Dim and Small Target Detection Based on Spiral Gradient Optimization Estimation and High-Order Correlation Enhancement

XIANGSUO FAN [1,2], **LEI MIN** [3], **JULIU LI** [2], **HONGWEI GUO** [4], **(Member, IEEE),**
LINPING FENG [5], **(Member, IEEE), AND ZHIYONG XU** [3]

[1]Faculty of Intelligence Manufacturing, Yibin University, Yibin, Sichuan 644000, China
[2]School of Electrical, Electronic and Computer Science, Guangxi University of Science and Technology, Liuzhou 545006, China
[3]Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, China
[4]School of Engineering, Honghe University, Mengzi 661100, China
[5]School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China

Corresponding authors: Lei Min (minlei1986@163.com) and Hongwei Guo (ghw202@foxmail.com)

ABSTRACT In order to reduce the influence of strong halo effect on dim small target detection in daytime, a dim small target detection algorithm based on spiral gradient optimization estimation and high-order correlation enhancement is proposed in this paper. In this paper, we first design a spiral motion model to obtain the local gradient information in the central image point by perturbing the designed motion direction, then estimate the optimal background by establishing a gradient optimization model to achieve background suppression while effectively removing the halo phenomenon. Considering the original high-order correlation model only uses a single pixel for motion correlation, there is insufficient information utilization, an improved high-order correlation energy enhancement model is proposed to enhance the target signal, the algorithm first constructs an attention discrimination model based on inner and outer windows to obtain the salient region of the image, and then carries out multi frame high-order motion correlation of neighborhood blocks to enhance the target energy. After experiments, it is shown that compared with the traditional algorithm, the algorithm proposed in this paper can effectively weaken the halo effect while suppressing most of the background and can effectively enhance the local signal-to-noise ratio of the target.

INDEX TERMS Dim and small target detection, background suppression, spiral gradient, high order correlation, energy enhancement.

I. INTRODUCTION

Dim and small target signal detection is of great importance for high altitude warning [1], [2]. These targets are isolated points in the image, occupy fewer image pixels, lack of relevant texture, and are often accompanied by atmospheric turbulence, clouds and various noise clutter around the target, which makes the target often submerged and brings difficulties to target detection [3]–[7]. For such detection scenarios, researchers have proposed many feasible detection algorithms. It includes anisotropic filtering [8], [9], top hat filtering [3], [10], [11], Max mean filtering [12],

The associate editor coordinating the review of this manuscript and approving it for publication was Mira Naftaly [ID].

spatiotemporal significance model [13] and greedy bilateral factorization model [14], as well as detection methods constructed according to the low rank characteristics of image background, such as IPI model [15], RPCA model [16], [17], MPCM model [18], TV-PCP model [18], LCM model [20], DPA model [21], CDAE model [22] and other algorithms have contributed to the detection of dim and small targets. For example, the spatiotemporal saliency model [13] and the greedy bilateral factorization model [14] proposed by Pang *et al.* Literature [13] fully fuse the time domain information and spatial domain information of the image to complete the background modeling of the image, and obtain the saliency region in the image, so as to retain and enhance the target signal. Finally, using the motion characteristics

of target to detect and extract real target through adaptive threshold segmentation, and the effect of background modeling in low altitude and long-distance scene is better. Another example is the real-time controlled weak and small target detection algorithm [21] proposed by sun *et al.* By using the processing method of tracking before detection, the image is accumulated to obtain a priori information to extract the target, and then the image is accumulated through the continuous time domain characteristics of multiple frames of images to enhance the signal of target, and then the edge information is fused by constructing a new adaptive model, Finally, the background modeling of the image is completed by combining the prior information with the adaptive model, which not only suppresses the background, but also enhances the signal value of the target to achieve the effect of target detection. In addition, there are also low rank and sparse decomposition model detection algorithm based on regularization [23]–[25] and dim small target detection algorithm based on deep neural network [26]–[29]. For instance, the background suppression model of spatial multiple subspace learning and spatial-temporal arrangement tensor proposed by sun *et al.* [23], firstly, the STIPT model is constructed by using the spatial-temporal information of sequence images and tensor to obtain the relevant features of the image to highlight the target characteristics, and then the learning strategy of multiple subspace learning is extended to the constructed tensor model to complete the background suppression of the image, the WSNM model is also combined to make the tensor avoid the constraint of pixel matrix dictionary, so as to better retain the low rank components of the image. Finally, the overall framework is optimized by constructing a new tensor singular value decomposition and alternating direction multiplier model (ADMM), which ensures that the target components are recovered more accurately and efficiently. For another example, the RISTD deep learning network model [28] proposed by Hou *et al.* Constructs the image feature extraction framework through the combination of manual features and convolution neural network to train the image scene, expose the prominent features of the image and improve the robustness of the model. Then, a mapping network is constructed based on the similarity between the feature image and the target to determine the final target. Finally, the target can be extracted by appropriate threshold segmentation. The above algorithms have good detection effect in scenes with high image signal-to-noise ratio, which mainly use scene data to estimate model parameters or train models. When the scene changes dynamically, the model parameters will change, and the model needs to be optimized. However, due to high-altitude and long-distance imaging, the image background changes dynamically and is accompanied by strong light. When the target moves near the strong light position or along the backlight, the image signal-to-noise ratio is less than 0dB, and the target information is completely submerged by the strong light halo, so the above detection algorithm is greatly limited, simple background modeling of the image can not well separate the target and

background, resulting in poor background modeling effect and failure of target detection. Therefore, it is necessary to put forward new research ideas to process the image with strong light background. In these scenes, the halo effect has a very serious impact on target detection [30]–[32]. In order to reduce the impact of halo on weak and small target detection, the camera halo in the image must be corrected to eliminate the optical vignetting effect, so that the output image signal becomes uniform, and then the next step of detection can be carried out. There are feasible methods to correct the optical vignetting effect of the image, and the effect is good, which can effectively correct the halo. For example, the infrared system vignetting effect correction method based on polynomial approximation proposed by Li [30], first use the square difference of the image to extract the image background, Then the correction factor of the image is obtained by polynomial approximation method to correct the halo. However, in the face of dynamic scene, because the image variance will change with the scene, the polynomial parameters fitted by image variance will also change, which affects the halo correction to a certain extent. Another example is the nonlinear compensation method for optical vignetting effect of the visual system proposed by Ding *et al.* [32], which uses the characteristics of camera exposure combined with coupled sampling and decoupled reconstruction to solve the response function of the camera, finds the irradiance of the image element through exposure time and image element illumination to build the relative exposure model, then builds the vignetting illumination compensation function through the relative exposure, and finally builds the camera response function through the correspondence between camera exposure and image grayscale. The algorithm will be converted to the use of camera illumination to correct only by the position of the image element, which can better correct the halo, but the method requires high control of exposure time by analyzing the exposure of the image element, otherwise it will cause overexposure phenomenon and not benefit in target detection.

In the strong light background, even after vignetting correction preprocessing, the target signal is still weak. In order to accurately extract the target signal, the common processing method is to model the background of the preprocessed image to obtain the background image and differential image, and then threshold segment the differential image to select candidate target points, and multi frame accumulation is carried out to enhance the target energy. However, this kind of algorithm only uses the spatial information of the image to process the image when modeling the background, which often weakens the signal of the target and suppresses the target as the background. When facing the target detection with high light intensity in the daytime, this kind of algorithm is obviously not applicable. In order to detect the target signal smoothly, researchers make full use of these such as air domain information and time domain information of the image, such as the algorithm proposed by Jiang *et al.* for the application of hidden Markov model in multi-target

automatic detection and tracking [33], [34], which predicts the motion of the target by calculating the transfer probability of the target at the current moment and the next moment, and then determines the detected target by the forward-backward local optimal algorithm, which combines The time-domain and space-domain information of the image is fully combined, but because the Markov random field model mainly deals with single image element, it has great limitations when facing images where the strongly illuminated target is not visible. The improved high-order correlation target energy accumulation algorithm [35] proposed by Yong, which uses the inter frame correlation and time-domain continuity of the target in the sequence image to complete the target detection and extraction, and improves the target energy through the iterative accumulation operation between adjacent frames, so as to provide good conditions for subsequent target detection and extraction.

