Research Article

Infrared Dim and Small Targets Detection Method Based on Local Energy Center of Sequential Image

Xiangsuo Fan,^{1,2,3} Zhiyong Xu,^{1,3} Jianlin Zhang,^{1,3} Yongmei Huang,^{1,3} and Zhenming Peng²

¹*Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, China*

²*School of Optoelectronic Information, University of Electronic Science and Technology of China, Chengdu 610054, China*

³*University of Chinese Academy of Sciences, Beijing 100039, China*

Correspondence should be addressed to Zhiyong Xu; xzy158@163.com and Yongmei Huang; huangym@ioe.ac.cn

Received 19 February 2017; Revised 2 May 2017; Accepted 10 May 2017; Published 4 July 2017

Academic Editor: Aimé Lay-Ekuakille

Copyright © 2017 Xiangsuo Fan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to detect infrared (IR) dim and small targets in a strong clutter background, a method based on local energy center of sequential image is proposed. This paper began by using improved anisotropy for background prediction (IABP), followed by target enhancement by improved high-order cumulates (HOC). Finally, on the basis of image preprocessing, the paper constructs a sequential image energy center detection algorithm that integrates the neighborhood, continuity, area, and energy and other motion characteristics of the target. Experiments showed that the improved anisotropic background predication could be loyal to the true background of the original image to the maximum extent, presenting a superior overall performance to other background prediction methods; the improved HOC significantly increased the signal-noise ratio of images; when the signal-noise ratio (SNR) is lower than 2.5 dB, the proposed method could effectively eliminate noise and detect targets.

1. Introduction

At present, detection of the IR dim and small moving target under the strong clutter environment is a core technology for infrared search and tracking system and has been hot and difficult research topic in this field. According to the precedence relationship between detection and tracking, the traditional detection algorithms can be divided into two categories: detect before track (DBT) and track before detect (TBD). The DBT algorithm is to extract the candidate targets from the preprocessed images and then use the sequence trajectory analysis method to confirm the target. As the key to sequence trajectory analysis is based on the motion continuity of the target and the randomness of noises to eliminate the influence of false targets, adverse to the detection of dim and small moving targets submerged in various noises and clutters, the DBT algorithm is only applied to scenes with high signal-to-noise ratio (SNR > 5 dB). In order to detect the targets in the case of low SNR, researchers also propose TBD algorithm [1–5]. The TBD algorithm firstly searches all possible trajectories of the target and applies the appropriate method to complete the interframe energy cumulates. By comparing the posterior

probability of each trajectory, the threshold method is used to judge the real trajectory of the target [6].

Either the DBT algorithm or the TBD algorithm is IR dim and small target detection algorithm based on multiframe image processing, except that both algorithms differ in the sequential processing order of interframe information. Various methods proposed in literature [1–5] need to know a priori knowledge of the target, including the movement state and trajectory. In the actual infrared scene, scope of application of those methods is undoubtedly. In addition, those methods only focus on studying the target judgment method and ignore study of the target criteria, namely, continuity, neighborhood, speed, area, energy, and other target characteristics. Thus, there is an underutilization of target movement information, and so forth, and their detection performance fails to meet the requirements in case of low SNR.

To this end, this paper attempts to start from the motion characteristics of IR dim and small targets and analyze how to rationally use motion information of the target and explore the criteria which are applicable for IR dim and small targets to achieve the target detection. It requires using the various

motion characteristics of the target as much as possible but minimizes the excessive restrictions on the target's motion characteristics and reliance on a priori knowledge. The analysis result shows that the target is relatively stable concentrated in a certain neighborhood in the sequential image, forming a certain area and concentration characteristics. However, the noise presents a random discrete distribution in the image sequence. Therefore, the energy center where the composite features, such as target neighborhood of sequential image and area and concentration degree, are integrated is proposed to realize the multiframe motion correlation detection of the target.

2. Related Works

The paper launches its related works from the discussion on the target detection algorithm and background estimation of the dim and small target.

2.1. Target Detection Algorithm. The paper conducts research on the typical detection algorithms of dim and small target in the last dozen years, which emphatically analyzes their application scenes and disadvantages, discovers problems from the disadvantages, and then proposes detection algorithm applied to scenes with low SNR. The typical DBT algorithm includes a pipeline filter method [7], which is simple and easy to implement. This method has better detection effect when the signal-noise ratio is high (>5 dB) but fails when the SNR is low and the target position makes no change. In addition, in the pipeline, once the frame detection error occurs, target detection of the next frame in the pipeline has a certain degree of risk, because the incorrect pipeline location increases the possibility of the target falling outside the pipeline, resulting in the failure of the target detection. This case is likely to happen in the practical application. The classical TBD algorithms include the following: the three-dimensional matching filter method proposed by Reed et al. [1]. This method is only applicable to the case where the velocity magnitude and direction are known. The unknown velocity may lead to speed mismatching, which leads to the decrease in the output SNR. In addition, this method is only applicable to scenes with small speed variations, due to limitations by computational complexity. The classical TBD algorithms also include the projection transformation method proposed by Falconer [2], which can effectively reduce the amount of data and the memory space in the three-dimensional search and detection process, but it may cause loss of SNR. In addition, when the noise is large and there is a large displacement between target frames, the detection performance of the method is rapidly reduced. It does not adapt to the target detection with low SNR and large inter-frame displacement. The dynamic programming proposed by Johnston and Krishnamurthy [3] is that this method can detect the point target trajectory of a linear motion in the case of a low SNR, but requires a priori knowledge of the velocity window parameter. If the target velocity is unknown, the parameter range of the velocity window must be liberalized, which will lead to the increase in computation and reduce the detection performance of the algorithm. A multilevel

