Dim and Small Target Detection Based on Gaussian Markov Random Field Motion Direction Estimation

LEI MIN1, XIANGSUO FAN2,3, JULIU LI2, ZHENGRONG XIANG2, AND QIQI WU2

1Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, China
2School of Electrical, Electronic and Computer Science, Guangxi University of Science and Technology, Liuzhou 545006, China
3Division of Intelligent Manufacturing, Yibin University, Yibin, Sichuan 644600, China

Corresponding authors: Xiangsuo Fan (100002085@gxust.edu.cn) and Zhengrong Xiang (221042138@stdmail.gxust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 62001129, in part by the Guangxi Natural Science Foundation under Grant 2021GXNSFBA075029, and in part by the Guangxi Science and Technology Base and Talent Project under Grant AD19245130.

ABSTRACT To improve the detection of dim and small targets in low signal-to-noise ratio (SNR < 3dB) scenarios, energy accumulation along the estimated direction by estimating the direction of the target’s motion at different moments is a proven approach, and considering that the Gaussian Markov random field can well describe the correlation of image element points in their air and time domains, Gaussian Markov random field is introduced into the paper to process the image. Firstly, by combining the proximity of image element spatial coordinates and the correlation of image element values to construct a new adaptive Gaussian weighted Markov random field filtering model, which fully incorporates the spatio-temporal information of image element coordinates and the correlation degree information of image element values for weighting to describe the image element neighborhood system. Second, to further enhance the energy of weak targets, the energy accumulation of targets is achieved by establishing a Gaussian Markov random field motion direction estimation model on the basis of obtaining filtered images. Through experiments, it is shown that the algorithm proposed in this paper can effectively accumulate the weak signal of the target between frames, and the target signal-to-noise ratio can reach more than 6dB compared with that before enhancement, effectively improving the detection ability of weak targets.

INDEX TERMS Dim and small targets, detection, Gaussian Markov random field, energy accumulation, motion estimation.

I. INTRODUCTION

Dim and small target detection is an important in computer vision research, which is widely used in aerospace, defense and security engineering. However, due to the lack of relevant texture information with targets in long distance imaging, the inherent characteristics of target is with fewer image elements and lower signal-to-noise ratio which pose certain challenges for detection. In addition, combined with the mechanism of high-altitude imaging considerations, the image also contains atmospheric turbulence, complex and variable clouds and a variety of photoelectric noise and other interference information often submerge target, that causing difficulties to the target detection [1], [2]. In order to overcome the above-mentioned difficulties in target detection and tracking, we need to suppress the background of the image and divide the image into background area, target area and noise area [3] as far as possible. And the target signal is retained while the background and other kinds of interference noise are suppressed to the maximum extent possible, which becomes the core of the research of each algorithm. Nowadays small target detection algorithms can be generally divided into three research directions: traditional spatiotemporal filtering, low rank sparse algorithms and deep learning respectively. The spatiotemporal domain filtering can be divided into two broad directions are detect before track (DBT) and track before detect (TBD) [4], [5]. and whether it is DBT or TBD, these two types of algorithms first use preprocessing to eliminate most of the background interference on the target, and then combine multi-frame motion correlation to detect the target.
The order before and after is different in algorithms, such as DBT are simple to operate, but when the target is submerged by various noises and clutter, the detection effect is not ideal, and most of the TBD algorithms need to know the prior knowledge of the target’s motion model, motion direction, and trajectory. Therefore, the scope of application of these detection methods is bound to be limited, and such algorithms are not effective for detection and tracking in low signal-to-noise ratio scenarios.

