more intricate, LSS-targets without fixed moving trajectories are often submerged in heavy noises or complex backgrounds. Thus, it is a concerned and challenging topic in the infrared detection field.

In order to detect such LSS-targets accurately, scholars in various countries have done a lot of researches and put forward diverse algorithms. Hu *et al.* [1] used the non-local mean filter based on circular mask to establish a background estimation model. By linking the gray scale distribution of the image to the temporal information, infrared dim targets can be extracted successfully. Yang *et al.* [2] simplified a two-dimensional median filter to a

one-dimensional median filter, which could eliminate the background and false targets. He *et al.* [3] realized small target detection by local entropy (LE) and one-dimensional empirical mode decomposition (EMD). Abdelkawy *et al.* [4] used two-dimensional Gauss function to construct a dictionary and proposed a two-dimensional orthogonal search (2D-FOS) algorithm. The time complexity of the algorithm is mainly dependent on the number of candidates and the size of the image in the dictionary. Compared with other orthogonal methods, the computation time is significantly reduced. An infrared small target detection algorithm based on the peer group filter (PGF), two-dimensional empirical mode decomposition (TDEMD), and local inverse entropy (LE) was proposed by Xie *et al.* [5]. Huang *et al.* [6] improved a dynamic programming track-before-detect algorithm (DP-TBD), which can simultaneously detect and track maneuvering dim targets. Gao *et al.* [7] built a co-detection model based on nonlinear weight and entry-wise weighted robust principal component analysis (RPCA), which can extract real targets accurately and suppress background clutters efficiently. Li *et al.* [8] proposed an infrared dim target detection approach based on sparse representation on a discriminative over-complete dictionary, which can not only capture significant features of background clutters and dim targets, but also strengthen the sparse feature difference between the background and target. Kim *et al.* [9] analyzed the characteristics of regional cluster and removed the false detection by means of spatial attribute-based classifications, the heterogeneous background removal filter, and temporal consistency filter. Motivated by the background classification and coastal region detection, Kim *et al.* [10] proposed a novel scene-dependent small target detection strategy involving the relationships between the geometric horizon and the image horizon. By classifying the infrared background types and detecting the littoral regions in omni-directional images, coastal regions can be detected by fusing the region map and curve map.

As discussed above, these methods can be classified into two categories: one is for the single-frame image target detection, which uses a filtering algorithm to eliminate complex backgrounds and estimate the foreground target. Morphology top-hat transform [11], high-pass filter [12], and two-dimensional entropy [13] can achieve real-time detection of targets, but the accuracy is not high under low signal to noise ratio (SNR) conditions. The detection accuracies of the matched filter [14], the wavelet transform [15], partial differential equations (PDE) [16], and the probabilistic principal component analysis matrix [17] (PCA) are high, but it is almost impossible to achieve real-time detection. Other methods such as particle filtering [18], mobile weighted pipeline filtering [19, 20], and likelihood ratio test [21] are based on the multi-frame image, which need to achieve target detection through inter-frame context information. This kind of algorithms has superior accuracy, but aims can't be found effectively if the targets are submerged in the backgrounds or noises [22].

Statistical background modeling is a fundamental and important part of many visual searching systems and other computer vision applications. This article focuses on how to remove noise and smooth the background for infrared images and then detect the real target according to the inter-frame difference from static complex background. Firstly, this paper analyzes the characteristics of infrared images. Secondly, a novel adaptive texture complexity and information entropy (TCAIE-LGM) smoothing filter is proposed, which can remove stripe noises and salt-and-pepper noises of infrared images. Finally, a pixel-based background subtraction is introduced to remove the complex background and detect the real target.

2. IR noise suppression

Because of the abnormity of material's internal structure and crystal defects, non-uniformity is a natural defect in the infrared imaging system. What's more, the LSS-target detection is ordinarily

applied under the background of the sky or sky-earth junction in many cases. The infrared image shows obviously inhomogeneous character due to low temperature radiation from the sky. Therefore, noise suppression must be done before using infrared cameras. This section reviews the related papers in terms of noise suppression and proposes the TCAIE-LGM algorithm, which can adaptively suppress the low frequency details and noises of the IR image.