hypothesis testing method proposed by Blostein and Huang [4] can detect multiple targets of linear motion at the same time. However, in the case of low SNR, in order to reduce the false alarm rate, there are many candidate trajectory starting points, resulting in a sharp increase in tree branches behind and a rapid increase in computational complexity. Furthermore, it needs to limit the target to the extent where the target must do the local uniform linear motion, which it is difficult for IR dim and small targets to meet in most cases. Thus, this method has limited scope of application. There is also a high-order correlation method proposed by Liou and Azimi-Sadjadi [5] which can detect straight or curved trajectories from a noisy three-dimensional image, without the need to know the a priori knowledge of the number of targets, the initial conditions, and so forth. This method can be applicable to multitarget detection under different clutter density. However, if the order is too high, the amount of computation and storage increase; if the order is too low, the false alarm rate increases. In recent years, as the machine learning algorithm has been rapidly developed in the field of target detection, some researchers have proposed small and dim moving target detection methods that are based on visual significance [8, 9] and sparse representation [10–13]. The method which is based on the notion that the significance performs well only when there is a great difference between the target and the background, while practicality of the sparse representation-based method is undoubtedly limited, when the target signal is seriously polluted by the noise and the sparse feature between the target and the background is demined seriously. It is difficult for these methods to meet the requirements of detection of small and dim targets, in case of low SNR ($\text{SNR} < 3$ dB).

All the detection algorithms proposed above exert poor detection effects on scenes with low SNR ($\text{SNR} < 3$ dB). In order to explore a new algorithm of multiframe motion associated detection applied to scenes with low SNR, based on the summary and analysis about advantages and disadvantages of each detection method (refer to Table 1 for details), the paper proposes a local energy central detection algorithm of sequence image that takes full advantage of the motion information, neighborhood gray scale, area, energy, and other characteristics embodied by the target in the time-space domain.

2.2. Background Estimation. Due to the low SNR of the target, plus the serious interference from noises, in order to improve the follow-up detection ability, the background estimation method is often required to estimate background pixel from the image first and then subtract the estimated part from the original image, so as to obtain an image containing the target components and part of the noises, followed by subsequent detection processing. Representative methods are Top-Hat, TDLMS [14], and so on; Top-Hat is a kind of practical nonlinear background estimation method, which tends to be affected by the structural elements, with poor adaptability as a consequence. In order to enhance the adaptability of the algorithm, some scholars put forward adaptable filtering technology, such as two-dimensional minimum mean square error filter (TDLMS), which requires no understanding of

TABLE 1: Advantages and disadvantages about different detection algorithms.

	Algorithm	Advantages	Disadvantages
DBT	Pipeline filtering method [7]	Simple process; easy for engineering implementation	Failure when the position of target does not change and low SNR
TBD	3D matched filtering method [1]	High detection performance; able to detect multiple trajectories simultaneously	Only applied to the case of known speed and direction
	Project transformation method [2]	Effectively reducing the amount of data and storage during the 3D search and detection process	Not adapted to the target detection with low SNR and large inter frame displacement
	Dynamic programming method [3]	Able to detect the target trajectory of points in linear motion in the case of low SNR	Requiring a priori knowledge of the velocity window parameters
	Multistage hypothesis testing method [4]	Able to detect multiple targets in linear motion simultaneously	Only adapted to the scene of targets in local uniform linear motion
	High-order correlation method [5]	Able to detect the linear or curve trajectory, requiring no prior knowledge	Detection results being affected greatly by order
Latest algorithms	Visual saliency [8, 9]	Able to quickly locate the region of interest	Only adapted to scenes with big differences between the target and the background, and more obvious characteristics for the target
	Sparse representation [10–13]	Effectively enhancing the sparse feature difference between the target and the background and improving the detection accuracy through training	Only adapted to stable or slowly changing background, and scenes with high SNR
	The proposed method	Able to effectively detect scenes with low SNR ($\text{SNR} < 3 \text{ dB}$)	Requiring large amount of computation

prior knowledge for the image and has a simple structure but requires statistic characteristics of the background to be constant or slowly changing, dramatically limiting the application scope. There are also background estimation methods based on statistics, such as single Gaussian background estimation method proposed by Benezeth et al. [17], which can deal with simple scenes with tiny and slow changes, except precisely describing the background when the background changes substantially or suddenly or background pixels present multimodal distribution. In order to solve the background of multimodal distribution, Bouwmans et al. [19] propose mixed of Gaussian model; the algorithm is adapted to the dynamic background estimation of long time series, whose disadvantages are a certain amount of training data to be required and rapid changes against the illumination, resulting in poor effects on the shadow processing. For the uncertainty of model parameters brought by the mixed of Gaussian during the background estimation process due to noise interference or deficiency in training data, Sigari et al. [18] propose a fuzzy running average method, which is suitable for the background estimation in camera shake and dynamic scenes but difficult to get rid of the shadow. Compared with the background estimation methods based on statistics, the nonparametric background model method owns the following advantages: it requires no potential model to be specified or explicit estimation parameters. Therefore, they can adapt to any unknown data distribution. For instance, Liu et al. [15, 22] adopt the model described by influencing factors to describe the changes of the background and then deduce the most reliable background status with the potentially distributed local extremism, so as to find out the point gathering most intensive data in the density distribution of data. This model is robust, able to adapt to the background under the scene that is chaotic and incompletely static, but contains small disturbance. With the complexity of application scenes, such as the illumination variation, camera shake, and dynamic background, Sobrala and Vacavant have reviewed background estimation methods for these scenes over the past dozen years [23] and developed an open source background estimation library called BGSLibrary [24], laying the foundation for subsequent scholars to conduct corresponding researches. However, all these algorithms are only applied to the circumstances like fixed scene shooting, slight shake, or slow moving of cameras, showing poor adaptability for the scenes of camera fast moving along with the small dim target, as well as the environment with low SNR.