Traditional detection algorithms include bilateral filtering [6], [7], anisotropic filtering [8]–[11], TDMS filtering [12]–[14], gradient inverse filtering [15]–[17], Markov random field model [18]–[23] etc. These algorithms that preserve the target signal while processing the image background have achieved relatively good results. However, with the development of science and technology, the imaging distance of small targets in the time domain is enlarging, and the various interference information encapsulated in their images is increased, the traditional detection algorithms will cause incomplete background suppression, and bring difficulties to target detection in low signal-to-noise ratio and high complexity images [24]. For example researcher Zeng et al. who proposed the improved bilateral filtering to achieve the background suppression, which can effectively enhance the signal of the target by grayscale weighting the image element neighborhoods and then filtering the image with a fixed filter template. However, its reliance on a fixed filtering template is more restrictive in face of changing complex spatio-temporal domain background, which will result the phenomenon in incomplete background suppression, and the image still retains more edge contour information, which is not conducive to the detection of targets [6]. The anisotropy-based edge detection algorithm proposed by Preona and Malik [8] achieves better performance in background suppression, and its method of using the combination of diffusion function and filter function for background suppression which effectively removes the background, and the algorithm provides a new research idea for the detection of small targets, and the improvement of the anisotropic kernel diffusion function for background suppression is respected by many scholars. For example, the method of infrared small target detection based on nuclear anisotropic diffusion proposed by Lin et al. [10] achieved better results, which by constructing a new diffusion function for background suppression of images, indicating that the anisotropic filtering algorithm has a greater advantage in background suppression. However, the diffusion function adaptive is not implemented, and the algorithm will suppress the target as background while background suppression converting to more complex scenes, which will eventually cause the target detection to fail. The two-dimensional mean square difference filtering based on neighborhood information [12] by Wan and Wang has achieved better results, which using the obvious difference between the target image element and the background to achieve the prediction of the background, using the correlation of the image in the time domain to adaptively generate different filter weights to predict the image element one by one, fully utilizing the image detail information. However, the information components held by the target in different directions are different, and the method uses the same weight matrix for all directions to calculate the next predicted weights, which tends to destroy the information of the target at different moments. In addition, the two-dimensional mean square (TDLMS) filtering is performed only on the X, Y axes of the image, without considering the image is captured in whole-frame pattern rather than in a direction-by-direction acquisition, which makes the signal of the target not fully retained when filtering is performed and is not conducive to the next step of detection. The gradient inverse filtering model in traditional algorithms is more typical, and the principle of the small target detection algorithm based on adaptive filtering with strong undulating background proposed by Li and Dong, [15] is to use the gradient inverse for background prediction. The gradient inverse is used to distinguish the background from the target by taking the inverse of the difference between the image elements. The inverse value of the target and the background is smaller when the inverse is taken, while the inverse value of the other image elements is larger after the difference is taken. This method has good background suppression characteristics for large areas, but the fixed filtering parameters make even the filtering parameters obtained for elements with small differences are not enough to suppress the background completely, there still some noise or edge contours in the image which affect the detection. The above traditional algorithm mainly detects and extracts target after background suppression from the perspective of the spatio-temporal domain. In order to make full use of the time-domain correlation between the front and back frames of sequential images to achieve target detection and extraction. Zhang et al. proposed a motion target detection method based on edge information and spatio-temporal Markov model [21], whose method using Markov random fields to do marker fields for three consecutive frames in an image to extract image edges, which provides a solid theoretical basis for the detection of small targets, and for this, Wang et al. proposed an MRF-based adaptive regularized infrared background clutter suppression algorithm [22] to background suppression of images achieved better results, which can effectively suppress the background of single-frame images and preserve the target signal, indicating that Markov random fields can be applied to the background processing of images. In addition, Jiang et al. proposed an algorithm for the application of hidden Markov model in automatic multi-target detection and tracking achieved significant results [19], which estimates the direction of target motion through the time-domain information of sequence images to achieve target tracking. It is shown that Markov random field algorithm can be better adapted to the detection of small targets in space-time domain sequence images, which provides a better idea for the algorithm research in this paper.

Low-rank sparse algorithm for background prediction mainly uses the low-rank property of the image to decompose
the image to find the rank [25], so as to obtain the low-rank part of the image and the sparse part to complete the background prediction, and such algorithms need to train the samples of the image and then compose the feature dictionary, which makes the sparse matrix of the relevant calculation extremely complicated and increases the operation cycle of the algorithm, and such algorithms are mainly for single-frame image and the algorithm mainly deals with single-frame images, and requires a large dictionary of feature information to support multi-frame sequence images, which greatly slows down the efficiency of detection. But compared with traditional algorithms, these algorithms have a greater advantage in background suppression and can detect target accurately, and their typical algorithms are IPI model [26], TV-PCP model [27], RPCA model [28], [29], etc. The IPI accurately, and their typical algorithms are IPI model [26], advantage in background suppression and can detect target with traditional algorithms, these algorithms have a greater advantage in both time and space domains. Firstly, we construct a new Markov random field filtering model based on Gaussian weighting to incorporate the coordinate information of image elements and the grayscale correlation between image elements into the distribution probability of the image element domain system to improve the detection adaptability of Markov random field. Second, in order to make the target extracted accurately, this paper also constructs a Markov random field based motion direction estimation model to accumulate the energy of the target and prepare for target extraction. The detailed description of the algorithm is shown in the algorithm section of this paper.

II. CONSTRUCTION OF GAUSSIAN MARKOV RANDOM FIELD

Markov random fields effectively describe the correlation of a random variable with other random variables in its neighborhood range, conditionally independent of other variables outside the neighborhood range [23]. Based on this property of Markov random field, it can be employed to describe the degree of correlation between individual image elements in an image and their neighborhood ranges. In practical applications, the variation of an image element in an image with respect to its neighborhood range can be understood as a random field, there is spatial correlation between image pixels, and a pixel can be determined by the set of its neighboring pixels [18]. Setting $S = \{s_1, s_2, \cdots, s_n\}$ is a set in the neighborhood of an image element, $s_i$ is random variable, and the $X = \{x_{s_i}, s_i \in S\}$ is indicates a random field. When the assemblies $P(x_{s_i} | x_{s_i}, \cdots, x_{s_n})$ is link to the $x_{s_j},$ the $s_j$ is a neighborhood point of $S_i.$ The $D_{s_i}$ is a neighborhood set of $S$ and $D = \{D_{s_i}, s_i \in S\}$ is the neighborhood systems of $S$ also. Assuming the $\Lambda_{s_i}$ is the value domain of $x_{s_i},$ $\Omega = \{x = (x_{s_1}, x_{s_2}, \cdots, x_{s_n}) : x_{s_i} \in \Lambda_{s_i}, 1 \leq i \leq n\}$ is the set of states for each random variable. For any $s_i, s_j \in S$ and $x \in \Omega,$ there is $P(x) > 0$ and with follow formula as:

$$P(x_{s_i} | x_{s_j}, s_j \neq s_i, s_j \in S) = P(x_{s_i} | x_{s_j}, j \in D_{s_j})$$

in the formula the $X$ is called Markov Random Field(MRF) relate to the neighborhood domain system of $D,$ named the One-dimensional Markov chains also. And if take the Gaussian distribution to indicate the joint probability of MRF, the $X$ is represent the Gaussian Markov Random Field(GRMF) [18], and the specific formula is as follows:

$$P(x | D) = \frac{1}{(2\pi \sigma^2)^{n/2}} \exp \left( -\frac{\sum_{s_j \in D} w_{s_j} x_{s_j}^2}{2\sigma^2} \right)$$

where the $x_{s_i}$ is the element’s value, $s_j \in D$ is the neighborhood set of pixel, $\sigma$ represent the neighborhood standard deviation, and the $w_{s_j}$ is weighting parameters. In the literature [18], when the $N = \{(0, 1), (0, −1), (−1, 0), (1, 0), (1, −1), (−1, 1), (−1, −1), (1, 1)\},$ the weight parameters of GMRF model are as follows: in order to effectively enhance the target signal from the difference image, the target signal enhanced is necessary. Thus,
the target’s motion model was built based on literature [32], which possess the definition of transition probability. This model can describe the images’ pixel motion transition probability in moment \( K \) to the next \( K + 1 \) moment, and utilize the weighted form to enhance target signal. The specific defined as follows:

**Definition 1:** Setting \( \{X(n) : n \geq 0\} \) is the Markov chain, the state space is \( E = \{0, 1, 2, \ldots\} \), called the conditional probability, the matrix follow as:

\[
p_{ij}^{(k)}(m) = P \{X(m+k) = j \mid X(m) = i\} \quad (3)
\]

where \( X(m) \) is take value \( i \) in \( m \) moment and after \( k \) steps of transfer, at \( m + k \) moments, the conditional probability of \( X(m+1) \) taking the value \( j \), that defined as the \( k \) step transfer probability of a Markov chain at time \( m \), and is a one-step transfer probability when \( k = 1 \).

**Definition 2:** Setting \( \{X(n) : n \geq 0\} \) is the Markov chain, the matrix

\[
P^{(k)}(m) = p_{ij}^{(k)}(m) = \begin{bmatrix} p_{00}^{(k)}(m) & p_{01}^{(k)}(m) & \cdots & p_{0n}^{(k)}(m) \\ p_{10}^{(k)}(m) & p_{11}^{(k)}(m) & \cdots & p_{1n}^{(k)}(m) \\ \vdots & \vdots & \ddots & \vdots \\ p_{n0}^{(k)}(m) & p_{n1}^{(k)}(m) & \cdots & p_{nn}^{(k)}(m) \end{bmatrix} \quad (4)
\]

is the \( k \)-step transfer matrix of the Markov chain at moment \( m \), when \( k = 1 \), the matrix \( P(m) = p_{ij}(m) \) represent one-step transfer matrix and for any \( i \in E \) of transfer matrix. For any row in one-step transition matrix \( i \in E \), the \( p_{ij}(m) \geq 0 \), the sum of all column elements in this row is 1, that is \( \sum_{j \in E} p_{ij}(m) = 1 \).

### III. ALGORITHMS

The dim target image is generally composed of three parts, smooth background region + target region + non-smooth background region, which are shown in Fig. 1. The analysis reveals that the grayscale difference between the target region and its neighborhood range is large, the grayscale difference between the smooth background region and its neighborhood range is smallest, and the non-smooth background region and its neighborhood range are not much different from the central image element because some image elements are in the non-smooth edge contour region. The weighted filtering function can be constructed with the help of image differences within different regions to enhance the target signal while suppressing most of the background information.

In order to effectively suppress most of the background and enhance the target signal at the same time, a new adaptive Gaussian weighted Markov random field filtering model is first established in the paper to obtain the difference image of the image. When the difference between the target area of the central image element and the neighboring background area is small, the algorithm can effectively filter out this part of the background area after filtering the image because the difference between the central image element and the neighboring area is not large; after filtering the image by the above algorithm, on the basis of obtaining the difference image, in order to further enhance the target signal, the paper proposes a Gaussian Markov random field motion direction estimation model based on In order to further enhance the target signal, a Gaussian Markov random field motion direction estimation model is proposed in the paper, which firstly establishes the motion model of the target in the neighborhood, and then constructs a Markov transfer probability model of the target based on the established motion model to estimate the motion direction of the target, and adopts the form of transfer probability weighting along the motion direction to enhance the target signal. The specific flow of the algorithm in this paper is shown in Figure 2 below.