2.1 Related work in terms of noise suppression

Scribner *et al.* [23, 24] proposed two adaptive non-uniformity correction algorithms based on the human visual system. One is the time domain high pass filter, which refers to the low-pass filtering characteristics of human eye horizontal cells in optical signals and constructs time-domain high pass filter to correct the offset coefficient. This method is simple and easy to implement, but leads to target degradation and ghost phenomenon in stationary scene. The other algorithm is the representation of space operation, which takes advantages of the neural network structure and the steepest descent method to adaptively suppress the stripe noise. The algorithm has a good effect on the spatial high frequency noise suppression, but the rate of convergence is slow. Harris and Chiang studied the non-uniformity correction algorithm with constant parameter statistics, in which thousands of images were needed for training [25]. Qin *et al.* [26] utilized wavelet transform in terms of clutter rejection, but the speed of convergence was slow as well.

2.2 L0 gradient minimization

Image smoothing is an important instrument for computing photography. Its function is to eliminate unimportant details and retain larger image edges. Xu *et al.* [27] proposed the L0 gradient minimization (LGM), which was a global smoothing filter based on the sparse strategy. LGM suppresses low-amplitude details and sharpens salient edges, which can remove noise, unimportant details, and make the results immediately usable in background

subtraction. Distinct from other filters, this method can faithfully maintain small-resolution objects and thin edges.

The LGM enhances the highest contrast edges and removes small amplitude gradients globally by confining the number of non-zero gradients. In the 2D image, what needs to be done is to constrain the number of gradients in the horizontal and vertical directions. The function representation is as

$$\min_{S}\left\{\sum_{p}\left(S_p - I_p\right)^2 + \lambda C(S)\right\} \quad (1)$$

$$C(S) = \#\left\{p \,\middle|\, \left|\partial_x S_p\right| + \left|\partial_y S_p\right| \neq 0\right\}, p \neq 0 \quad (2)$$

where I is the input image, and S is the computed result. $\nabla S = (\partial_x S_p, \partial_y S_p)^{\mathrm{T}}$ is the gradient of the image, which is calculated between neighboring pixels along horizontal and vertical directions. p is the counts of pixels whose magnitude is not zero.

2.3 Adaptive TCAIE-LGM smoothing

Although LGM can remove the noise and eliminate details, the degree of smoothness can't be controlled effectively. Adaptive L0 gradient minimization smoothing based on texture complexity and information entropy (TCAIE-LGM) is purposed in this paper in order to remove noise and suppress low-amplitude details in the infrared image abstraction.

2.3.1 Information entropy of image

Entropy was first proposed by Shannon [28] and applied in thermodynamics, which refers to the degree of chaos in the system. It has significant applications in the fields of cybernetics, probability theory, number theory, astrophysics, life science, and so on. It is a very critical parameter in various fields and has a more specific definition as well. Information entropy is a random measure of information contained in an image. It is a statistical form of characteristics, which reflects the average amount of information in an image.

The one-dimensional information entropy of the image represents the information contained in the

gray scale distribution of the image, which represents the aggregate characteristics of the image. The entropy of a gray image is defined as

$$ENT_1 = -\sum_{i=0}^{255} P_i \log P_i \tag{3}$$

where P_i indicates the proportion of pixels in the image with a gray value of i, which can be obtained from the histogram.

The one-dimensional information entropy of images can represent the aggregate characteristics of the gray scale distribution in the image, but can't reflect the spatial characteristics of the image. In order to demonstrate the spatial feature, the two-dimensional entropy of the image is composed on the basis of one-dimensional entropy, which can reflect the features of the spatial distribution of the gray level. The two-dimensional information entropy can address the complexity and the inhomogeneity of the image. When the image is a pure color graph, there is only one gray value. At this time, the two-dimensional information entropy of the image is the minimum, and the amount of information about the image is zero. If the gray value of each pixel is different, the two-dimensional information entropy of the image is the biggest, and the information of the image is the largest. For two-dimensional digital images in discrete form, the two-dimensional information entropy formula is as

$$P(i,j) = \frac{f(i,j)}{M^2} \tag{4}$$

$$ENT_2 = -\sum_{i=0}^{255} \sum_{j=0}^{255} P(i,j) \log\left[P(i,j) \right] \tag{5}$$

where i represents the gray value of current pixels, and j represents the gray value of current pixels in the neighborhood. The mean value of neighborhood in the image is chosen as the spatial characteristic of the gray distribution and forms a two-element feature group with the pixel gray level of the image, which is recorded as $f(i, j)$. M is the scale of the image, P is the frequency of appearance of each gray level, and $P(i, j)$ is the frequency of appearance of the two-element feature group (i, j). The two-dimensional entropy can reflect the position information of the gray pixel in the image and the comprehensive characteristics of the gray value distribution in the neighborhood under the premise that the amount of image information is not zero. If the probability distribution of pixel value is balanced, the outline of the object is clear, and the entropy value is large. Otherwise, if the elements of the co-occurrence matrix are different, the entropy value is small.