In recent years, some scholars have been trying to achieve background estimation by separating the “gray singularity” formed by the “gray disturbance” on the image caused by the target. For example, Song et al. [25] describe the “gray singularity” of the target area with gradient operators and achieve background estimation according to this. However, it is difficult for gradient operators to distinguish the strong

texture of the target and complex background, leading to poor suppressing effects of the algorithm on the background edge texture. Besides, Wang and Liu [26] adopt anisotropic diffusion filter to separate the gradient feature differences between the target and the background, to improve the SNR of the image. The algorithm belongs to the unidirectional diffusion, which cannot enhance the target signal but only reserve them negatively, so its improvement in image SNR is not sufficient. Considering that the target signals in the imaging system are the process of diffusion outwards from the center pixel and the gradient relationship of the anisotropy in different directions in each pixel is similar to that of the point spread function, the idea of anisotropy is introduced into background estimation in the paper and then improved to enhance the target signal.

3. Prediction of Anisotropic Background

3.1. Anisotropic Differential Principle. Background prediction is crucial for dim small target detection. On one hand, anisotropy is featured with the ability to smoothen and stabilize the background region and meanwhile reserve the marginal details and mutating zones in the background; the diffusion equation is

$$\frac{\partial u(x, y, t)}{\partial t} = \text{div} [c(\nabla u)] \nabla u, \quad (1)$$

where u is the grayscale image, ∇u is the gradient, $c(\nabla u)$ is the edge stop function, and div is the divergence operator. The edge stop function $c(\nabla u)$ calculates the smoothing coefficient based on relations of gradients in different directions. Literature [27] presents the anisotropic edge stopping function as follows:

$$\begin{aligned} c_1(\nabla u) &= \frac{1}{1 + (|\nabla u|/k)^2}, \\ c_2(\nabla u) &= \exp \left[- \left(\frac{|\nabla u|}{k} \right)^2 \right]. \end{aligned} \quad (2)$$

Wherein, k is a constant; for a flat region with small gradient, the $c(\nabla u)$ value is large and will use high level of smoothening; for a mutation region with large gradient, the $c(\nabla u)$ value is small and will use low level of or no smoothening, which will reserve these regions.

3.2. Improved Anisotropy for Background Prediction. Analysis on images of dim small targets shows that the differences of each direction's features between the target region and other regions could serve to realize differentiated disposal of different characteristic regions. The gradient operator of the local region where dim small targets are located is shown in the following:

$$\begin{pmatrix} c(f(i, j) - f(i - \text{step}, j)) \\ c(f(i, j) - f(i, j - \text{step})) & f(i, j) & c(f(i, j) - f(i, j + \text{step})) \\ c(f(i, j) - f(i + \text{step}, j)) \end{pmatrix}. \quad (3)$$

If the mean value of $\min_1$ and $\min_2$, the two smallest parameters of the original anisotropic edge stopping function's four directions, is used to carry out pixel-by-pixel filtering of the image, and the values of parameters of the stable background and nonstable background are found to be relatively big, with relatively small parameter values of the singularity region (target signals), the target signals can only be reserved but cannot be enhanced. The $\min_1$ and $\min_2$ equation is as follows:

$$\begin{aligned} \min_1 &= \min \{c(f(i, j) - f(i - \text{step}, j)), c(f(i, j) \\ &\quad - f(i + \text{step}, j)), c(f(i, j) - f(i, j - \text{step})), \\ &\quad c(f(i, j) - f(i, j + \text{step}))\}, \\ \min_2 &= \min \{c(f(i, j) - f(i - \text{step}, j)), \\ &\quad c(f(i, j) - f(i + \text{step}, j)), \\ &\quad c(f(i, j) - f(i, j - \text{step})), \\ &\quad c(f(i, j) - f(i, j + \text{step})) \mid \text{not } \min_1\}. \end{aligned} \quad (4)$$

$$\left(\begin{array}{c} c(f(i, j) - f(i - \text{step}, j)) \\ c(f(i, j) - f(i, j - \text{step})) \quad \frac{\min_1 + \min_2}{2} \quad c(f(i, j) - f(i, j + \text{step})) \\ c(f(i, j) - f(i + \text{step}, j)) \end{array} \right). \quad (6)$$

4. High-Order Cumulates Target Enhancement

4.1. High-Order Cumulates Principle. The high-order cumulates can effectively accumulate the space-time domain energy, suppress Gaussian noise, enhance the transient signal, and achieve the goal of enhancing dim and small target energy [20]. As the noise of the infrared sequential image can be regarded as Gaussian noise, the following binary hypothesis is applied to the image with the removed background:

$$\begin{aligned} H_0: F_0(x, y, k) &= N(x, y, k), \\ H_1: F_0(x, y, k) &= F_T(x, y, k) + N(x, y, k), \end{aligned} \quad (7)$$

where $F_0(x, y, k)$ is the image with the removed background, H_0 is the pixel of the area where the target is not located, and H_1 is the pixel of the area which the target passes through. M frame high-order cumulates can be described as follows:

$$\begin{aligned} C_{Mf} &= E \left\{ F_0 \left(x, y, \frac{1-M}{2} \right) + F_0 \left(x, y, \frac{3-M}{2} \right) + \dots \right. \\ &\quad \left. + F_0 \left(x, y, \frac{M-1}{2} \right) \right\} = C_{MT} + C_{MN}, \end{aligned} \quad (8)$$

where C_{MT} refers to cumulates of the target and C_{MN} refers to cumulates of noise. Because the noise $N(x, y, k)$ obeys the

To highlight the singularity region's signals of enhanced targets, the edge stopping function is improved as follows. Its function image is shown in Figure 1:

$$\begin{aligned} c_1(\nabla u) &= 1 - \frac{1}{1 + (|\nabla u|/k)^2}, \\ c_2(\nabla u) &= 1 - \exp \left[- \left(\frac{|\nabla u|}{k} \right)^2 \right]. \end{aligned} \quad (5)$$