#### A. ADAPTIVE GAUSSIAN WEIGHTED MARKOV RANDOM FIELD FILTERING MODEL

The traditional first-order GMRF model is difficult to adapt to the dynamically changing scene data due to the use of fixed filtering weight parameters. In order to obtain a difference image that can effectively retain the target signal and suppress most of the background parameters, an adaptive Gaussian weighted Markov random field filtering model is proposed in the paper, which fully incorporates the space domain information of image element coordinates.
FIGURE 2. Overall flow chart of the algorithm.

and the similarity degree information of image element values for weighting to better describe the joint probability distribution of the image element neighborhood system, as follows:

1) A weighting function that fully incorporates the coordinate space domain information of the neighboring system and the similarity information of the image element values is established with the following equation:

\[ w = \frac{(row - i)^2 + (col - j)^2}{2\sigma_1^2} \times \frac{(f(i+row, j+col) - f(i, j))^2}{2\sigma_2^2}, \]

in the formula the \((i, j)\) is the coordinates of the central image element point, \(f(i, j)\) is the gray value of center point, \(row, col\) represent the coordinates of \(R \times R\) neighboring domain of center pixel, \(f(i + row, j + col)\) is the gray value of \(R \times R\) neighboring domain of center pixel, \(\sigma_1\) indicate the standard deviation of neighborhood coordinates of \(R \times R\), \(\sigma_2\) indicate the standard deviation of neighborhood gray value of \(R \times R\), since \(\sigma_1\) and \(\sigma_2\) are automatically calculated based on the neighborhood area, they are able to adapt to the dynamic changes of the scene.

2) The image is weighted and filtered using a function with the following equation:

\[
\{ \begin{align*}
    f'(row, col) &= f(row, col) \times w(row, col) \\
    row, col &\in R \times R
\end{align*} \]

where the variables are defined as above.

3) Creating a new Gaussian Markov joint probability weighting function for the filtered image and the specific formula is as (7), shown at the bottom of the page, in this formula the \(\sigma\) is image’s standard deviation of \(R \times R\) behind filtering, \(w_p\) is the weights obtained by the Gaussian Markov weighting function, the other variables are defined as above.

4) Using Gaussian Markov weighting function to enhance the signal at the center image element point \((i, j)\), the specific formula is as (8), shown at the bottom of the page, in the formula \(F(i, j)\) is the difference image obtained after Gaussian Markov weighted filtering, and other variables are defined as above.

B. GAUSSIAN MARKOV RANDOM FIELD MOTION DIRECTION ESTIMATION MODEL

The difference image obtained after Gaussian Markov weighted filtering, the target energy in the image is still weak, in order to further enhance the target signal, a Gaussian Markov random field motion direction estimation model is
proposed in the paper to achieve the target energy enhancement. The algorithm first constructs the motion model of the target neighborhood, then constructs the target Markov state space and one-step transfer probability model based on the target neighborhood motion model to estimate the motion direction of the target, and finally adopts the form of transfer probability weighting along the motion direction to enhance the target signal. The details are as follows:

1) Constructing the motion model of the target neighborhood, due to the long-distance imaging, weak targets move slowly at different moments, usually less than or equal to 2 pixel/s. Thus, the motion model of the target in the adjacent frames in 9 directions, where “1, 2, 3, 4, 5, 6, 7, 8, 9” denotes the target at rest between adjacent frames, the target moving to the right, the target moving to the upper right, the target moving to the upper left, the target moving to the lower left, the target moving to the lower right, and the target moving to the upper right, respectively, as shown in Figure 3.

2) According to the target motion model to build the motion transfer probability model, corresponding formula as follows:

\[
F'_k(i, j) = Move^9(F_k(i, j)) \\
F'_t(i, j) = F'_k(i, j) + F_{k+1}(i, j) \\
p_{s_k, s_{k+1}} = \max \left( F'_1(i, j), F'_2(i, j), \ldots, F'_9(i, j) \right)
\] (9)

where \(Move(\cdot)\) is the moving function of 9 directions in neighborhood domain and it use the difference diagram \(F_k(\cdot)\) of k moments to get the motion model diagram by moving along 9 directions respectively, t is 9 directions of serial number, \(F'_t(\cdot)\) is the graph of the t th cumulative result at moment k and moment \(k + 1\), \(\max(\cdot)\) is take the maximum value function, \(I\) is the position of maximum pixel point \((i, j)\), \(p_{s_k, s_{k+1}}\) is motion transition probability matrix, with a matrix size of \(3 \times 3\), and defined as follows:

\[
p_{s_k, s_{k+1}} = \begin{cases} 
1, & \text{if } s_k = s_{k+1} = 1 \\
1/3, & \text{else}
\end{cases}
\] (10)

Give an example, if the maximum location of pixel \((i, j)\) is 9, the motion transition probability matrix as follow:

\[
p_{s_k, s_{k+1}} = \begin{bmatrix} 
1/3 & 1/3 & 1/3 \\
1/3 & 1/3 & 1/3 \\
0 & 0 & 1 
\end{bmatrix}
\]

3) One-step transfer probability matrix weighting is used to enhance the target signal, the formula are as follows:

\[
F_e(i, j) = \sum_{row=-R/2}^{R/2} \sum_{col=-R/2}^{R/2} F'_t(i + row, j + col) \times p_{s_k, s_{k+1}}(i + row, j + col)
\]

in the formula the \(F_e(\cdot)\) is enhanced results image, \(F'_t(\cdot)\) is totalized result of the position where the maximum value of \((i, j)\) in image is located, \(p_{s_k, s_{k+1}}\) is motion transition probability matrix.