2.3.2 Texture complexity

Texture is a kind of visual feature reflecting the homogeneous phenomenon in the image, which embodies the structure and arrangement properties of surface structure with gradual change or periodic change. Texture feature is a global feature, which describes the surface properties of the scene corresponding to the image or image area. As a statistical feature, texture features have rotation invariance and strong resistance to noise. The angular second moment is a measure of the gray scale change of the image texture, which reflects the uniformity of the gray scale distribution and the thickness of the texture. Therefore, this paper uses the angular second moment of the image histogram to calculate the texture complexity of the image.

Firstly, a normalized one-dimensional histogram of infrared images is necessary. The histogram of the image whose gray level is $[0, L–1]$ is a discrete function. The following statistics are made on the image as

$$h(z_i) = n_i, \ i = 0, \ 1, \ \cdots, \ L-1 \tag{6}$$

where z_i is the gray value of level i, and n_i is the number of pixels z_i in the image. The normalized histogram is calculated as

$$p(z_i) = \frac{n_i}{n} \tag{7}$$

where n is the total number of pixels in the image, and $p(z_i)$ is the probability value of the gray level z_i.

The mean value of the gray level can be calculated as follows:

$$mean = \sum_{0}^{L-1} z_i p(z_i). \tag{8}$$

Yanfeng WU *et al.*: Infrared LSS-Target Detection Via Adaptive TCAIE-LGM Smoothing and Pixel-Based Background Subtraction

183

The angular second moment of the histogram is as follows:

$$\mu(z) = \sum_{0}^{l-1} (z_i - \text{mean})^2 p(z_i). \qquad (9)$$

2.3.3 Improved methods

In our purposed method, information entropy is used to evaluate the energy distribution of the image, and the angular second moment is calculated to evaluate the texture complexity of the image. These two parameters are introduced into (1), which replace the regularization parameter λ in the original equation and adaptively smooth the IR image according to the complexity of the image.

$$\min_{S} \left\{ \sum_{p} (S_p - I_p)^2 + (\kappa_1 + \kappa_2) C(S) \right\} \qquad (10)$$

$$C(S) = \# \left\{ p \big| |\partial_x S_p| + |\partial_y S_p| \neq 0 \right\}, p \neq 0 \qquad (11)$$

$$\kappa_1 = \frac{1}{\phi} ENT_2 = -\frac{1}{\phi} \sum_{i=0}^{255} \sum_{j=0}^{255} p(i,j) \log[P(i,j)] \qquad (12)$$

$$\kappa_2 = \frac{1}{\varphi} \mu(z) = \frac{1}{\varphi} \sum_{0}^{l-1} (z_i - m)^2 p(z_i) \qquad (13)$$

where κ_1 and κ_2 are adaptive variables depending on the information entropy and the angular second moment and φ and φ are empirical parameters.

As the L_0-norm is not guided, the global optimal problem is a non-deterministic polynomial (NP) hard problem. So the variable splitting variable is applied here as a result turning the problem into the quadratic programming problem, each of which has its closed-form solutions as

$$\min_{S,h,v} \left\{ \sum_{p} (S_p - I_p)^2 + (\kappa_1 + \kappa_2) C(h,v) + \right.$$

$$\left. \beta \Big[(\partial_x S_p - h_p)^2 + (\partial_y S_p - v_p)^2 \Big] \right\} \qquad (14)$$

$$C(h,v) = \# \left\{ p \big| |h_p| + |v_p| \neq 0 \right\}, p \neq 0 \qquad (15)$$

where auxiliary variables, h_p and v_p, correspond to $\partial_x S_p$ and $\partial_y S_p$, respectively, and $C(h, v)$ is an automatically adapting parameter to control the similarity between variables (h, v) and their corresponding gradients. The function above is solved through alternatively minimizing. The S

estimation sub-problem corresponds to

$$E(S) = \sum_{p} (S_p - I_p)^2 + \beta \Big[(\partial_x S_p - h_p)^2 + (\partial_y S_p - v_p)^2 \Big]. \qquad (16)$$