In Figure 1, the horizontal axis is the direction gradient value and the vertical axis is the function value; the improved edge stopping function is a monotonically increasing function. As for the stationary and nonstationary regions of the infrared image, a small gradient value will cause a small edge stopping function value. As for singular region, the edge stop function value is larger when the gradient is larger. Substitute (5) into (3); then use (4) to evaluate the mean value of $\min_1$ and $\min_2$, the two smallest parameters of the original anisotropic edge stopping function's four directions; then apply pixel-by-pixel filtering of the image with this mean value, and the dim small targets would be able to be highlighted smoothly. The filter equation is as follows:

Gaussian distribution, $C_{MN} = 0 (M \rightarrow \infty)$, and $C_{Mf} = C_{MT}$. After using the M high-order cumulates as the detection statistic,

$$|C_{Mf}| \begin{cases} = 0, & \text{Not target,} \\ > 0, & \text{Target.} \end{cases} \quad (9)$$

4.2. Improved HOC. The original high-order cumulates only take the time domain's characteristics into consideration, which will inevitably affect the enhancement effect. To better accumulate its energy on the space domain, the space domain characteristics (motion information of the target) need to be considered. The target's movement in adjacent frames could be described as 12 forms in Figure 1. The first five forms are horizontal movement. The middle five forms are vertical movement and the last two forms are diagonal movement, as shown in Figure 2.

Whichever direction the target movement is, it always exists in the continuous neighborhood region of the neighboring frames. Therefore, the moving target's energy could be accumulated inside the moving neighborhood region of the

FIGURE 1: Images of edge stopping functions $c1$ and $c2$.

FIGURE 2: Target movement model.

movement of continuous M frames of images. The movement energy cumulates of the target could be described as follows:

$$f_p(x, y) = \sum_{m=-r, n=-r}^{m=r, n=r} t_p(x, y) \times F_0(x + m, y + n, k), \quad (10)$$

$$P_0(x, y) = \arg \max_p f_p(x, y).$$

Here, t_p is the template of the target neighborhood region, r is the accumulative window radius, and F_0 is the blocking-out image sequence with the purpose of extracting the maximum values of the adjacent frames' target neighborhood regions as the accumulative value. The improved M frames of HQS could be defined as follows:

$$C_{Mf} = E \left\{ P_0 \left(x, y, \frac{1-M}{2} \right) + P_0 \left(x, y, \frac{3-M}{2} \right) + \dots + P_0 \left(x, y, \frac{M-1}{2} \right) \right\}. \quad (11)$$

5. Local Energy Center of Sequential Image

Local energy center (LEC) of sequential image refers to the center of the target moving energy region formed by IR sequential image through continuous multiframe energy

cumulates and determined by the target moving area S , the target concentration degree I , and the moving area multiple G_s . The specific formula is as follows:

$$S = \sum_x \sum_y F_L(x, y),$$

$$F_L(x, y) = f_L(x, y, 1) \mid f_L(x, y, 2) \cdots \mid f_L(x, y, N),$$

$$I = \sum_{k=1}^N T(f_L(x, y, k)), \quad (12)$$

$$G_s = \frac{S}{\bar{S}},$$

$$\bar{S} = \frac{\sum_1^N \sum_y \sum_x f_L(x, y, k)}{I},$$

where $f_L(x, y, k)$ is a local binary image of the candidate target at the k frame; $F_L(x, y)$ is a region where the moving trajectory of the target is projected onto the binary image and is obtained by a multiframe image or an operation; $T(\cdot)$ refers to a function which is used to calculate the number of candidate targets in local region; I is the frequency at which the candidate target appear in a sequential region; S is the target moving area; $\bar{S}$ is the mean area of the candidate target and is obtained by comparing the accumulated N frame area in a local region with the degree of concentration.

5.1. Proposed Algorithm. Acquire the local energy center of sequential image, to seek for the cumulative frame length in $(2R+1) \times (2R+1)$ neighborhood of each candidate target point and use them as three elements of the $(N-1)/2$ frame before and after the current frame. Then, according to the formula (13), sort them by size. Select the mass center of the candidate with the smallest serial number as the local energy center of sequential image. The expression is as follows:

$$\text{LEC} = \arg \min (\text{order}(S) + \text{order}(I) + \text{order}(G_s)), \quad (13)$$

where $\text{order}(\cdot)$ refers to the sorting function, $\min(\cdot)$ represents the function which is used to achieve a minimum, and $\arg$ refers to the parameter satisfying the condition. The detection method is as shown in Algorithm 1.

6. Background Prediction Performance and Enhancement Effect Evaluation

For evaluation of background prediction result, in this study, we use three indexes, Mean Squared Error (MSE) [28], Structural Similarity (SSIM) [29], and local Signal-to-Noise Ratio Gain (GSRN) [30], to evaluate the effect of image background prediction. The enhancement effect of high-order cumulates is evaluated by using the target's average grayscale value (AGV) and the image's local signal-to-noise ratio (LSNR).

- (1) **Initialize the parameters:** R -target search neighborhood region, Num-number frames of images
- (2) Use IABP to acquire the difference graph. Then use improved HOC to enhancement the difference graph. Finally, use the local maximum value partitioning (LMVP) [21] to obtain the binary image of the sequence
- (3) **Input:** Input N frames of binary image.
- (4) **Output:** the target in every frame of image's center (x, y)
 - (1) FOR $i = 1:\text{Num}$
 - (2) Record the i 's width as m and high as n
 - (3) Record the i 's all candidate targets x_i ($i = 1, 2, 3, \dots$)
 - (4) FOR $i = 1:m$
 - (5) FOR $j = 1:n$
 - (6) IF $i < x - R \mid i > x + R \mid j < y - R \mid j > y + R$
 - (7) Use formula (12) to obtain the three elements- S , I and G_s
 - (8) END IF
 - (9) END FOR
 - (10) END FOR
 - (11) Use formula (13) to obtain the true target (x, y)
 - (12) END FOR
- (5) Record the points (x, y) of the target in every frame of image and export corresponding detection result

ALGORITHM 1

(1) MSE is used to calculate the average error between each pixel value of the predicted background image and the real background image. The equation is as follows:

$$\text{MSE} = \frac{1}{MN} \sum_{x=1}^N \sum_{y=1}^M [R(x, y) - F(x, y)]^2, \quad (14)$$

where F is the predict background image; R is the real background image (because of dim and small infrared target is very dim, so use infrared image as real background image); M and N are image width and height, respectively.