In summary, when detecting small targets in the daytime strong illumination scene, it is first necessary to correct the image halo to remove the impact of the strong halo on subsequent target detection, and then carry out background modeling and energy enhancement on the preprocessed image. In the paper, a spiral gradient optimal estimation model is proposed to suppress the background of the image. By constructing a new spiral gradient model, the background modeling of the image can be completed while correcting the image halo. Secondly, the improved high-order correlation energy enhancement model is used to enhance the target signal in the difference image to ensure that the target is accurately extracted. See sections 2 and 3 for detailed algorithm introduction.

II. DIM SIGNAL DETECTION ALGORITHM BASED ON SPIRAL GRADIENT OPTIMIZATION ESTIMATION

In the detection of dim and small targets in the daytime, with the change of incident light angle, the energy distribution on the target surface of the detector changes, resulting in uneven image brightness distribution. Because most of the light energy is distributed in the center of the image in the transmission process, the light energy is similar to the point diffusion energy distribution form, which diffuses from the center to both sides, and the central halo phenomenon is obvious. Therefore, this paper proposes a spiral gradient optimization estimation of dim signal detection algorithm, based on the pre-designed spiral motion model for perturbation to obtain local gradient information in a neighborhood of an image element, and then rely on the gradient optimization model to estimate the local optimal background to achieve background suppression while effectively removing the center of the halo, which greatly improves the detection ability of subsequent targets.

A. SPIRAL MOTION MODEL

(1) Firstly, set a spiral motion model, first number its coordinates with a pixel point as the center, and store the number in

an array, which is recorded as $Data_{2 \times 25} = \begin{bmatrix} 0 & 0 & \cdots & 2 \\ 0 & 1 & \cdots & 2 \end{bmatrix}$, show as fig 1.
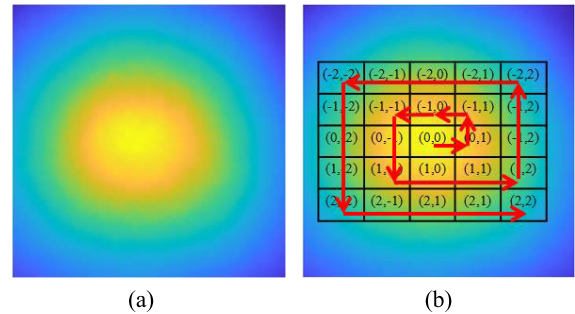

FIGURE 1. Spiral motion diagram. (a) center of the halo; (b) motion model.

Record the sequence number of the array as $k(k = 1, 2, \cdots, 25)$. Firstly, the moving coordinate information of the central pixel (row, col) is obtained by solving the residual function, Then take the newly moved coordinate center pixel as the center and take out the data block with width and height of $w \times h$. The specific formula is as follows:

$$\begin{cases} num = \mathrm{mod}(k - 1, 25) + 1 \\ row = x_0 + Data\,(1, num) \times step \\ col = y_0 + Data\,(2, num) \times step \\ f_k = f(row - h/2 : row + h/2, col - w/2 : col + w/2) \end{cases}$$
(1)

where num is the number of the array; (row, col) is the newly acquired move coordinate; $step$ is the step size of the central image coordinate point, f_k indicates the newly acquired k-th$(k = 1, 2, \cdots, 25))$ data block, with width and height of $w \times h$.

B. GRADIENT OPTIMASATION ESTIMATION MODEL

Calculating the gradient difference between the central pixel data block and the newly obtained data block f_k, then obtain the gradient minimum value of the neighborhood block of spiral motion by constructing the gradient optimization model, and record the center point coordinates corresponding to the gradient minimum value as (x_0', y_0'). The specific formula is as follows:

$$\begin{cases} F_k(row, col) = \dfrac{1}{L * L} \displaystyle\sum_{m=-fix(L/2)}^{m=fix(L/2)} \sum_{n=-fix(L/2)}^{n=fix(L/2)} \\ \qquad\qquad\qquad\qquad \times f(x_0 + m, y_0 + n) \\ \qquad - \dfrac{1}{L * L} \displaystyle\sum_{m=-fix(L/2)}^{m=fix(L/2)} \sum_{n=-fix(L/2)}^{n=fix(L/2)} f_k(row + m, col + n) \\ [F_{\min 1}, x_0', y_0'] = \min_{k=1,2,\cdots 25} (F_1(row, col), \cdots, F_k \\ \qquad\qquad (row, col), \cdots, F_{25}(row, col)) \\ F_{\min 1} = \dfrac{1}{L * L} \displaystyle\sum_{m=-fix(L/2)}^{m=fix(L/2)} \sum_{n=-fix(L/2)}^{n=fix(L/2)} f(x_0' + m, y_0' + n) \end{cases}$$
(2)

in formula the $L * L$ is the range size of the neighborhood value is $3 * 3, 5 * 5, 7 * 7, \cdots$; $F_k(row, col)$ is the gradient difference between the central image metadata block and the newly acquired data block f_k; (x'_0, y'_0) is the coordinate mark of the center point corresponding to the minimum gradient, $F_{\min 1}$ is the mean value of the data block of the central image element (x'_0, y'_0).

Substitute the newly obtained center point coordinates (x'_0, y'_0) into formula (1) to obtain a new data block f'_k, as follows:

$$
\begin{cases}
num = \mod(k - 1, 25) + 1 \\
row_1 = x'_0 + Data\,(1, num) \times step \\
col_1 = y'_0 + Data\,(2, num) \times step \\
f'_k = f(row_1 - h/2 : row_1 + h/2, col_1 - w/2 : col_1 + w/2)
\end{cases}
\tag{3}
$$

Immediately afterwards, equation (2) is used to obtain the gradient minimum of the neighborhood block with (x'_0, y'_0) as the central coordinate point for the spiral motion with the help of the gradient optimization model, and the central point corresponding to the gradient minimum is sat labeled as (x''_0, y''_0), the specific formula is as follows:

$$
\begin{cases}
F'_k(row_1, col_1) = \dfrac{1}{L * L} \displaystyle\sum_{m=-fix(L/2)}^{m=fix(L/2)} \sum_{n=-fix(L/2)}^{n=fix(L/2)} \\
\qquad\qquad\qquad \times f(x'_0 + m, y'_0 + n) \\
\quad - \dfrac{1}{L * L} \displaystyle\sum_{m=-fix(L/2)}^{m=fix(L/2)} \sum_{n=-fix(L/2)}^{n=fix(L/2)} f_k(row_1 + m, col_1 + n) \\
[F_{\min 2}, x''_0, y''_0] = \min_{k=1,2,\cdots 25} \big(F'_1(row_1, col_1), \cdots, \\
\qquad\qquad F'_k(row_1, col_1), \cdots, F'_{25}(row_1, col_1)\big) \\
F_{\min 2} = \dfrac{1}{L * L} \displaystyle\sum_{m=-fix(L/2)}^{m=fix(L/2)} \sum_{n=-fix(L/2)}^{n=fix(L/2)} f(x''_0 + m, y''_0 + n)
\end{cases}
\tag{4}
$$

Finally, a gradient-optimized background estimation model is constructed as follows:

$$
f'(i, j) = \begin{cases}
\dfrac{1}{L * L} \displaystyle\sum_{m=-fix(L/2)}^{m=fix(L/2)} \sum_{n=-fix(L/2)}^{n=fix(L/2)} f(x'_0 + m, y'_0 + n) \\
\quad \text{if } F_{\min 1} \leq F_{\min 2} \\
\dfrac{1}{L * L} \displaystyle\sum_{m=-fix(L/2)}^{m=fix(L/2)} \sum_{n=-fix(L/2)}^{n=fix(L/2)} f(x''_0 + m, y''_0 + n) \\
\quad \text{else } F_{\min 1} > F_{\min 2}
\end{cases}
\tag{5}
$$