Fixing h and v and solving the upper formula, S is calculated as

$$S = F^{-1} \left\{ \frac{F(I) + \beta \big[F(\partial_x) * F(h) + F(\partial_y) * F(v) \big]}{F(1) + \beta \big[F(\partial_x) * F(\partial_x) + F(\partial_y) * F(\partial_y) \big]} \right\} \qquad (17)$$

where F is an FFT operator, and $*$ denotes the complex conjugate. Fixing S and h, v is calculated as follows:

$$\min_{h,v} \left\{ \sum_{p} (\partial_x S_p - h_p)^2 + (\partial_y S_p - v_p)^2 + \right.$$

$$\left. \frac{(\kappa_1 + \kappa_2)}{\beta} C(h,v) \right\}. \qquad (18)$$

Equation (18) is accordingly decomposed as

$$\sum_{p} \min_{h_p, v_p} \left\{ (h_p - \partial_x S_p)^2 + (v_p - \partial_y S_p)^2 + \right.$$

$$\left. \frac{(\kappa_1 + \kappa_2)}{\beta} H(|h_p| + |v_p|) \right\} \qquad (19)$$

where $H(|h_p| + |v_p|)$ is a binary function:

$$H(|h_p| + |v_p|) = \begin{cases} 1, & |h_p| + |v_p| \neq 0 \\ 0, & \text{others} \end{cases}. \qquad (20)$$

Each single pixel in (19) is as follows:

$$E_p = \left\{ (h_p - \partial_x S_p)^2 + (v_p - \partial_y S_p)^2 + \right.$$

$$\left. \frac{(\kappa_1 + \kappa_2)}{\beta} H(|h_p| + |v_p|) \right\} \qquad (21)$$

when the upper is minimization, (22) is obtained as follows:

$$(h_p, v_p) = \begin{cases} (0, 0), & (\partial_x S_p)^2 + (\partial_x S_p)^2 \leq \dfrac{(\kappa_1 + \kappa_2)}{\beta} \\ (\partial_x S_p, \partial_y S_p), & \text{otherwise} \end{cases}. \qquad (22)$$

The preprocessing method can remove the small non-zero gradient, smooth the unimportant details, and enhance the saliency edge of the image, which plays an important role in the goal of the pixel-based

background subtraction (PBAS) algorithm. The preprocessing method can also effectively eliminate the texture information which may cause interference in the process of background modeling and highlight the boundary features of the target. This makes the detection result clearer and reduces the noise and background interference effectively.

3. LSS-target detection using PBAS

PBAS is to automatically generate a binary mask which divides a set of pixels into the set of foreground and a set of background pixels [29]. This paper presents an improved nonparametric method aiming at IR images, which can overcome the influence of background changes and smear phenomenon.

In our case, the front N frame pixels x_i are collected as background models $B(x_i)$.

$$B(x_i) = \{B_1(x_i), B_2(x_i), \cdots, B_k(x_i), \cdots, B_N(x_i)\}. \quad (23)$$

To calculate the foreground segmentation mask, the discriminant formula is expressed as follows:

$$F(x_i) = \begin{cases} 1, & \#\{\mathrm{dist}[I(x_i), B_k(x_i)] < R(x_i)\} < \#\min \\ 0, & \text{else.} \end{cases} \quad (24)$$

If the pixel x_i satisfying the formula is less than the threshold #min, x_i is decided to be the foreground. Otherwise, pixel x_i is defined as the background. $R(x_i)$ is represented as the judgment threshold.

The gradient magnitude at the pixel $I^m(x_i)$ is calculated as

$$I^m(x_i) = DoG * K(x, y) \quad (25)$$

$$DoG \overset{\Delta}{=} G_{\zeta_1} - G_{\zeta_2}$$
$$= \frac{1}{\sqrt{2\pi}}\left(\frac{1}{\zeta_1}e^{-(x^2+y^2)/2\zeta_1^2} - \frac{1}{\zeta_2}e^{-(x^2+y^2)/2\zeta_2^2}\right) \quad (26)$$

where G_{ζ_1} and G_{ζ_2} are Gauss functions, and ζ_1 and ζ_2 are the scale parameters of Gauss function. $K(x, y)$ is the average of difference of Gaussian gradient magnitude over the last observed frame. Correspondingly, $B_k^m(x_i)$ is the gradient magnitude

of the background history. For the distance calculation, we use the following equation:

$$\mathrm{dist}[I(x_i), B_k(x_i)]$$
$$= \frac{\alpha}{\overline{I_m}} \cdot |I^m(x_i) - B_k^m(x_i)| + |I^v(x_i) - B_k^v(x_i)| \quad (27)$$

where $I^v(x_i)$ is the pixel value, $B_k^v(x_i)$ is the pixel value of the background history, and $I^v(x_i)$ is the average gradient magnitude over the last observed frame. The currently observed minimal distance is expressed as follows:

$$d_{\min}(x_i) = \min_k \mathrm{dist}[I(x_i), B_k(x_i)]. \quad (28)$$

An array of minimal decision distances is created as follows:

$$D(x_i) = \{D_1(x_i), D_2(x_i), \cdots, D_k(x_i), \cdots, D_N(x_i)\}. \quad (29)$$