(2) SSIM is used to evaluate the degree of similarity of geometric structure information of the predicted and the real background; the parameters are very effective for the evaluation of the performance of the image background prediction. The equation is as follows:

$$\mu_R = \frac{1}{NM} \sum_{x=1}^N \sum_{y=1}^M R(x, y),$$

$$\sigma_R = \sqrt{\frac{1}{NM-1} \sum_{x=1}^N \sum_{y=1}^M [R(x, y) - \mu_R]^2}, \quad (15)$$

$$\sigma_{RF} = \frac{1}{NM-1} \sum_{x=1}^N \sum_{y=1}^M [R(x, y) - \mu_R] [F(x, y) - \mu_F],$$

$$\text{SSIM} = \frac{(2\mu_R\mu_F + \varepsilon_1)(2\sigma_{RF} + \varepsilon_2)}{(\mu_R^2 + \mu_F^2 + \varepsilon_1)(\sigma_R^2 + \sigma_F^2 + \varepsilon_2)},$$

where F , R , M , and N are as defined above; μ_R represents the real background pixels mean; σ_R represents the real background of standard deviation; σ_{RF} is the background covariance; ε_1 and ε_2 are a small constant to ensure that denominator is not 0.

(3) GSNR is the mean value of signal-to-noise ratio in the sequence frames. The equation is as follows:

$$\text{SNR} = \frac{(g_t - g_b)}{\sigma},$$

$$\text{GSNR} = \frac{1}{N} \sum_{i=1}^N \left(\frac{\text{SNR}_{\text{out}}^i}{\text{SNR}_{\text{in}}^i} \right), \quad (16)$$

where g_t is the maximum value of the target area; μ_b is the mean value of the local region of the target; σ is the standard deviation of the local region of the target.

(4) The formulas for AGV and LSNR are as follows:

$$\text{AGV} = \frac{\sum_{i=1}^m I_{i,j}}{m},$$

$$\text{LSNR} = \frac{\mu_t - \mu_k}{\sigma_k}, \quad (17)$$

where $I_{i,j}$ is the grayscale value of the candidate target at row and column of (i, j) . m is the total number of pixels occupied by the candidate target. μ_t is the local mean value of the candidate target. μ_k is the local background mean value, and σ_k is the local background standard deviation. The size of the local background area is generally 3 times of the target area.

7. Experimental Results and Analysis

7.1. Background Prediction Results and Analysis. Infrared image sequence obtained in an actual scene and 6 image frames with different SNR are used for experiments. Here the SNR is defined as $\text{SNR} = 10\log_{10}((u_t - u_b)/\sigma_b)$, with unit being dB, where u_t is the target area mean value, u_b is the background area mean value, and σ_b is the background area standard deviation. The background area is generally 3 times of the target area. For example, when the target size is

TABLE 2: Signal-noise ratio of 6 frames of images.

Number	1	2	3	4	5	6
SNR	2.45	1.05	0.86	2.21	2.19	2.13

3×3 , the background area is the 9×9 range centering on the target. In this paper, the improved anisotropic background prediction method is used to predict the background. The edge stop function c_2 is selected, with $k = 120$, step = 4. Comparative analysis is compared with Top-Hat, TDLMS [14], nonparametric background method [15], anisotropic background prediction (ABP) method [16], single Gaussian (SG) [17], fuzzy running average (FRA) [18], and mixed of Gaussian (MoG) [19]. The three indicators MSE, SSIM, and GSNR are used to evaluate the background prediction effect of the infrared images. A 5×5 "square" structure is adopted for Top-Hat. The settings for other methods are referenced from the literature [14–19]. The experimental results are listed from Tables 2–5.

The smaller the MSE value is, the smaller the error is, indicating that the background prediction effect is better. The closer the SSIM value is to 1, the closer the predicted background is to the real background. The larger the GSNR value is, the better the target enhancement effect of the difference image obtained from the background prediction is. Through comparison of the three performance indicators of MSE, SSIM, and GSNR, it can be seen that the improved anisotropic background prediction method is better than other background prediction algorithms in terms of the prediction effect.

Meanwhile, an image frame whose SNR is 0.86 is selected and the above methods in this paper are used to predict the background. The results are shown in Figure 3, in which (a) is the original infrared image where the target has been marked with a red rectangle; (b) shows the background prediction and difference and three-dimensional graphs of Top-Hat; (c) shows those of TDLMS; (d) shows those of the nonparametric method; (e) shows those of the anisotropic method; (f) shows those of the improved anisotropic background prediction method; (g) shows those of the single Gaussian method; (h) shows those of the fuzzy running average method; (i) shows those of the mixed of Gaussian method.

As can be seen from Figure 3, the backgrounds of traditional background prediction methods (Top-Hat and TDLMS) are blurred and there is a significant block effect. The nonparametric method needs to increase the number of training background frames in order to obtain a clearer background image. However, this will severely undermine the adaptability of the background model. The difference graphs obtained by single Gaussian, fuzzy running average, and mixed of Gaussian are prone to target drift or losses. They need different number of training frames in different scenes to obtain the background model, and they are only applicable to scenes where the background changes very slowly or the background is stationary. The anisotropic method can only negatively retain the target signal but cannot enhance the target signal. The improved anisotropic background prediction method can effectively eliminate most of the background

in the image. It not only preserves the edge contours of stationary background and nonstationary background but also eliminates the problems of block effect and target drift. After calculating the difference with the original image, it can extract the candidate target and reduce the false alarm rate.