III. IMPROVED HIGH ORDER CORRELATION ENERGY ENHANCEMENT MODEL

A. HIGHER ORDER CORRELATION THEROY

In continuous multi frame images, the target moves along a certain direction, so the target has a certain correlation between adjacent frames [35]. The high-order correlation

detection model mainly uses the inter frame correlation of the image to enhance the signal of the target. The target can be further highlighted through the accumulation before and after the frame. The corresponding calculation formula is as follows [35]:

$$
Y(x, y, t_n) = \left[\sum_{i=-|v|}^{|v|} \sum_{j=-|v|}^{|v|} f(x, y, t_n) f(x + i, y + j, t_{n+1}) \right]
\tag{6}
$$

where $Y(x, y, t_n)$ represents the correlation values of time t_n and time t_{n+1}, $f(x, y, t_n)$ and $f(x + i, y + j, t_{n+1})$ represent the image element value at the corresponding moment in the image, respectively, x, y indicates the coordinates of the image element in the image, v indicates the moving neighborhood radius of the target. The first-order correlation of two frames of images is obtained by summing the product of pixels at the front and back times. Using this idea, the correlation of multiple frames of images can be extended to enhance the target point signal. The specific calculation formula is as follows [35]:

$$
\begin{aligned}
&Y^{(1)}(x, y, t_n) \\
&= \left[\sum_{i=-|v|}^{|v|} \sum_{j=-|v|}^{|v|} Y^{(0)}(x, y, t_n) Y^{(0)}(x + i, y + j, t_{n+1}) \right]
\end{aligned}
\tag{7}
$$

in formula the $Y^{(0)}(x, y, t_n) = f(x, y, t_n)$, $Y^{(1)}(x, y, t_n)$ Represents the first-order correlation function of two adjacent frames. In order to increase the energy of the target more obviously. Continuing the accumulation of $Y^{(1)}(x, y, t_n)$ with the next image frame means completing the second order operation of higher order correlation, the specific mathematical formula is as follows:

$$
\begin{aligned}
&Y^{(2)}(x, y, t_n) \\
&= \left[\sum_{i=-|v|}^{|v|} \sum_{j=-|v|}^{|v|} Y^{(1)}(x, y, t_n) Y^{(1)}(x + i, y + j, t_{n+1}) \right]
\end{aligned}
\tag{8}
$$

where $Y^{(2)}(x, y, t_n)$ represents the second-order correlation value, and v represents the moving neighborhood radius of the target.

B. IMPROVED HIGH ORDER CORRELATION MODEL

Due to the dim small target in the daytime in the bright light environment, there is a serious lack of target information, which is not conducive to the detection of the target, and there is more noise in the bright light environment, and the differential image in the image contains a lot of noise outside the out target, which is not conducive to the identification of the target. In view of this situation, this paper uses the improved high-order correlation energy accumulation model to enhance the image target signal, but the previous high-order correlation accumulation model mainly accumulates a single pixel in the front and rear frames of the image, which makes insufficient use of the information retained in

the target signal. In this paper, the attention discrimination model of the inner and outer windows is constructed by using the gradient difference characteristics between the target and the background area to obtain the significant area with large regional gradient difference, and then the high-order motion correlation of the significant area block is carried out by using the motion correlation of the target between adjacent frames. The specific calculation formula is as follows:

$$
\begin{cases}
k = round(l/2) \\
k_1 = floor(l_1/2) \\
A = f(m_1 - (k-1) : m_1 + (k-1), n_1 - (k-1) : \\
\qquad n_1 + (k-1), t) \\
B = f(m_1 - k_1 : m_1 + k_1, n_1 - k_1 : n_1 + k_1, t) \\
A\_value = \dfrac{\sum\limits_{m=-k}^{m=k} \sum\limits_{n=-k}^{n=k} A - \sum\limits_{m=-k_1}^{m=k_1} \sum\limits_{n=-k_1}^{n=k_1} B}{\left[(2(k-1)+1)^2 - (2(k_1-1)+1)^2 \right]} \\
B\_value = \dfrac{\sum\limits_{m=-k_1}^{m=k_1} \sum\limits_{n=-k_1}^{n=k_1} B}{\left[(2(k_1-1)+1)^2 \right]} \\
D\_value = |A\_value - B\_value|
\end{cases} \quad (9)
$$

where k and k_1 represent the radius of the inner and outer windows respectively; l and l_1 are the sizes of the inner and outer windows, $round$ and $floor$ are rounding functions up and down respectively; A, B denote the attention areas of the inner and outer windows, respectively. f represents the difference image at time t, (m_1, n_1) is the image element coordinate, A_value, B_value is the mean value of attention area of inner and outer windows respectively, D_value is the gradient difference of attention area between inner and outer windows.

The attention discrimination model is constructed by the gradient difference between the inner and outer windows in above formula. When the gradient difference between the inner and outer windows is compared with the preset threshold Th, if the gradient difference between the inner and outer windows is greater than the preset Th, the region is regarded as the significant region, and then the significant region block is subjected to high-order motion correlation by using the motion correlation of the target between adjacent frames. To enhance the target energy, the detailed attention discrimination model is as follows:

$$
\begin{cases}
result\_A = f(m_1 - 1 : m_1 + 1, n_1 - 1 : n_1 + 1, t) \\
result\_A_1 = f_1(m_1 - 1 : m_1 + 1, n_1 - 1 : n_1 + 1, t+1) \\
if \ D\_value \le Th \\
Y^{(1)}(m_1, n_1, t) = f(m_1, n_1, t) \\
else \ Y^{(1)}(m_1, n_1, t) = \dfrac{\left[\sum\limits_{m=-L}^{m=L} \sum\limits_{n=-L}^{n=L} result\_A \times result\_A_1 \right]}{L \times L}
\end{cases} \quad (10)
$$

where Th is a preset threshold, f_1 denotes the difference image at moment $t+1$, $result\_A$ and $result\_A_1$ represents the

neighborhood block with (m_1, n_1) as the central coordinate at time t and time $t+1$, respectively, and L represents the radius of the neighborhood block. $Y^{(1)}(m_1, n_1, t)$ represents the enhanced image obtained after high-order motion correlation of adjacent frame images, which is regarded as the first-order correlation image. Other variables are defined as above.

In order to further enhance the target energy, the paper also redefines the second-order higher-order correlation cumulative equation on the basis of the first-order correlation image in order to maximize the enhancement of the target signal with the following mathematical calculation model.

$$
\begin{cases}
result\_A_2 = f_2(m_1 - 1 : m_1 + 1, n_1 - 1 : n_1 + 1, t+2) \\
\\
if \ D\_value \le Th \\
Y^{(1)}(m_1, n_1, t+1) = f(m_1, n_1, t+1) \\
else \ Y^{(2)}(m_1, n_1, t+1) \\
= \dfrac{\left[Y^{(1)}(m_1, n_1, t) \times \sum\limits_{m=-L}^{m=L} \sum\limits_{n=-L}^{n=L} result\_A_1 \times result\_A_2 \right]}{L \times L}
\end{cases} \quad (11)
$$

where $Y^{(2)}(m_1, n_1, t+1)$ is the second-order correlation image, which is obtained by performing high-order operation again using the first-order correlation image and the next frame image on the basis of obtaining the significant region by the first-order correlation operation, $result\_A_2$ is the neighborhood block with (m_1, n_1) as the center coordinate at the moment $t+1$, $Y^{(1)}(m_1, n_1, t)$ is the first-order correlation image, and $result\_A_1$ represents the neighborhood block of the central coordinate at the current t time.

C. THIS PAPER SUMMARIZS THE ALGORITHM
In this paper, a spiral gradient optimization model is designed to estimate the optimal background and effectively remove the halo phenomenon while suppressing the background. Then the difference image is obtained by differential processing between the original image and the background image. Finally, the improved high-order correlation energy enhancement model is used to enhance the significant region (target signal) in the difference image, and the enhanced image is segmented and extracted to obtain the target point. The specific flow chart of the algorithm in this paper is shown in Figure 2.