Whenever an update of $B_k(x_i)$ is carried out, $d_{\min}(x_i)$ is written to the array $D(x_i)$. The average of these values is expressed as follows:

$$\overline{d_{\min}(x_i)} = \frac{1}{N}\sum_k D_k(x_i). \quad (30)$$

When the minimum distance between the new pixel value and the sample set is smaller than the threshold, it is suggested that the pixel may be the background. At this time, the minimum distance can describe the complexity of the background. The larger the distance is, the more complex the background is. Therefore, the average value of the minimum distance in the sample set is invoked as a measure of background complexity. The decision threshold can be dynamically adapted as follows:

$$R(x_i) = \begin{cases} R(x_i) \cdot (1 - R_{\mathrm{inc/dec}}), & R(x_i) > \overline{d_{\min}(x_i)} \cdot R_{\mathrm{scale}} \\ R(x_i) \cdot (1 + R_{\mathrm{inc/dec}}), & \text{else.} \end{cases} \quad (31)$$

In order to respond to changes in the background, it is necessary to update the background template $B(x_i)$. Updating means that for a certain index, the corresponding background model value is replaced by the current pixel value. This means that the current pixel value may be placed into the background model with the update rate $T(x_i)$. We define the update rate as

Yanfeng WU *et al.*: Infrared LSS-Target Detection Via Adaptive TCAIE-LGM Smoothing and Pixel-Based Background Subtraction

185

$$T(x_i) = \begin{cases} \dfrac{\overline{\overline{d_{\min}(x_i)}} \cdot T(x_i)}{\overline{\overline{d_{\min}(x_i)}} + T(x_i) \cdot T_{\text{inc}}}, & F(x_i) = 1 \\[4mm] \dfrac{\overline{\overline{d_{\min}(x_i)}} \cdot T(x_i)}{\overline{\overline{d_{\min}(x_i)}} + T(x_i) \cdot T_{\text{dec}}}, & F(x_i) = 0 \end{cases} \tag{32}$$

$$T_{\text{lower}} < T(x_i) < T_{\text{upper}} \tag{33}$$

where T_{inc} and T_{dec} are fixed parameters, and the observed minimal distance can be calculated as (1). $F(x_i) = 1$ implies foreground, and $F(x_i) = 0$ implies background. The higher $T(x_i)$ is, the likely a pixel will be updated. Because the foreground area can't be used to update the model, the background model can only be used to update those pixels that currently belong to the background. The updating process is comparable to that of the visual background extractor (ViBe) algorithm [30]. When a pixel is considered as a background point, it is updated. Different from the ViBe algorithm, the update rate $T(x_i)$ is adaptive, and the neighboring pixel can be updated by its current pixel value randomly as well.

4. Experiments and results

In order to verify the effectiveness of our algorithm, 4 groups of video are tested, and the experimental results are compared with the three frame difference algorithms, mixture of Gaussian [31], and Kernel density estimation (KDE) algorithm [32]. As shown in Table 1, Quadrocopter_1 (video 1) is from real-time shooting of our infrared camera, Quadrocopter_2 (video 2) is from VOT-LTIR 2015 dataset, bat_seq1 (video 3) is from OTCBVS Benchmark, and plane (video 4) is from Terravic Motion IR Database.

In order to ensure the fairness and the accuracy of the experimental results, this article uses the open source code from cited articles to run the test video. Our method is written in C language. On the computer of Intel (R) Core (TM) i7-8650 CPU @ 4.3 GHz, 16.00 GB memory, NVDA GTX1070 graphics card, and Windows 10 system, the experimental environment is built with OpenCV 3.4.2 library and Microsoft visual C++ 2017 software.

Table1 Test sequences in our experiments.

Video	Frame number	Scale	$\overline{SCR}$
Video 1	215	640×512	0.34
Video 2	1010	640×480	1.02
Video 3	760	320×240	0.68
Video 4	1100	640×512	1.41

Our method consists of eleven tunable parameters which need to be well tuned for specific test videos under specific scenarios. This paper gives a set of optimized parameters, which has a good effect in the test dataset, as shown in Table 2.

Table 2 Parameters setting in our experiments.