7.2. Enhancement Results and Analysis

7.2.1. Parameter Selection Analysis. The main parameters of high-order cumulates include radius r of the cumulative window and length M of the cumulative frame. To achieve effective accumulation of the energy of a moving target, the cumulative window radius r , the cumulative frame length M , and the target moving velocity v must satisfy the following equation:

$$r \geq (M - 1) \times v. \quad (18)$$

In order to reduce the accumulation of noise energy in the image, it is desirable to select the window radius r that is the minimum value satisfying (18). The relationship between the SNR gain and the cumulative length M after the accumulation is shown in Figure 4. It can be seen that when the cumulative length M is set to 4 frames, the SNR enhancement effect is better. In the experiments, the cumulative frame length is $M = 4$ and considering that the movement of a small target in long distance infrared imaging is slower (usually $v \leq 2$ pix/s), the cumulative window radius is $r = 4$.

7.2.2. Results and Analysis. In order to verify the enhancement effect of high-order cumulates, a simulation experiment is done on an image frame whose SNR is 1.05. The main parameters of the algorithm are as follows: cumulative window radius $r = 4$ and cumulative frame length $M = 4$. In Figure 5, (a) shows the difference graph and the corresponding 3D graph obtained by using the improved anisotropic method; (b) shows the image enhanced from (a) by using the original high-order cumulates (HOC) method and its corresponding 3D graph; (c) shows the image enhanced from (a) by using the improved high-order cumulates (IHOC) method and its corresponding 3D graph. The target position has been marked with a red rectangle. Table 6 describes the AGV and LSNR of the image after the original high-order cumulates and the improved high-order cumulates methods are used, respectively. It can be seen that the original and the improved high-order cumulates methods both enhance the dim and small target. On the whole, the improved high-order cumulates method provides a better enhancement effect as its image SNR improves more obviously and its overall performance is better.

7.3. Detection Results and Analysis

7.3.1. Parameter Selection Analysis. The detection effect of the local energy center of sequential image is related to the length N of the cumulative frame and the cumulative neighborhood size R . Figure 6 shows the relationship curve between the detection rate (Pd) and the cumulative frame length N . It can be seen that the target detection rate is the highest when

TABLE 3: MSE comparison between various background prediction methods.

MSE	1	2	3	4	5	6
Top-Hat	164.88	166.64	168.76	164.54	164.74	165.73
TDLMS [14]	116.24	119.84	165.18	116.73	118.52	118.89
NPB [15]	50.67	54.88	63.14	50.86	51.32	52.63
ABP [16]	11.45	13.12	13.75	11.78	12.16	12.64
<i>IABP</i>	9.28	9.87	10.43	9.59	9.62	9.79
SG [17]	14.65	15.93	18.71	14.76	15.22	15.88
FRA [18]	12.53	13.57	14.17	12.86	13.45	13.36
MoG [19]	13.24	14.55	16.34	13.54	14.12	14.35

TABLE 4: SSIM comparison between various background prediction methods.

SSIM	1	2	3	4	5	6
Top-Hat	0.575	0.592	0.485	0.563	0.552	0.567
TDLMS [14]	0.694	0.518	0.507	0.652	0.546	0.539
NPB [15]	0.769	0.642	0.635	0.726	0.686	0.654
ABP [16]	0.969	0.925	0.889	0.958	0.942	0.937
<i>IABP</i>	0.979	0.935	0.932	0.968	0.956	0.943
SG [17]	0.949	0.879	0.753	0.936	0.929	0.913
FRA [18]	0.968	0.915	0.885	0.954	0.935	0.921
MoG [19]	0.957	0.893	0.873	0.942	0.923	0.915

TABLE 5: GSNR comparison between various background prediction methods.

	Top-Hat	TDLMS [14]	NPB [15]	ABP [16]	<i>IABP</i>	SG [17]	FRA [18]	MoG [19]
GSNR	5.24	6.43	7.32	9.16	10.75	8.11	8.72	8.34

TABLE 6: Comparison between the enhancement effects of the original high-order cumulates and the improved high-order cumulates.

Difference graph		HOC		IHOC	
AGV	LSNR	AGV	LSNR	AGV	LSNR
90	3.45	118	7.19	235	13.05

the frame is $N = 10 \sim 12$. This paper selects $N = 11$ for accumulation. The size of the cumulative neighborhood R is closely related to the size of the target, the cumulative frame length, and the target velocity. In these experiments, $R = 20$.

7.3.2. Results and Analysis. To verify the effectiveness of the proposed detection method for different scenes, three scenes A, B, and C with frame lengths of 85, 114, and 245, respectively, are selected for experiments. The target in scene A moves around a certain point randomly. The target in scene B is strongly mobile, which moves upward first and downward next, then suddenly accelerates to go upward obliquely, and

finally turns around and goes downward. The target in scene C just moves obliquely and downward in a uniformly accelerated rectilinear motion. Figure 7(a) is the first frame image of scenes A; (b) is a trajectory image sequence obtained by superimposing the binary images obtained by improving the anisotropic and the high-order accumulation methods; (c) is the trajectory image obtained by removing the noise from the local energy center of sequential image and superimposing the target detection results. Figure 8(a) is the first frame image of scenes B; (b) and (c) are images obtained by the same methods. Figure 9(a) is the first frame image of scenes C; (b) and (c) are images obtained by the same methods. It can be seen from Figures 7, 8, and 9 that using the local energy center of the sequence image can effectively eliminate the noise from the image and accurately detect the target.

In order to evaluate the detection performance of proposed detection method, the infrared images of 10 different sequences are compared by using the pipeline filtering method [7], Wu et al.'s method [20], and proposed method, respectively. The average SNR, number of sequence frames, the detection rate (Pd), and the false alarm rate (Pf) are

FIGURE 3: Continued.