IV. ANALYSIS OF TEST RESULT
In order to further evaluate the detection effect of this paper, four groups of real scene data are used to experiment the dim and small targets of sequence images. The sequence images of these four groups of real scenes are stored in 16 bit bytes, and the image format is BMP. The four groups of real scenes are the daytime detection of long-distance star targets by photoelectric detection system. The targets in the scene are near the center of the image and stationary. Due to the influence of strong light background and uneven energy distribution
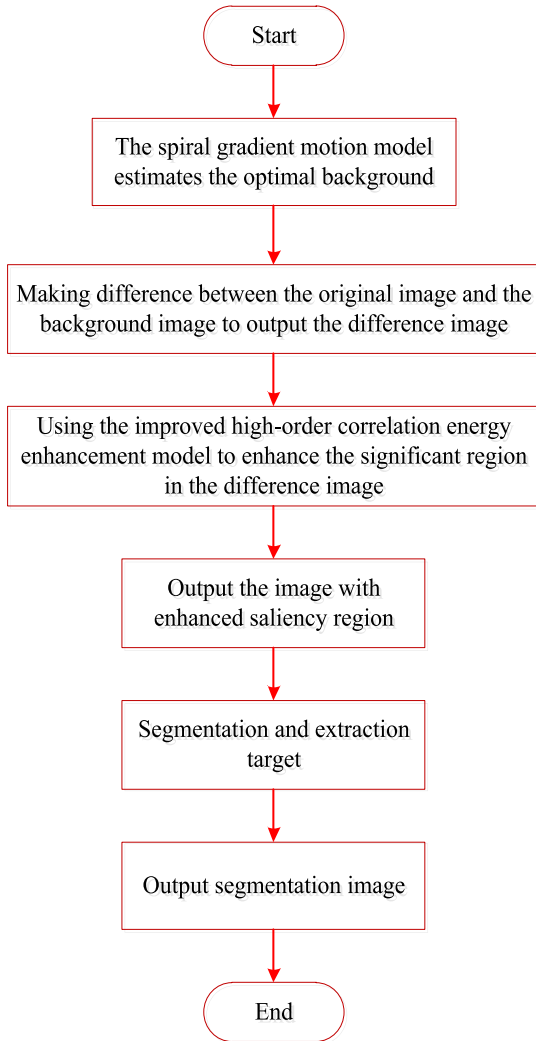
Start

The spiral gradient motion model estimates the optimal background

Making difference between the original image and the background image to output the difference image

Using the improved high-order correlation energy enhancement model to enhance the significant region in the difference image

Output the image with enhanced saliency region

Segmentation and extraction target

Output segmentation image

End

FIGURE 2. Algorithm flow chart of this paper.

TABLE 1. Attributes of each scene.

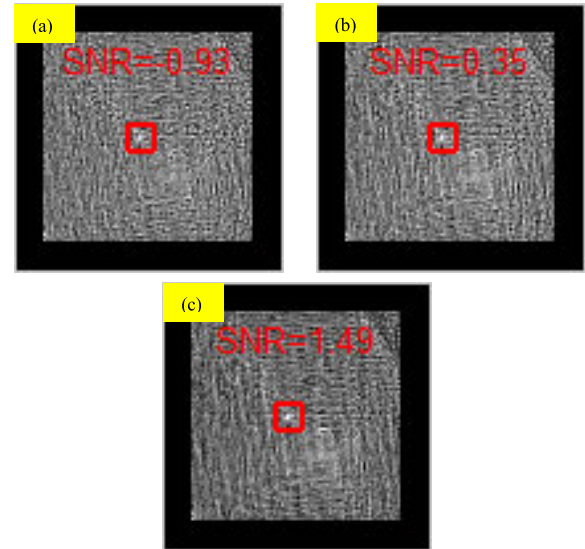| Scene	Average SNR	Frame length	Image size	Target size
Scene 1	-6.42dB	500	256x256	2x2
Scene 2	-6.41dB	500	256x256	2x2
Scene 3	-8.57dB	500	256x256	2x2
Scene 4	-8.36dB	500	256x256	3x3

FIGURE 3. Difference diagram of different frame number accumulated by this algorithm. (a) Accumulate 10 frames; (b) Accumulate 20 frames; (c) Accumulate 50.

of light field, the image has strong halo effect. The targets in the four groups of scenes are affected by strong halo, which brings great difficulties to target detection, Therefore, it is necessary to remove the halo effect of the image before detecting the target. At the same time, four different algorithms are used for comparative analysis, including weak and small target detection based on nonuniformity correction and multidirectional gradient [31], two-point correction detection algorithm [32] and high-order operation correlation detection algorithm [35]. The average signal-to-noise ratio and image attributes of each scene are shown in Table 1.

A. DIFFERENT CUMULATIVE FRAME DIFFERENCE DIAGRAMS

In this paper, the algorithm of this paper is used to evaluate the accumulation effect of different accumulation frames on the target. From the difference results after accumulation, the local signal-to-noise ratio is effectively improved with the increase of the accumulation frames, and the local signal-to-noise ratio is −0.93dB when accumulating 10 frames, 0.35dB when accumulating 20 frames, and 1.49dB when

accumulating 50 frames, see Figure 3, which strongly This strongly indicates that the local signal-to-noise ratio of the target can be enhanced by multiframe accumulation, thus favorably improving the detection capability of the target. In the actual field observation, it is also necessary to choose the appropriate number of accumulated frames for detection by compromising the execution efficiency.

B. DETECTION RESULT OF DIFFRERNT SCENES

The above four scenes are used to obtain the detection result plots of dim small and weak target detection based on non-uniform correction with multi-directional gradient [24], two-point correction detection algorithm [23] and high-order running correlation detection algorithm [32] and this paper's algorithm, respectively, from the detection result plots, it can be seen that since this paper's algorithm uses a spiral motion model, the local gradient information within the central image element point is obtained by perturbation based on the designed motion direction. Then the local optimal background is estimated by establishing a gradient optimization model to achieve background suppression while effectively removing the halo center, see Figure 4-Figure 7.

From the above Figures 3-6, it can be seen that the effect of the differential map obtained by the algorithm in this paper is better than that of other algorithms, and
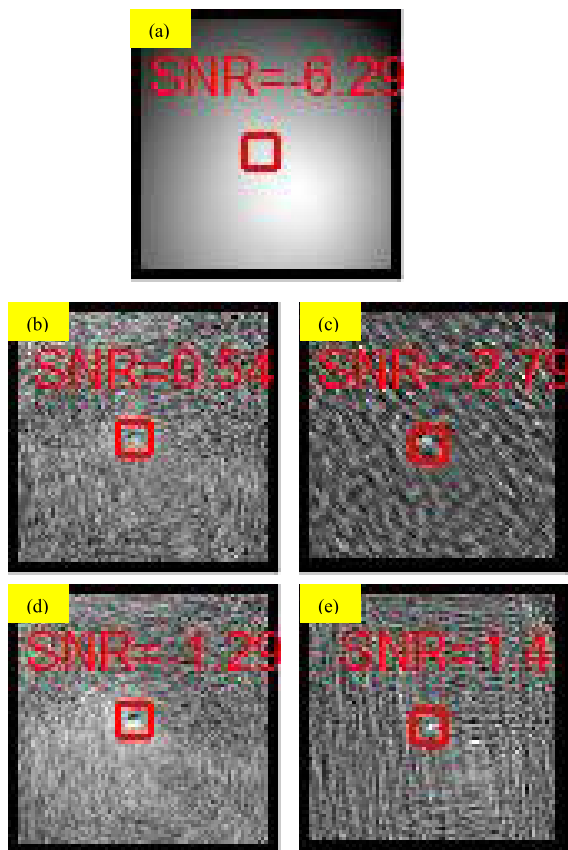
FIGURE 4. Detection results of different algorithms. (a) original image; (b) non-uniform corrected multi-directional gradient difference map signal-to-noise ratio; (c) traditional high-order correlation model difference map noise ratio; (d) two-point corrected model difference map signal-to-noise ratio; (e) algorithm difference map signal-to-noise ratio in this paper.