Serial number	Parameter	value
(a)	φ	100
(b)	φ	1000
(c)	$R_{\text{inc}}/R_{\text{dec}}$	0.04
(d)	R_{lower}	20
(e)	T_{dec}	0.1
(f)	R_{scale}	10
(g)	ζ_1	0.1
(h)	ζ_2	0.8
(i)	T_{inc}	2
(j)	T_{lower}	20
(k)	T_{upper}	50

The calculation method of the local signal to clutter ratio (SCR) is as follows:

$$SCR = \frac{|\delta_t - \delta_b|}{\sigma_b}. \tag{34}$$

Among them, δ_t is the average value of the pixels in the local area of targets, and δ_b and σ_b are the average value and standard deviation of the pixels in the neighborhood of targets, respectively. The average signal to heterozygous ratio ($\overline{SCR}$) [33] is as

$$\overline{SCR} = \frac{1}{J_t} \sum_{i=1}^{N_t} SCR_i \tag{35}$$

where J_t is the number of targets, and SCR_i is the SCR of target i.

In order to test the degree of noise suppression, we use Video 1, which includes obvious stripe noise, to test our algorithm. The result is shown in Fig. 1. Figure 1(a) is the origin image, and Fig. 1(b) is the experimental result of our method. Through our method, the SNR of image is obviously improved. By calculating the correlation between two points at the same distance and to the same direction in either

image, the gray level co-occurrence matrix (GLCM) can reflect the comprehensive information including the direction, interval, and amplitude of the image. GLCM is a symmetric matrix, and the more convergent the diagonal matrix is, the better the correlation between pixels is. Figure 2(a) is the GLCM of the origin image, and Fig. 2(b) is the GLCM of the result image through our method. It can be seen from the contrast that noises have been greatly suppressed.

(a)   (b)
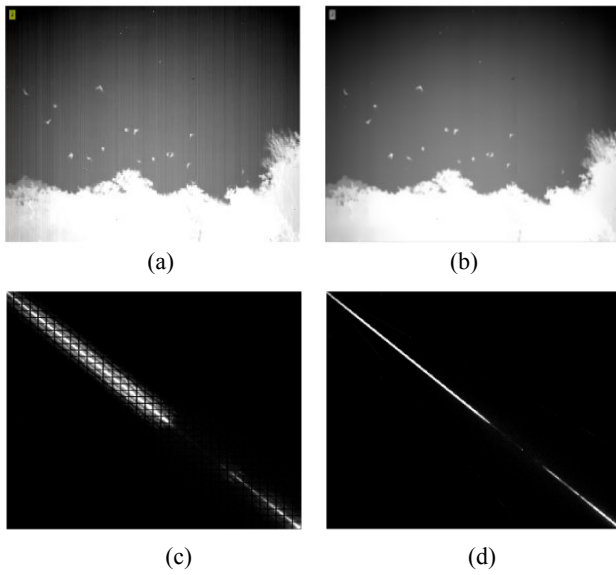

(c)   (d)

Fig. 1 Background suppression result: (a) the original image, (b) the result after adaptive TCAIE-LGM smoothing, (c) GLCM of the original image, and (d) GLCM of the smooth image.

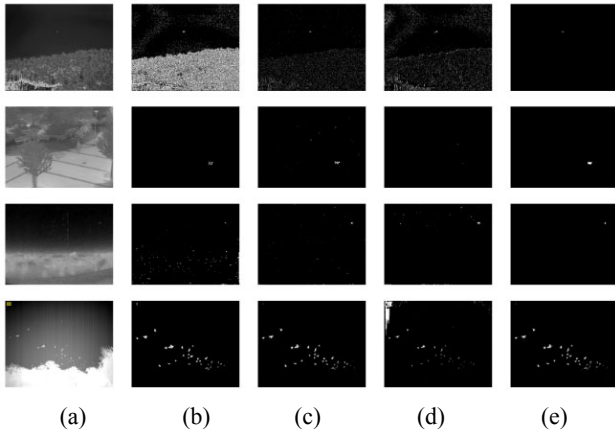(a)   (b)   (c)   (d)   (e)

Fig. 2 Experimental result: (a) the original images, (b) TFD algorithm, (c) mixture of Gaussian algorithm, (d) KDE algorithm, and (e) our method.

In order to quantitatively analyze the performance of the detection algorithm, the detection probability (P_d), the false alarm rate (F_a), and the time complexity of the algorithm (time complexity) are used in the experiment to evaluate the detection effect of the algorithm. The detection probability and the false alarm rate are defined as follows:

$$P_d = \frac{\text{number of true detections}}{\text{number of actual targets}} \quad (36)$$

$$F_a = \frac{\text{number of false detections}}{\text{number of frame insequence}} . \quad (37)$$

Table 3 Detection probability (%).