C. ALGORITHM SUMMARY

Based on the above steps, the pseudo-code of the algorithm in this paper is summarized and shown in Table 2:

D. DATASET APPLICATION

In this paper, three data sets are used to verify the feasibility and progress of the proposed algorithm model. Among them, data set A and data set B are collected by our team in the field, and data set C is the public data set in reference 34. The relevant data sets are described in the table below.

IV. RESULTS AND ANALYSIS

A. ANALYSIS OF BACKGROUND SUPPRESSION RESULTS

In the paper, the background suppression effect of this paper is evaluated from qualitative and quantitative perspectives.
and the quantitative evaluation is analyzed from the suppression effect of three different regions in the original figure, and it can be seen from Figure 3 that the smooth background region and the non-smooth region, the adaptive Gaussian weighted Markov random field filtering model proposed in this paper can effectively suppress the smooth background region and the non-smooth edge contour region, while the target region, the algorithm in this paper suppresses the background while effectively retains the target signal.

In order to better reflect the effect of the proposed algorithm in this paper after background suppression and target energy enhancement, two quantitative evaluation metrics, BSF and SNR, are used in the paper to calculate the effect of this algorithm and other algorithms on background suppression, and target energy preservation. The specific mathematical models of the two are as follows [5], [35]:

\[
\begin{align*}
BSF &= \frac{\sigma_{in}}{\sigma_{out}} \\
SNR &= 10 \times \log_{10} \left( \frac{E_T - E_B}{\sigma_B} \right) 
\end{align*}
\]

where \(\sigma_{in}\) and \(\sigma_{out}\) are the mean squared differences between the input image and the difference image, respectively, BSF is the background suppressors. \(E_T\) and \(E_B\) denote the mean values of the target and background regions, respectively, SNR is the signal-to-noise ratio of the image. The specific algorithm comparison data are shown in Table 4.

As shown in the table, the algorithm in this paper has good results in the background suppression and target signal-to-noise ratio of the three scenes, which indicates that the weighted filtering model constructed by the image element neighborhood coordinate information and image element grayscale information can better suppress the background of the image. Meanwhile, as seen from the table, using the image time domain information combined with the direction prediction model constructed by Markov random field to calculate the one-step transfer probability of the target can retain the target signal well and enhance the contrast between the target signal and the image. In order to better illustrate the actual effect of the algorithm in this paper on background suppression and target signal retention in the image, this paper will give the corresponding difference images of the image in the next section, the difference images three-dimensional images, and the three-dimensional images after energy enhancement.

### B. ANALYSIS OF ENERGY ENHANCEMENT RESULTS

In order to be able to effectively illustrate the different processing effects of the algorithm proposed in this paper and other algorithms, this section uses the dataset in the literature [34] as the scene C of this paper’s experiments, one scene collected by the team as the scene A of this paper’s experiments, and then one of the dataset in the literature [34] as the scene C of this paper’s experiments, in terms of the improved bilateral filtering [6], the gradient weighted filtering model [15], the TDLMS filtering [12], anisotropic filtering [8], RPCA model [28], PSTNN model [36], FRKW model [37] and the algorithm proposed in this paper are compared, and the original images of the three experimental scenes, the difference images of each algorithm after background suppression and the difference images in 3D are compared visually to highlight the effect achieved by the algorithm of this paper on the image processing, which are shown in figures (A1)-(A6), (B1)-(B6), (C1)-(C6) and (a1)-(a6), (b1)-(b6), (c1)-(c6) respectively to represent the results obtained from the experiments in the above filters in the three scenarios, and the specific experimental results are shown in the following figure 5, 6 and 7.