FIGURE 5. Detection results of different algorithms. (a) original image; (b) non-uniform corrected multi-directional gradient difference map signal-to-noise ratio; (c) traditional high-order correlation model difference map noise ratio; (d) two-point corrected model difference map signal-to-noise ratio; (e) algorithm difference map signal-to-noise ratio in this paper.

the background image elements obtained in the differential image are smoother and the target signal is more obvious. In contrast, the conventional algorithms proposed in the literature [23], [24], [32] rely on the scene data to estimate the correction parameters, and when the scene is in dynamic change, the inherent correction parameters will be difficult to meet the demand of dynamic correction, and the obtained differential images also have more streaking information and the target energy is weakened. The algorithm in this paper relies on the pre-designed spiral motion model to obtain the local gradient information of the neighborhood image, and then estimates the local optimal background according to the gradient optimization model. It does not need to estimate the correction parameters of the model, and the fringe information of the image can be smoothed after the spiral motion, which can meet the correction requirements of the dynamic scene.

Since the signal-to-noise of the target in the image is still relatively low even after spiral gradient estimation under strong daylight illumination, which is not conducive to target detection and extraction, this paper proposes an improved higher-order correlation energy accumulation model to enhance the image target signal, which mainly uses the motion correlation of the target between adjacent frames
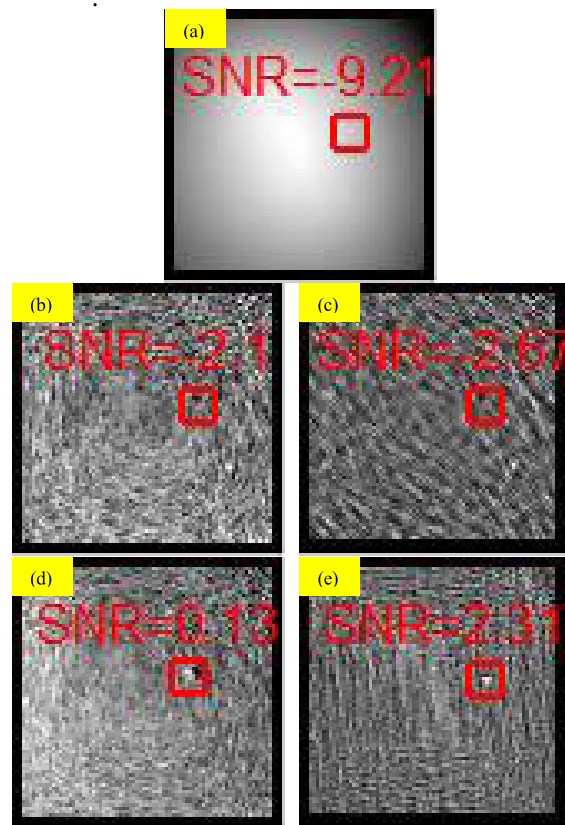
to perform higher-order motion correlation for significant blocks of regions. In order to intuitively reflect the effect of the proposed algorithm on target enhancement, two scenes are selected to show the effect before and after target enhancement. Fig 8(A) and (A1) are the difference images and 3D plots acquired after the spiral gradient estimation in this paper is adopted for scene 1; Fig 8(B) and (A1) are the difference images and 3D plots acquired after the improved higher-order correlation energy accumulation model is adopted for scene 1 to enhance. Fig 8(C) and (C1) are the difference images and three-dimensional plots obtained after adopting the spiral gradient estimation in this paper for scene 2; Fig 8(D) and (D1) are the difference images and three-dimensional plots obtained after adopting the improved higher-order correlation energy accumulation model for scene 2 to enhance, and the specific effects are shown in Figure 8 below.

As can be observed from Fig 8 above, after the improved higher-order correlation energy enhancement model proposed in this paper, a significant enhancement of the target energy can be observed from the signal-to-noise ratio of the images and the 3D plots before and after enhancement. In order to further illustrate the effect of the algorithm in this paper, several scenes of data are selected in the paper to quantitatively evaluate the differences before and after

TABLE 2. Signal-to-noise ratio of images before and after enhancement in 4 scenes.

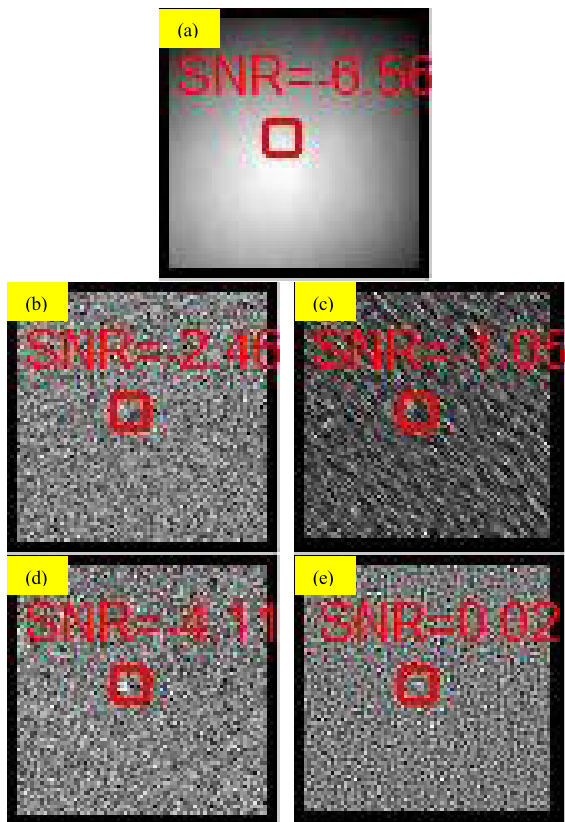| Scene\Index	Target signal	SNR value				
Scene 1	Before	1.4000	1.3700	1.4400	1.2800	1.4000
	After	**2.0200**	**2.0600**	**2.0900**	**2.9800**	**2.1200**
Scene 2	Before	2.3100	2.2900	2.4100	2.3600	2.3000
	After	**2.3900**	**2.3500**	**2.6900**	**2.6300**	**2.5700**
Scene 3	Before	0.0200	-0.0400	0.0100	-0.0700	-0.0200
	After	**0.9300**	**0.8700**	**0.9800**	**0.9600**	**0.9100**
Scene 4	Before	-1.2400	-1.2600	-1.4100	-1.7500	-1.7200
	After	**1.0400**	**0.9000**	**0.9500**	**0.2300**	**0.2600**

FIGURE 6. Detection results of different algorithms. (a) original image; (b) non-uniform corrected multi-directional gradient difference map signal-to-noise ratio; (c) traditional high-order correlation model difference map noise ratio; (d) two-point corrected model difference map signal-to-noise ratio; (e) algorithm difference map signal-to-noise ratio in this paper.

FIGURE 7. Detection results of different algorithms. (a) original image; (b) non-uniform corrected multi-directional gradient difference map signal-to-noise ratio; (c) traditional high-order correlation model difference map noise ratio; (d) two-point corrected model difference map signal-to-noise ratio; (e) algorithm difference map signal-to-noise ratio in this paper.