Test video	Detection probability (Pd)			
	Three frame difference	Mixture of Gaussian	KDE	Ours
Video 1	76.3	82.6	81. 7	95.6
Video 2	80.4	85.2	74.5	93.2
Video 3	69.6	89.8	86.3	94.3
Video 4	85.3	90.5	77.4	97.8

The detection probability and false alarm rate are obtained and shown in Tables 3 and 4. The time complexity of the algorithm is shown in Table 5. The average time is as follows:

$$t_{\text{average}} = \frac{\text{total time consumption}}{\text{frame number of video}} . \quad (38)$$

Table 4 False alarm rate (%).

Test video	False alarm rate (Fa)			
	Three frame difference	Mixture of Gaussian	KDE	Ours
Video 1	46.3	36.3	40.3	1.6
Video 2	20.4	10.8	16.7	4.3
Video 3	32.1	12.1	14.2	7.0
Video 4	13.3	9.4	35.3	7.8

Table 5 Average time consuming of the algorithm (ms).

Test Video	Average time (t_average)			
	Three frame difference	Mixture of Gaussian	KDE	Ours
Video 1	4.3	36.5	52.8	73.6
Video 2	4.1	33.2	50.9	72.7
Video 3	3.4	17.6	21.3	32.8
Video 4	3.8	35.9	53.2	71.4

In order to illustrate the comparison results more explicitly, the experimental results of our method and the comparison algorithm are given in Fig. 2. Among them, Fig. 2(a) is the original test video sequence (Video 1–Video 4), Fig. 2(b) is the experimental result of the frame difference

Yanfeng WU *et al.*: Infrared LSS-Target Detection Via Adaptive TCAIE-LGM Smoothing and Pixel-Based Background Subtraction

187

algorithm, Fig. 2(c) is the experimental result of the mixture of Gaussian algorithm, Fig. 2(d) is the experimental result of KDE algorithm, and Fig. 2(e) is the experimental result of our method.

5. Conclusions

In order to detect the low altitude, slow speed, and small target from the infrared complex background, this paper proposes a novel detection method via TCAIE-LGM smoothing and improved pixel-based background subtraction. Complexity and entropy are calculated and introduced into L0 gradient minimization smoothing, which can remove noises and suppress low-amplitude details in infrared image abstraction. In addition, difference of Gaussian map is integrated into the adaptive pixel-based segmentation background modeling algorithm, which can differentiate LSS-Target from the sophisticated background. Experimental results show that the proposed method significantly outperforms the existing methods in detection accuracy and can detect the infrared LSS-target with the static complex background.

However, our method also has some limitations. Once the target remains stationary for a long time, the moving target will gradually be absorbed by the background. The updating of our model occupies a great deal of memories and resources, which leads to low computational efficiency of the algorithm. Our method needs to be further optimized to solve the above problems in the future.

Acknowledgment

References

[1] J. Hu, Y. Yu, and F. Liu, "Small and dim target detection by background estimation," *Infrared Physics & Technology*, 2015, 73: 141–148.

[2] W. P. Yang, X. P. Lu, J. C. Li, and Z. L. Zhang, "Fast algorithm of infrared small target detection in jitter background," *SPIE*, 2015, 9476: 947614-1–947614-6.

[3] H. Deng, J. G. Liu, and Z. Chen, "Infrared small target detection based on modified local entropy and EMD," *Chinese Optics Letters*, 2010, 8(1): 24–28.

[4] E. Abdelkawy and D. Mcgaughy, "Small infrared target detection using two-dimensional fast orthogonal search (2D-FOS)," *SPIE*, 2003, 5094: 179–185.

[5] T. Xie, Z. Chen, and R. Y. Ma, "A novel method for infrared small target detection based on PGF, BEMD and LIE," *Journal of Infrared & Millimeter Waves*, 2017, 36(1): 92–101.

[6] D. Y. Huang, A. K. Xue, and Y. F. Guo, "Penalty dynamic programming algorithm for dim targets detection in sensor systems," *Sensors*, 2012, 12(4): 5028–5046.

[7] J. L. Gao, C. L. Wen, and M. Q. Liu, "Robust small target Co-detection from airborne infrared image sequences," *Sensors*, 2017, 17(10): 2242-1–2242-21.

[8] Z. Z. Li, J. Chen, Q. Hou, H. X. Fu, Z. Dai, R. Z. Li, *et al.*, "Sparse representation for infrared dim target detection via a discriminative over-complete dictionary learned online," *Sensors*, 2014, 14(6): 9451–9470.

[9] S. Kim, "Sea-based infrared scene interpretation by background type classification and coastal region detection for small target detection," *Sensors*, 2015, 15(9): 24487–24513.

[10] S. Kim and J. Lee, "Small infrared target detection by region-adaptive clutter rejection for sea-based infrared search and track," *Sensors*, 2014, 14(7): 13210–13242.