As the above three figures can be clearly observed that the algorithm in this paper has better retention effect on the target signal after background suppression than other algorithms. The figures (A1), (B1), (C1) represent the images after the background suppression of the images by bilateral filtering, respectively. From the figures, it can be found that the bilateral filtering can hardly find the obvious weak targets exist in the difference images after the background suppression of the three sequences, and the target points can be observed to exist and be more prominent in the three corresponding 3D images of (a1), (b1), (c1), but there are also more noisy dendrites, which is not conducive to the subsequent further processing of the image. According to the experimental results, it can be shown that the improved bilateral filter, which relies on the filter template structure elements to generate filter weights, is less effective in dealing with images with complex backgrounds and low signal-to-noise ratios; Figures (A2), (B2), (C2) represent the result plots of the gradient inverse weighted filtering model for three scenes after background suppression, and it can be observed in the difference plots of the three scenes that the gradient inverse filtering is more uniform and gentle in the background II in which the background is more completely suppressed and the target is not obvious in the difference images. However, the isolated prominence of the target point on the 3D image of Fig. (b2) is obvious, indicating that the algorithm can suppress most of the background while preserving the target point (as shown
in Fig. 5(B2) above), but it also retains more noise with similar energy values to the target point, as shown in Fig. (A2), (C2), and obvious target highlights can be observed in the difference images of both scenes A and C, but in Fig. (a2) and (c2), there are more interference noises stronger than the target signal present in the 3D images of (c2). It is shown that the weighted normalized filtering method using the inverse of the gradient difference between individual image elements will be retained by the algorithm as a candidate target point when facing multiple continuous noises with little difference, resulting in difficulties in target identification and little differentiation between targets and noises, which brings difficulties to target detection. Figure (A3), (B3), (C3) shows the TDLMS two-dimensional mean square difference filtered difference images after background suppression. There are obvious traces of boundary processing in the difference images of the three scenes, and the target location can be obviously found in the difference images in scene A. However, in the corresponding three-dimensional images (a3), (b3), (c3) of the algorithm, the algorithm has insufficient ability to retain the target signal, and the weak target signal is seriously confused with other noisy signals, which is not conducive to the screening of the target. In addition, if the target is at the edge position, the edge processing ability of the algorithm is not enough to extract the target, resulting in the phenomenon of target detection loss; Figures (A4), (B4), (C4) show the difference images of the background of the three experimental scenes processed by the anisotropic algorithm, as shown in the figure, the use of the kernel diffusion function combined with the gradient information of the four neighborhoods of the image element to filter can effectively suppress the non-smooth edge contours in the background, and the corresponding difference three-dimensional images have fewer noise dendrites and obvious target prominence (Figure (a4), (b4)). The corresponding difference image’s 3D image has less noise dendrites and the target is obvious (Figure (a4), (b4)), which is a good basis for target extraction. However, when the background contains both air domain information and ground information, the algorithm has poor processing effect, such as the figure (C4) and (c4) under the processing of scene C. It can be seen that the air domain part and the ground part under the processing of the algorithm have obvious stratification phenomenon, which indicates that the way of using inter-image gradient information for filtering in the face of the scene of ground conjoined arrangement, the gradient information between the elements is not very different, and in the use of gradient This part of information will be retained when processing in the form of gradient diffusion, which makes the difference images retain more edge contours and noise, and the processing effect is poor, and there is still improved for adjustment. Figures (A5), (B5), (C5) represent the difference result plots of the RPCA low-rank decomposition model after the same filter step and window processing, from which it can be seen that the processing effect is different in all three scenes, and the processing effect is worse in scenes A and C where the background is more complicated, and the presence of the target cannot be obviously observed in the difference plots, and the processing effect is better in scene B where the target is obvious. The different processing effects are zoomed in using the corresponding 3D images (a5), (b5), (c5) of the three scenes, reflecting the fact that the algorithm still has more noise left in the 3D images of scene A and scene C than the target signal, especially in the processing of scene C, where the target has been flooded by the surrounding neighborhood information and the overall signal is maintained at a low 0.2 or less. It shows that the method of background suppression of images using the low-rank property is not very applicable to images with low contrast and background containing various types of interference information.

### Table 4. Comparison of BSF and SNR index of each algorithm in scenarios A, B and C.

| Scene/Algorithm | BSF | SNR | BSF | SNR | BSF | SNR | BSF | SNR | BSF | SNR |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Scene A         |     |     |     |     |     |     |     |     |     |     |
| **BSF**         | 54.4313 | 104.1440 | 22.2786 | 136.8164 | 77.8000 | 68.0686 | 42.3086 | 169.2818 |
| **SNR**         | 5.2600  | 6.6200  | 0.9200  | 6.2800  | 4.8800  | 12.0100 | 16.2900 | 10.4300  |
| Scene B         |     |     |     |     |     |     |     |     |     |     |
| **BSF**         | 6.1348 | 4.9588 | 8.8570 | 25.3510 | 23.0272 | 28.8934 | 11.6948 | 43.3639 |
| **SNR**         | 11.9600 | 9.7500 | 1.3400 | 38.0600 | 9.0800 | 13.1300 | 0.0300 | 39.0400 |
| Scene C         |     |     |     |     |     |     |     |     |     |     |
| **BSF**         | 26.7645 | 20.6638 | 9.5323 | 33.7160 | 20.7767 | 51.4925 | 59.4062 | 51.8290 |
| **SNR**         | 4.4200 | 3.9400 | 0.8300 | 6.5500 | -3.2700 | 5.2200 | 9.4900 | 7.9000 |