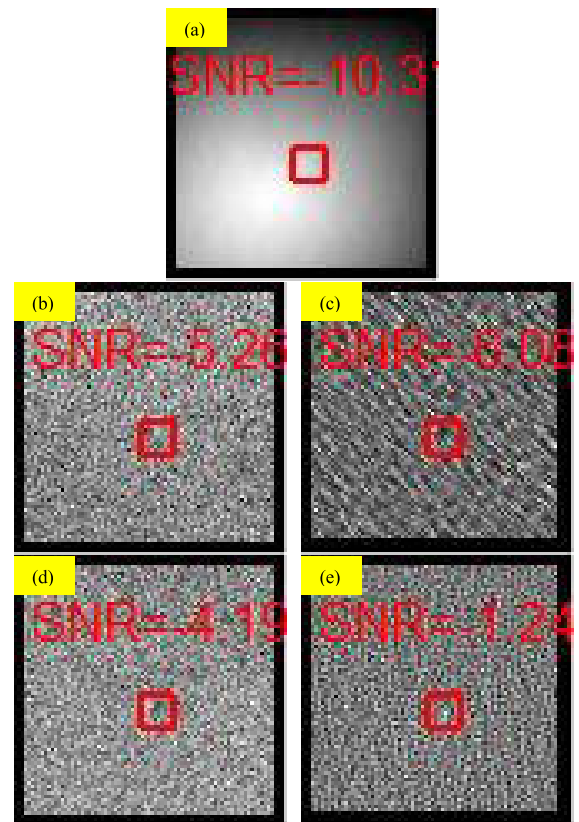enhancement using the image signal-to-noise ratio, as shown in Table 2.

As shown in the image signal-to-noise ratio in the table 2, it can be clearly observed that the algorithm in this paper
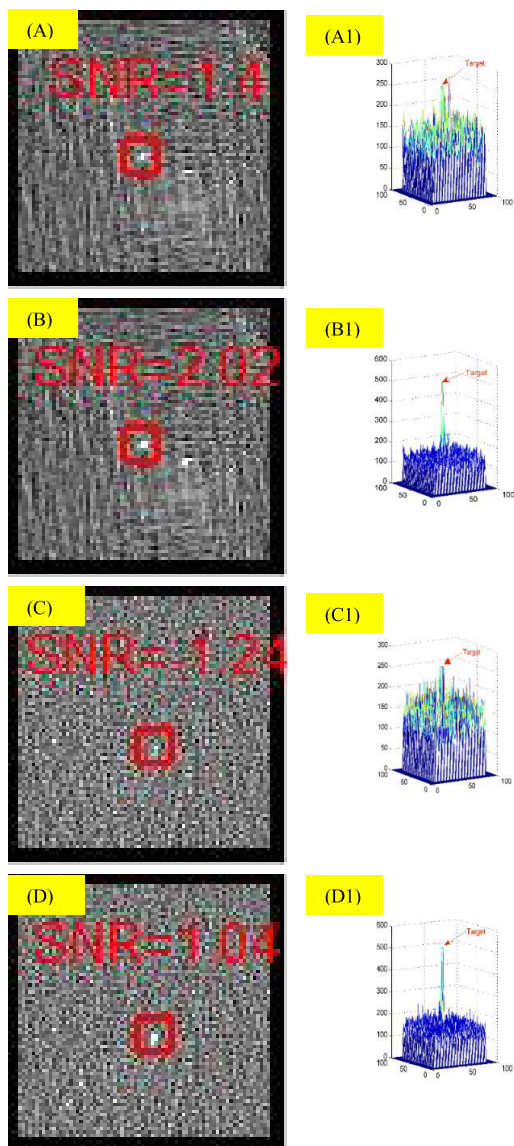
FIGURE 8. Comparison before and after target energy enhancement in two scenes.

FIGURE 9. The operation results of each algorithm in scenario 3. (A), (B), (C), (D) denote the detection results of the non-uniform corrected multi-directional gradient model, the traditional high-order correlation accumulation model, the two-point correction model and the algorithm in this paper, respectively.

has obvious effects on target signal retention and signal enhancement, which indicates that the spiral gradient estimation model is feasible for background suppression in strong daytime illumination scenes. The data in the above table also indicate that the target energy is further enhanced after the improved higher-order correlation energy enhancement model in this paper, which is a better basis for subsequent target detection and extraction.

At the same time, in order to further illustrate the retention and enhancement effect of the algorithm in this paper on the target energy, different algorithms are compared with the algorithms proposed in this paper, including non-uniform correction and multi-directional gradient weak and small target detection [24], two-point correction detection algorithm [23] and high-order operation correlation detection algorithm [32]. The specific experimental results are as follows:
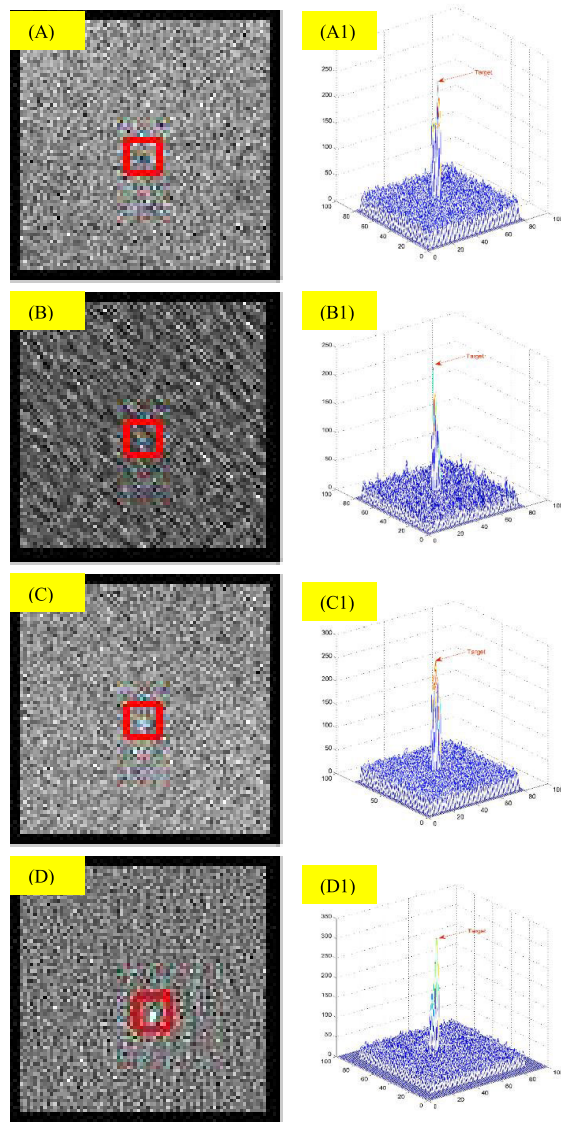
From Fig 9 and 10 above, it can be observed that, compared with the traditional algorithm, the target discrimination in the difference map of this paper's algorithm in the same scene is clearer, and it can be observed that the target is present in an isolated point in the image, indicating that this paper's algorithm is better in suppressing the strong background while retaining the target energy, and it can be observed from the 3D map of the same scene that, compared with the other three algorithms, this paper's algorithm is higher for the target signal energy It shows that the target signal enhancement idea proposed in this paper is feasible.

C. ALGORITHM ROC ANALYSIS

In order to quantitatively evaluate the target detection ability of this algorithm in strong light environment, ROC curves
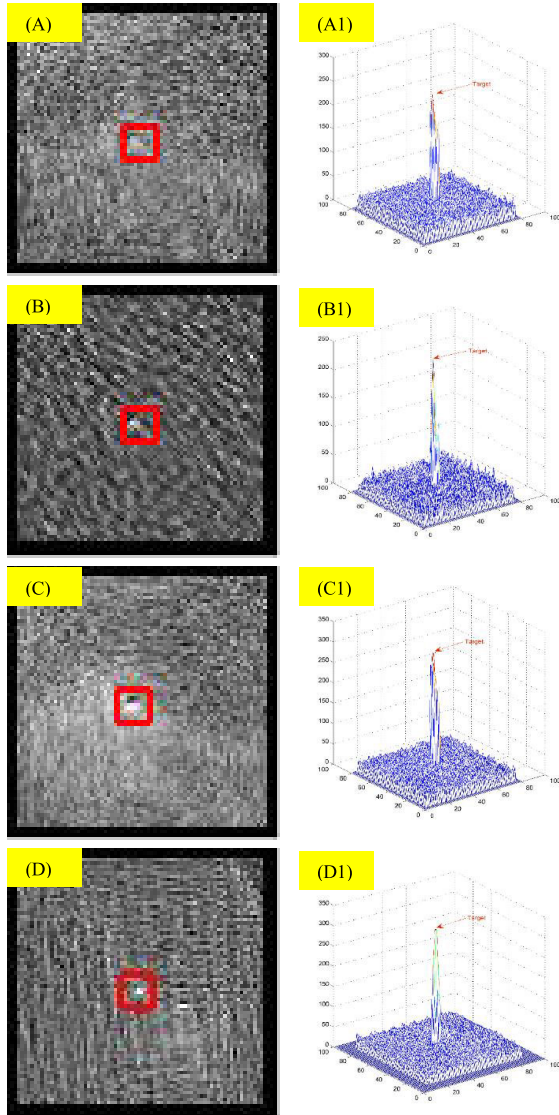
FIGURE 10. The operation results of each algorithm in scenario 4. (A), (B), (C), (D) denote the detection results of the non-uniform corrected multi-directional gradient model, the traditional high-order correlation accumulation model, the two-point correction model and the algorithm in this paper, respectively.