FIGURE 3: From left to right, background graph, difference graph, and 3D graph corresponding to the difference graph of different background prediction methods.

described in Table 7. In Figure 10, the detection abilities of the three methods are compared. In (a), the abscissa is the average SNR of the image sequence, the ordinate is the detection rate, the circle indicates the detection rate of pipeline filtering, the

asterisk indicates the detection rate of Wu et al.'s method, and the plus sign represents the detection rate of the local energy center of image sequence. (b) is the curve fitting graph for (a), in which P is the detection rate fitting curve of the pipeline

FIGURE 4: Relationship between SNR gain and cumulative frame length M .

TABLE 7: Simulation result of each sequence.

Seq.	1	2	3	4	5	6	7	8	9	10
SNR	2.21	1.58	1.32	2.47	2.36	1.23	1.24	0.87	1.43	0.96
Frames	132	200	85	124	324	117	210	235	75	89
Pipeline filter [7]										
Pd	92.7%	91.6%	89.4%	94.7%	94.3%	89.6%	88.3%	77.5%	85.4%	87.3%
Pf	4.7%	5.03%	4.65%	4.89%	3.18%	5.97%	4.38%	10.42%	7.32%	6.45%
Wu et al.'s method [20]										
Pd	96.1%	95.6%	93.4%	96.7%	95.2%	90.8%	91.3%	82.5%	92.5%	88.2%
Pf	3.95%	4.52%	4.23%	2.57%	2.24%	5.64%	4.32%	7.89%	5.57%	5.89%
Proposed method										
Pd	97.6%	96.5%	95.3%	97.6%	98.4%	94.8%	95.4%	85.4%	94.8%	89.2%
Pf	2.78%	3.03%	3.55%	1.72%	1.43%	5.21%	4.24%	6.56%	4.32%	5.43%

filtering method, W is that of Wu et al.'s method, and E is that of proposed method. Figure 11 shows a comparison between the average SNR and the false alarm rates. The symbols in the figure are similar to those of Figure 10.

It can be seen from Table 7 that, under the same SNR, the detection rate of proposed method in this paper is the largest, and its false alarm rate is the lowest, followed by the proposed method in the literature [20]. As can be seen from Figures 10 and 11, the detection rates of the three methods increase as the SNR increases, while the false alarm rate decreases as the SNR increases. For a dim and small target with local SNR less than 2.5 dB in the infrared sequence image, the proposed method can detect it well, and the detection rate is obviously improved and the false alarm rate is reduced compared with the other two methods under the same SNR.

8. Conclusions

In order to improve the detection and recognition ability of small targets in images, this paper first uses the improved

anisotropy to predict the background and then adopts the improved high-order cumulates to enhance the target, and finally, on the basis of image background suppression and target enhancement, this paper proposes a new motion feature of local energy center of sequential image as a multiframe motion correlation detection algorithm. The algorithm does not need to predict a priori knowledge in advance, such as the motion velocity and direction. In addition, this method requires no excessive restrictions. Thus, compared to traditional methods, this method has a wider range of applications and is more in line with the needs of the actual infrared scene. The simulation experiment shows the following:

(1) Overall performance of the improved anisotropy is better than other background prediction methods. For different SNR images, MSEs of the improved anisotropy are all less than 10.45, and MSE is lower when SNR is higher. SSIM of the improved anisotropy are all greater than 0.93 for different SNR images. For low SNR images, such as

FIGURE 5: Results before and after enhancement and corresponding 3D views.

$SNR = 0.86$, SSIM also achieves good results, up to 0.932. GSNR achieves good results among the improved anisotropy, reaching 10.75.

(2) For the image without the background, the target energy is still very dim and there is noise interference, which is not easy to be accurately segmented and extracted. By reinforcing the high-order cumulate method and fully considering the energy cumulates in the neighborhood of the target's movement, the effect is better than the original

high-order cumulates. The average grayscale and the local SNR of the target are, respectively, 118 and 7.19 before the improvement, while those are, respectively, 235 and 13.05 after the improvement, which significantly improve image grayscale of small targets and the local SNR.

(3) This paper constructs a sequential image energy center detection algorithm that integrates the neighborhood, continuity, area, and energy and other motion characteristics of the target. The method proposed can better detect the dim

FIGURE 6: The relationship curve between the detection rate and the number of cumulative frames.

FIGURE 7: Detection results of image scene A with the proposed method.

FIGURE 8: Detection results of image scene B with the proposed method.

and small target in infrared sequential image, of which the local SNR is less than 2.5 db. Under the same SNR condition, the detection rate of the method is obviously improved and

the false alarm rate is reduced, compared with the pipeline filtering method. The method can detect a target with the lowest SNR of 0.86.

FIGURE 9: Detection results of image scene C with the proposed method.

FIGURE 10: Detection rate comparison diagram of the pipeline filter method, Wu et al.'s method, and the proposed method.

FIGURE 11: False alarm rate comparison diagram of the pipeline filter method, Wu et al.'s method, and the proposed method.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was partly supported by the Institute of Optics and Electronics Chinese Academy of Sciences, by the National Natural Science Foundation of China (61571096), and by Foundation of Key Laboratory of Beam Control Chinese Academy of Sciences (2014LBC002).