[11] X. Z. Bai and F. G. Zhou, "Analysis of new top-hat transformation and the application for infrared dim small target detection," *Pattern Recognition*, 2010, 43(6): 2145–2156.

[12] P. Wang, J. W. Tian, and C. Q. Gao, "Infrared small target detection using directional highpass filters based on LS-SVM," *Electronics Letters*, 2009, 45(3): 156–158.

[13] E. Guariglia, "Entropy and fractal antennas," *Entropy*, 2016, 18(3): 1–17.

[14] I. S. Reed, R. M. Gagliardi, and L. B. Stotts, "Optical moving target detection with 3-D matched filtering," *IEEE Transactions on Aerospace & Electronic Systems*, 2002, 24(4): 327–336.

[15] X. M. Shen and L. Deng, "Game theory approach to discrete H, filter design," *Signal Processing IEEE Transactions*, 1997, 45(4): 1092–1095.

[16] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, 1997, 22(1): 61–79.

[17] Z. J. Liu, C. Y. Chen, and X. B. Shen, "Detection of small objects in image data based on the nonlinear principal component analysis neural network," *Optical Engineering*, 2005, 44(9): 093604-1– 093604-9.

[18] C. Yao and W. Chen, "Research on infrared dim and small target detection based on improved particle algorithm," *Progress in Laser and Optoelectronics*, 2017(11): 143–148.

[19] J. Liu and H. Ji, "Infrared dim and small target detection based on mobile weighted pipeline filtering," *Journal of Xi'an Electronic and Science University* (*NATURAL SCIENCE EDITION*), 2007, 34(5): 743–747.

[20] Y. Huang, X. F. Zhang, and Y. U. Xin, "Pipeline filtering detection method for photon imaging stationary point target," *Chinese Optics*, 2013, 6(1): 73–79.

[21] H. S. Nie, Z. J. Huang, J. T. Diao, J. Chen, H. J. Liu, and Q. Li, "A Wiener filter based infrared small target detecting and tracking method," in *Proceeding of International Conference on Intelligent System Design and Engineering Application*, Changsha, China, 2010, pp. 184–187.

[22] Y. F. Wu, H. J. Sun, and P. X. Liu, "A novel fast detection method of infrared LSS-Target in complex urban background," *International Journal of Wavelets Multiresolution & Information Processing*, 2018, 16(01): 1619–1632.

[23] D. A. Scribner, K. A. Sarkady, and M. R. Kruer, "Adaptive retina-like preprocessing for imaging detector arrays," *Proceeding of IEEE International Conference on Neural Networks*, 1993(3): 1955–1960.

[24] D. A. Scribner and J. T. Caulfield, "Nonuniformity correction for staring IR focal plane arrays using scene-based techniques," in *Proceedings of SPIE: The International Society for Optical Engineering*, San Francisco, CA, USA, 1990, pp. 224–233.

[25] J. G. Harris, "Nonuniformity correction of infrared image sequences using the constant-statistics constraint," *Image Processing IEEE Transactions*, 1999, 8(8): 1148–1151.

[26] H. L. Qin, S. Q. Liu, H. X. Zhou, and R. Lai, "Nonuniformity correction algorithm based on wavelet transform for infrared focal plane arrays," *Acta Optica Sinica*, 2007, 7(9): 1617–1620.

[27] L. Xu, C. W. Lu, Y. Xu, and J. Y. Jia, "Image smoothing via L_0 gradient minimization," *ACM Transactions on Graphics*, 2011, 30(6): 1–12.

[28] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, 1948: 27(3): 379–423.

[29] M. Hofmann, P. Tiefenbacher, and G. Rigoll, "Background segmentation with feedback: the pixel-based adaptive segmenter," in *Proceeding of 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Providence, RI, USA, 2012, pp. 38–43.

[30] O. Barnich and M. V. Droogenbroeck, "ViBe: a universal background subtraction algorithm for video sequences," *IEEE Transactions on Image Processing*, 2011, 20(6): 1709–1724.

[31] Sadarangani and Nikhil, "An improved Gaussian mixture model algorithm for background subtraction," *Massachusetts Institute of Technology*, 2002, pp. 1–72.

[32] A. Elgammal, R. Duraiswami, D. Harwood, and L. S. Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," *Proc. of the IEEE*, 2002, 90(7): 1151–1163.

[33] C. Y. Wang and S. Y. Qin, "Adaptive detection method of infrared small target based on target-background separation via robust principal component analysis," *Infrared Physics & Technology*, 2015, 69: 123–135.