FIGURE 11. ROC curves of different scenarios.

of the above algorithms are drawn in combination with ROC index. The specific definitions are as follows:

$$P_d = \frac{NTDT}{NT} \times 100\%$$
$$P_f = \frac{NFDT}{NP} \times 100\% \quad (12)$$

where, P_d is the detection rate, P_f is the false alarm rate. $NTDT$ is the number of real targets detected, and $NFDT$ is the number of false alarm targets detected. NT is the total number of real targets in the image, and NP is the total number of targets detected in the image. The specific ROC curve is shown in the figure below.

As seen in Fig 11 above, the detection of targets in four different daytime bright illumination scenes by the algorithm in this paper is better, and the detection rate of targets in the
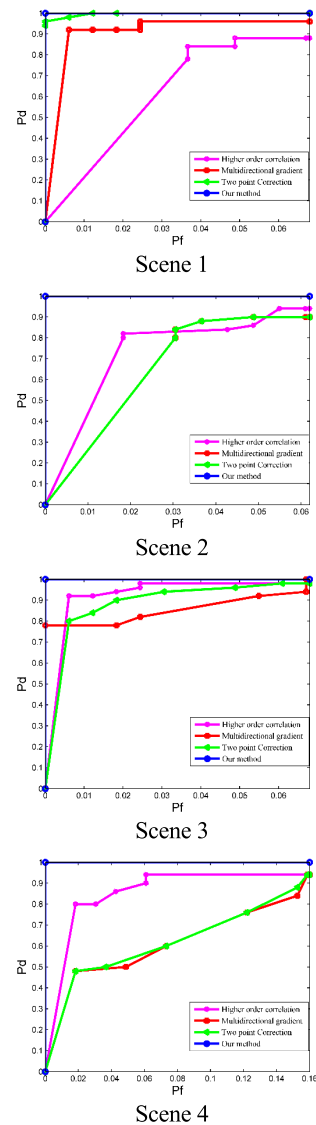
four scenes reaches 100%. In the case of controlling the same segmentation threshold, the detection rate of this algorithm is 100% and the false alarm rate is 0 in scene 1, followed by the two-point correction detection rate of $P_d = 94\%$ and $P_f = 0.0183\%$, and finally the traditional high-order correlation energy enhancement model, the highest detection rate is not 90% and the target is seriously lost. And in scenes 2 and 4, the effect of this paper's algorithm is clearly compared with other algorithms, ranking first with a detection rate of 100%. The traditional high-order correlated single-point energy enhancement model both ranked second with detection rate $P_d = 94\%$ and false alarm rate $P_f = 0.0183\%$, indicating that the two-point correction model was more effective than the other two compared algorithms in these two strong light scenes. In addition, there is not much difference between the weak target detection and the two-point correction model with non-uniform correction and multi-directional gradient, where these two algorithms rank after

TABLE 3. Comparison of operation efficiency of each algorithm.

Scene\Algorithm	Literature [31]	Literature [32]	Literature [35]	Ours
Scene A	4.0428	3.1212	3.0953	83.4861
Scene B	9.7053	4.3212	3.2476	86.4889
Scene C	7.7090	6.3748	6.3854	81.8207
Scene D	3.4245	3.8925	3.6396	86.8354

the conventional high-order with the same detection rate and false alarm rate in scene 2. In Scenario 3, the traditional high-order correlated energy enhancement model and the two-point correction model play smoothly, and the overall detection rate and false alarm rate of both algorithms are better than the non-uniform correction and multi-directional gradient weak target detection, with the traditional high-order correlated energy enhancement model ranking second with an average detection rate of 94.3333% and the two-point correction ranking third with a detection rate of 90.3333%.

D. D TIME CONSUMPTION
In this paper, the time consumed by one frame image is used (Frame/Second) to measure the execution efficiency of different algorithms. The algorithm consumption of each model is shown in Table 3 below. It can be seen from table 3 that although the traditional algorithm takes a short time, the target detection rate is low. The execution efficiency of the algorithm in this paper is relatively low, but the detection rate is high. In the actual detection, parallel computing can be used to improve the detection efficiency.

V. CONCLUSION
After the above experimental verification, it can be observed that this algorithm has a good effect on target detection in strong light scene, which fully shows the applicability of this algorithm for target detection in strong light scene in the daytime. After experimental verification, relevant conclusions can be drawn as follows:

(1) In this paper, the spiral gradient estimation model is used to model the background of the image, which can remove the halo and effectively retain the target signal. The experimental results show that the target is clearly visible in the difference map and the signal-to-noise ratio is good, which shows that this method is feasible and effective.

(2) In this paper, an improved high-order correlation energy accumulation model is proposed to enhance the image target signal. Firstly, the attention discrimination model of the inner and outer windows is constructed by using the gradient difference between the target and the background region to obtain the significant region with large regional gradient difference, and then the significant region block is

subjected to high-order motion correlation. Compared with the other three algorithms, In this paper, the target signal energy is high, and the local signal-to-noise ratio of the target is significantly improved.