In addition, the algorithm is affected by the filtering step and the filtering window size, so that the filtering parameters need to be adjusted continuously when the scene is changed to obtain the corresponding better effect, which increases the debugging time of the algorithm, and there is a large room for improvement in the scene adaptability. Figures (A6), (B6), (C6) is the result of PSTNN model, it can show that most background is suppressed and the target is preserved evidently. But it also with noises and edge retained in difference image, which show the PSTNN model still improvable. Figures (A7), (B7), (C7) show the result of FRKW model after background modeling, from the image can found that this method with better ability to retain target signal, which finish background suppression by Random Walker (RW)
model with adaptive threshold operation, so the target is brighter than other algorithms. But the RW model is suitable for image with less noises, it remain lot of interfere messages in difference image that lead target identification difficulty. Figures (A8), (B8), (C8) show the difference images obtained by the algorithm in this paper after three scenes processing, as shown in the figure, the presence of the target can be obviously observed on the difference images of all three scenes without energy enhancement, indicating that the adaptive Gaussian-weighted Markov random field filtering model constructed in this paper makes full use of the image detail information, and has high adaptivity in the face of different scenes, and the background suppression The effect is better. And it is observed by the corresponding 3D plots (a8), (b8), (c8) that the algorithm in this paper can preserve the signal of the target while suppressing most of the background, which lays a better foundation for the subsequent segmentation and extraction of the target.

To further demonstrate the effect achieved by the algorithm design idea of this paper, the actual effect of the Gaussian Markov random field motion direction estimation model constructed in this paper on the target signal enhancement will be given below to highlight the progressive nature of the algorithm, the actual effect is shown as follows:
As shown in the figure can well compare the effect of the original, difference image and difference image’s three-dimensional image of the algorithm in this paper before and after energy enhancement, from which it can be observed that the signal-to-noise ratio is enhanced from 9.000 to 22.5300 in scene A and from 29.6300 to 36.2000 in scene B. It shows that the Markov random field motion direction estimation model constructed in this paper well enhances the signal of the target and makes sufficient preparation for target detection. In order to more intuitively reflect the effect of this paper’s algorithm on the target signal enhancement, the corresponding statistical tables are produced to compare the signal-to-noise ratios of the three scenes in this paper on the original, difference image and enhanced difference image to reflect the progress of this paper’s algorithm. The specific data are shown in the following table.

### V. DETECTION RESULT

In order to further highlight the progress of the algorithm in this paper, we refer to the ROC (receiver operating curve) curve evaluation index defined in [26] and [27] to evaluate the detection performance of different algorithms in three sequence scenes. The specific formula is as follows:

\[
\begin{align*}
    P_d &= \frac{\text{NTD}}{\text{NT}} \times 100\% \\
    P_f &= \frac{\text{NFD}}{\text{NI}} \times 100\% 
\end{align*}
\]

where PD is the detection rate; NTD represents the number of real target points detected by the algorithm in the sequence image, and NT represents the number of real target points contained in the sequence image; Pf represents false alarm rate; NFD represents the number of false target points contained in the sequence image.
detected in the sequence image, and Ni represents the number of all points detected in the sequence image.

According to the definition of ROC curve, the same sequence image is segmented with different thresholds to calculate the total number of all points segmented in the whole sequence image (target points + false points). Finally, the relevant data are calculated and the corresponding ROC curve is drawn as follows:

As shown above, after calculating the corresponding ROCs for the three scenes, it can be observed that the Markov random field filtering model and the motion direction estimation model constructed in this paper can better suppress
the image background and enhance the target energy in the difference images, which is well prepared for target segmentation extraction. According to the image scene changes, the ROC results of all three scenes as shown are the data results obtained after applying the same segmentation thresholds for image segmentation, respectively. From Figure 10 Scene A, it is obvious to observe that the proposed algorithm in this paper has a detection rate $P_d = 94\%$ at a false alarm rate $P_f = 0.0087$, followed by the RPCA filtering model with a detection rate $P_d = 92\%$, but the corresponding false alarm rate $P_f = 0.0116$. In Scene B, the advantage achieved by this algorithm and other algorithms in the control of the same segmentation threshold situation is greater, in the false alarm rate $P_f = 0$, the detection rate $P_d = 100\%$, when $P_f = 0.0004$ detection rate can also reach 99\%, indicating that the algorithm of this paper is better in this scene, ranked second is the FKRW model, the detection rate $P_d = 99\%$, false alarm rate $P_f = 0.0005$, the third is gradient inverse model and this algorithm data gap is larger, the highest The detection rate $P_d = 95\%$, corresponding to the false alarm rate $P_f = 0.0022$, the fourth is the bilateral filtering, whose detection rate $P_d = 93\%$ when its false alarm rate $P_f = 0.0031$, the image also contains more false target points, and the target detection loss is high, but the overall average detection rate also reached 90.5\%, then the next in line is the RPCA model, TDLMS filtering and anisotropic filtering. In Scene C, our method, gradient inverse model and the bilateral filtering algorithm are better than other algorithms. Under the same segmentation threshold control, the detection rate $P_d$ of these three algorithms can reach 100\%, among which this algorithm and the gradient inverse filtering model reach 100\% under each threshold, while the bilateral filtering will still produce a weak loss of frame with the change of the threshold. The anisotropic filtering algorithm follows, and the detection rate $P_d$ reaches the highest 99\% under the same threshold, and the corresponding false alarm rate $P_f$ is also at a low 0.0025, and then there is FKRW model with detection rate $P_d = 98\%$, false alarm rate $P_f = 0.0051$, and the TDLMS filtering with the detection rate $P_d$ can reach 98\%, in which case this algorithm can still achieve a slight advantage, which again shows the progress and desirability of this algorithm.