References

- [1] I. S. Reed, R. M. Gagliardi, and H. M. Shao, "Application of three-dimensional filtering to moving target detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 19, no. 6, pp. 898–905, 1983.
- [2] D. G. Falconer, "Target Tracking With The Hough Transform," in *Proceedings of the 11th Conference on Circuits, Systems and Computers*, pp. 249–252, Pacific Grove, CA, USA, November 1977.
- [3] L. A. Johnston and V. Krishnamurthy, "Performance analysis of a dynamic programming track before detect algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 1, pp. 228–242, 2002.
- [4] S. D. Blostein and T. S. Huang, "Detecting small, moving objects in image sequences using sequential hypothesis testing," *IEEE Transactions on Signal Processing*, vol. 39, no. 7, pp. 1611–1629, 1991.
- [5] R. J. Liou and M. R. Azimi-Sadjadi, "Multiple target detection using modified high order correlations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 2, pp. 553–568, 1998.
- [6] H. Y. Zhang, H. Duan, and M. H. Liao, "The TBD method for dim targets based on multi-level crossover and matching operator," *Journal of Harbin Institute of Technology*, vol. 18, no. 1, pp. 57–61, 2011.
- [7] G. Wang, R. M. Inigo, and E. S. McVey, "Pipeline algorithm for detection and tracking of pixel-sized target trajectories," *Signal and Data Processing of Small Targets*, vol. 1305, no. 1, pp. 167–178, 1990.
- [8] W. Wang and C. M. Li, "A robust infrared dim target detection method based on template filtering and saliency extraction," *Infrared Physics & Technology*, vol. 73, pp. 19–28, 2015.
- [9] X. P. Shao, H. Fan, and G. X. Lu, "An improved infrared dim and small target detection algorithm based on the contrast mechanism of human visual system," *Infrared Physics & Technology*, vol. 55, no. 5, pp. 403–408, 2012.
- [10] X. Wang, S. Q. Shen, and C. Ning, "A sparse representation-based method for infrared dim target detection under sea-sky background," *Infrared Physics & Technology*, vol. 71, pp. 347–355, 2015.
- [11] J. J. Zhao, Z. Y. Tang, and J. Yang, "Infrared small target detection using sparse representation," *Journal of Systems Engineering and Electronics*, vol. 22, no. 6, pp. 897–904, 2011.
- [12] D. Hu, J. J. Zhao, and Y. Cao, "Infrared small target detection based on saliency and principal component analysis," *Journal of infrared and millimeter waves*, vol. 29, no. 4, pp. 303–306, 2010.
- [13] J. J. Zhao, Z. Y. Tang, J. Yang et al., "Infrared small target detection based on image sparse representation," *Journal of Infrared and Millimeter Waves*, vol. 30, no. 2, pp. 156–161, 2011.
- [14] T. W. Bae, Y. C. Kim, S. H. Ahn et al., "An efficient two-dimensional least mean square (TDLMS) based on block statistics for small target detection," *Journal of Infrared, Millimeter, and Terahertz Waves*, vol. 30, no. 10, pp. 1092–1101, 2009.
- [15] Y. Liu, H. Yao, W. Gao, X. Chen, and D. Zhao, "Nonparametric background generation," *Journal of Visual Communication and Image Representation*, vol. 18, no. 3, pp. 253–263, 2007.
- [16] H. X. Zhou, Y. Zhao, H. L. Qin et al., "Infrared dim and small target detection algorithm based on multi-scale anisotropic diffusion equation," *Guangxi Xuebao/Acta Photonica Sinica*, vol. 44, no. 9, Article ID 0910002, pp. 26–32, 2015.
- [17] Y. Benzeeth, P. M. Jodoin, B. Emile, H. Laurent, and C. Rosenberger, "Review and evaluation of commonly-implemented background subtraction algorithms," in *Proceedings of the 19th International Conference on Pattern Recognition (ICPR '08)*, pp. 1–4, Tampa, Fla, USA, December 2008.
- [18] M. H. Sigari, N. Mozayani, and H. R. Pourreza, "Fuzzy running average and fuzzy background subtraction: concepts and application," *International Journal of Computer Science and Network Security*, vol. 8, no. 2, pp. 253–259, 2008.
- [19] T. Bouwmans, F. E. Baf, and B. Vachon, "Background Modeling using Mixture of Gaussians for Foreground Detection - A Survey," *Recent Patents on Computer Science*, vol. 1, no. 3, pp. 219–237, 2008.
- [20] B. Wu, H. B. Ji, and P. Li, "New method for moving dim target detection based on third order cumulate in infrared image," *Hongwai Yu Haomibo Xuebao/Journal of Infrared and Millimeter Waves*, vol. 25, no. 5, pp. 364–367, 2006.
- [21] Q. Zhang and J. J. Cai, "Small dim infrared targets segmentation method based on local maximum value partitioning," *Infrared Technology*, vol. 33, no. 1, pp. 124–130, 2011.
- [22] Y. Z. Liu, "Nonparametric Background Generation," *IEEE Transactions on Image Processing*, vol. 19, no. 7, pp. 25–35, 2010.
- [23] A. Sobrala and A. Vacavant, "A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos," *Computer Vision and Image Understanding*, vol. 122, pp. 4–21, 2014.
- [24] A. Sobral, "BGSLibrary: an opencv c++ background subtraction library," in *Proceedings of the IX Workshop de Viso Computacional (WVC '13)*, Rio de Janeiro, Brazil, June 2013, <http://code.google.com/p/bgslibrary/>.
- [25] S. G. Song, J. G. Wang, and Q. S. Chen, "Infrared dim and small target detection under sea and sky complex background," *Opto-Electronic Engineering*, vol. 32, no. 4, pp. 9–12, 2010.
- [26] Y. H. Wang and W. N. Liu, "Dim target enhancement algorithm for low-contrast image based on anisotropic diffusion," *Opto-Electronic Engineering*, vol. 35, no. 6, pp. 15–19, 2012.
- [27] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [28] Y. B. Tong, Q. S. Zhang, and Y. P. Qi, "Image quality assessing by combining PSNR with SSIM," *Journal of Image and Graphics*, vol. 11, no. 12, pp. 1758–1763 (Hungarian), 2006, (Chinese).
- [29] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [30] Q. Cao and D. Bi, "Characteristic-selecting filtering in infrared small target detection," *Guangxue Xuebao/Acta Optica Sinica*, vol. 29, no. 9, pp. 2408–2412, 2009 (Chinese).

Hindawi

Submit your manuscripts at
<https://www.hindawi.com>