REFERENCES
[1] Y. Lu, S. Huang, and W. Zhao, "Sparse representation based infrared small target detection via an online-learned double sparse background dictionary," *Infr. Phys. Technol.*, vol. 99, pp. 14–27, Jun. 2019.
[2] H. Qi, B. Mo, F. Liu, Y. He, and S. Liu, "Small infrared target detection utilizing local region similarity difference map," *Infr. Phys. Technol.*, vol. 71, pp. 131–139, Jul. 2015.
[3] L. Deng, H. Zhu, Q. Zhou, and Y. Li, "Adaptive top-hat filter based on quantum genetic algorithm for infrared small target detection," *Multimed. Tools Appl.*, vol. 77, no. 9, pp. 10539–10551, May 2018.
[4] J. Jiao and L. D. Wu, "Infrared dim small target detection method based on background prediction and high-order statistics," in *Proc. Int. Conf. Image, Vis. Comput.*, vol. 2, 2017, pp. 53–57.
[5] L. Deng, J. Zhang, and H. Zhu, "Infrared moving point target detection using a spatial-temporal filter," *Infr. Phys. Technol.*, vol. 95, pp. 122–127, Dec. 2018.
[6] X. Fan, Z. Xu, J. Zhang, Y. Huang, Z. Peng, Z. Weia, and H. Guod, "Dim small target detection based on high-order cumulant of motion estimation," *Infr. Phys. Technol.*, vol. 99, pp. 86–101, Jun. 2019.
[7] Y. H. Wang and W. N. Liu, "Dim target enhancement algorithm for low-contrast image based on anisotropic diffusion," *Opto-Electron. Eng.*, vol. 35, no. 6, pp. 15–18, 2008.
[8] Q. Lin, S. C. Huang, X. Wu, and Y. Zhong, "Infrared small target detection based on nuclear anisotropic diffusion," *High Power Laser Part. Beams*, vol. 27, pp. 011014-1–011014-6, Jan. 2015.
[9] H. Zhu, Y. Guan, L. Deng, Y. Li, and Y. Li, "Infrared moving point target detection based on an anisotropic spatial-temporal fourth-order diffusion filter," *Comput. Elect. Eng.*, vol. 68, pp. 550–556, May 2018.
[10] X. Bai and F. Zhou, "Analysis of new top-hat transformation and the application for infrared dim small target detection," *Pattern Recognit.*, vol. 43, no. 6, pp. 2145–2156, 2010.
[11] L. Z. Deng, J. Zhang, G. Xu, and H. Zhu, "Infrared small target detection via adaptive M-estimator ring top-hat transformation," *Pattern Recognit.*, vol. 112, pp. 1–8, Apr. 2021.
[12] O. E. Drummond, S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," *Proc. SPIE*, vol. 3809, pp. 74–83, Oct. 1999.
[13] D. Pang, T. Shan, P. Ma, W. Li, S. Liu, and R. Tao, "A novel spatiotemporal saliency method for low-altitude slow small infrared target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.
[14] D. Pang, T. Shan, W. Li, P. Ma, and R. Tao, "Infrared dim and small target detection based on greedy bilateral factorization in image sequences," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3394–3408, 2020.
[15] C. Q. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4996–5009, Dec. 2013.
[16] C. Wang and S. Qin, "Adaptive detection method of infrared small target based on target-background separation via robust principal component analysis," *Infr. Phys. Technol.*, vol. 69, pp. 123–135, Mar. 2015.
[17] J. Fan, Y. Gao, Z. Wu, and L. Li, "Infrared dim small target detection technology based on RPCA," in *Proc. 3rd Int. Conf. Electron. Inf. Technol. Intellectualization (ICEITI)*, 2017, pp. 748–752, doi: 10.12783/dtcse/iceiti2017/18914.
[18] Y. Wei, X. You, and H. Li, "Multiscale patch-based contrast measure for small infrared target detection," *Pattern Recognit.*, vol. 58, pp. 216–226, Oct. 2016.
[19] X. Wang, Z. Peng, D. Kong, P. Zhang, and Y. He, "Infrared dim target detection based on total variation regularization and principal component pursuit," *Image Vis. Comput.*, vol. 63, pp. 1–9, Jul. 2017.
[20] M. X. Zhao, L. Cheng, X. Yang, P. Feng, L. Liu, and N. Wu, "TBC-Net: A real-time detector for infrared small target detection using semantic constraint," Tech. Rep., 2015.
[21] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2014.
[22] X. Sun, X. Liu, Z. Tang, G. Long, and Q. Yu, "Real-time visual enhancement for infrared small dim targets in video," *Infr. Phys. Technol.*, vol. 83, pp. 217–226, Jun. 2017.

[23] Y. Sun, J. Yang, and W. An, "Infrared dim and small target detection via multiple subspace learning and spatial-temporal patch-tensor model," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 5, pp. 3737–3752, May 2020.

[24] L. D. Zhang and Z. M. Peng, "Infrared small target detection based on partial sum of the tensor nuclear norm," *Remote Sens.*, vol. 11, no. 4, pp. 1–4, 2019.

[25] T. C. Zhang, Y. Q. Wu, and F. Zhou, "Infrared small target detection based on robust principal component analysis joint directional derivative penalty," *J. Phys., Conf. Ser.*, vol. 1873, no. 1, pp. 1–6, 2021.

[26] M. Shi and H. Wang, "Infrared dim and small target detection based on denoising autoencoder network," *Mobile Netw. Appl.*, vol. 25, no. 4, pp. 1469–1483, Aug. 2020.

[27] J. J. Bai, H. Y. Zhang, and Z. J. Li, "The generalized detection method for the dim small targets by faster R-CNN integrated with GAN," in *Proc. IEEE Int. Conf. Commun. Inf. Syst.*, vol. 3, Dec. 2018, pp. 1–5.

[28] Q. Hou, Z. Wang, F. Tan, Y. Zhao, H. Zheng, and W. Zhang, "RISTDnet: Robust infrared small target detection network," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[29] X. Liu, X. Li, L. Li, X. Su, and F. Chen, "Dim and small target detection in multi-frame sequence using Bi-conv-LSTM and 3D-conv structure," *IEEE Access*, vol. 9, pp. 135845–135855, 2021.

[30] Z. L. Li, T. S. Sheng, and S. L. Lou, "Correction method of vignetting effect in infrared system based on polynomial approximation," *Infr. Laser Eng.*, vol. 45, no. S1, 2016, Art. no. S104002.

[31] Y. Z. Zhang, Z. W. Hu, H. J. Meng, and X. Zhi, "Influence of vignetting on the accuracy of array CCD pyrometer and its correction method," *J. Northeastern Univ.*, vol. 35, no. 5, pp. 609–612, 2014.

[32] Y. Ding, J. T. Tao, W. Quan, and H. Cheng, "Nonlinear compensation method for optical vignetting effect of visual system," *J Tsinghua Univ.*, vol. 57, no. 7, pp. 702–706, 2017.

[33] X. Y. Jang and S. Z. Zhou, "Hidden Markov models for automatic detection and tracking of multiple targets," *Acoust. Electr. Eng.*, vol. 124, no. 4, pp. 29–33, 2016.

[34] J. X. Zhang and C. R. Wang, "Motion target detection method based on edge information and spatio-temporal Markov model," *J. Yanshan Univ.*, vol. 35, no. 2, pp. 124–127, 2011.

[35] Y. Yong, "Research on small target detection and recognition technology," Ph.D. dissertation, D, Inst. Optoelectron. Technol., Chin. Acad. Sci., pp. 51–60, 2005.

JULIU LI received the bachelor's degree from the School of Electrical, Electronic and Computer Science, Guangxi University of Science and Technology, Liuzhou, China. He is currently a Graduate Student with the School of Electrical, Electronic and Computer Science. His research interests include signal processing and image detection and tracking algorithms.

HONGWEI GUO (Member, IEEE) received the B.S. and M.S. degrees from Yunnan University, Kunming, China, in 2002 and 2010, respectively, and the Ph.D. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2019. Since 2020, he has been a Professor with the School of Engineering, Honghe University, Mengzi, China. From 2013 to 2014, he was a Visiting Scholar with the Electrical Engineering Department, Tsinghua University, Beijing, China. His current research interests include video compression and video communication. He is currently a Senior Member of China Computer Federation. He was a co-recipient of a Best Student Paper Award and a Best Poster Award at the IEEE BMSB 2018 and the ChinaMM 2018, respectively.

XIANGSUO FAN received the B.S. degree in automation from Hainan Normal University, Haikou, China, in 2012, and the M.S. and Ph.D. degrees from the University of Electronic Science and Technology of China, Chengdu, China, in 2015 and 2019, respectively. Since 2019, he has been a Research Assistant with the School of Electrical and Information Engineering, Guangxi University of Science and Technology, Liuzhou, China. His research interests include signal processing and image detection and tracking algorithms.

LINPING FENG (Member, IEEE) received the M.S. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2015, and the Ph.D. degree in electrical and computer engineering from the University of Macau, Macau, China, in 2018. He is currently with the School of Electronic and Information Engineering, Xi'an Jiaotong University, as an Assistant Professor. His current research interests include millimeter-wave monolithic integrated power amplifier and systems and high-sensitivity detection technology.

LEI MIN received the B.S. and M.S. degrees in mathematics from Sichuan University, Chengdu, China, in 2009 and 2012, respectively, and the Ph.D. degree from the University of Electronic Science and Technology of China, in 2019. Since 2019, he is currently a Researcher with the Institute of Optics and Electronics, Chinese Academy of Sciences. His research interests include signal processing, image super-resolution, image detection, and tracking algorithms.

ZHIYONG XU was born in Chengdu, Sichuan, China, in 1767. He received the B.S. degree from Sichuan University and the M.S. degree from the University of Electronic Science and Technology of China. He is currently a Researcher with the Institute of Optics and Electronics, Chinese Academy of Sciences and a Master's Supervisor of the University of Chinese Academy of Sciences. His research interests include signal processing and image reconstruction algorithms.

• • •