**VI. CONCLUSION**

After the above experimental data show that the adaptive Gaussian Markov filtering model based on Markov random field proposed in this paper utilizes the spatial coordinate information of the image and the correlation between image elements to achieve the background suppression of the image, fully integrates the space domain information and time domain information of the image into the image, so that the detail information of the image is fully utilized and finally more target information can be retained in the difference images, which provides the conditions for the subsequent processing of the image. In order to improve the final detection rate of the target, this paper uses the variation of the target’s spare position and time on the pre- and post-frame motion of the scene combined with Markov’s advantages in time and space domain processing to construct a new motion direction prediction model to enhance the energy of the target, highlighting the target region so that it exists as an isolated point, while increasing the contrast and discrimination of the image. After the above series of experiments and analysis, the following points are summarized for the algorithm proposed in this paper:

1) In conventional image filtering, the key is to retain the weak signal target while performing background suppression to lay the foundation for the final extraction of the target. After comparing the data of background suppression evaluation index (BSF) and signal-to-noise ratio (SNR) evaluation index of each algorithm, it is found that the algorithm in this paper can achieve the highest in both indexes, which indicates that the image background suppression is better and the target energy is more completely retained.

2) The fundamental of weak signal target detection is to extract the target smoothly in the complex environment, and one of the conditions to ensure that it can be detected and extracted is to increase the signal-to-noise ratio of the image, and the enhanced motion direction estimation model of the target energy constructed in this paper can be found that the model determines the target by determining the information of the nine directions of the image elements for the target energy enhancement of consecutive multi-frame scenes is more effective. The model enhances the signal-to-noise ratio of the image very well and makes the target identification more fresh.

3) After the algorithm is experimentally verified by the above three scenarios it is obvious to observe that the algorithm in this paper has a better preservation of target energy in the process of image background suppression to obtain image difference images. Secondly, combining the time-domain information and space-domain information of the image for image background processing on the detection of weak signals from long-distance imaging can well improve the detection capability of weak targets.

**REFERENCES**

[1] M. 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,” in *Proc. Comput. Vis. Pattern Recognit.*, 2015, vol. 14, no. 8, pp. 1–17.

[2] J. Liu and H. B. Ji, “Infrared weak target detection based on moving weighted pipe filtering,” *J. Xidian Univ.*, vol. 34, no. 5, pp. 743–745, 2007.

[3] J. G. Sun, “Sequential image infrared small target detection and tracking algorithm research,” Ph.D. dissertation, Univ. Chin. Acad. Sci., Beijing, China, 2014, pp. 1–18.

[4] D. B. Wang, “Research on infrared weak target detection and tracking technology in complex background,” Ph.D. dissertation, Xidian Univ., Xi’an, China, 2010, pp. 1–12.

[5] 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.

[6] Y. Q. Zeng and Q. Chen, “Single-frame IR weak target background suppression based on improved bilateral filtering,” *Infr. Technol.*, vol. 33, no. 9, pp. 537–539, 2011.

[7] T.-W. Bae, “Small target detection using bilateral filter and temporal cross product in infrared images,” *Infr. Phys. Technol.*, vol. 54, no. 5, pp. 403–411, Sep. 2011.
P. Perona and J. Malik, “Scale-space and edge detection using anisotropic diffusion,” IEEE Pattern Anal. Mach. Intell., vol. 12, no. 7, pp. 629–634, Jul. 1990.

H. X. Zhou, Y. Zhao, H. L. Qin, S. M. Yin, G. Liu, D. Zhao, X. Yan, and S. H. Rong, “Multi-scale anisotropic diffusion equation for infrared weak targets Detection algorithm,” Actaphotonicasinaica, vol. 44, no. 9, pp. 0910002-1–0910002-5, 2015.

Q. Lin, S. C. Huang, X. Wu, and Y. Zhong, “Infrared small target detection based on nuclear anisotropic diffusion,” Intense Excitation Part. Beams, vol. 27, no. 1, pp. 0110141–0110144, 2015.

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. Electre. Eng., vol. 68, pp. 550–556, May 2018.

L. L. Wan and M. Wang, “Infrared small target detection based on partial response fusion and local contrast enhancement,” Optik, vol. 191, pp. 116–122, 2020.

J. C. Li and Z. K. Sheng, “Detection method for moving point targets in strong undulating backgrounds,” J. Huazhong Univ. Sci. Technol., vol. 43, no. 1, pp. 178–181, 2015.

H. L. Qin and Q. Wang, “Infrared small target detection using tensor based least mean square,” Comput. Electre. Eng., vol. 91, pp. 1–14, May 2021.

Y. Cao, R. Liu, and J. Yang, “Small target detection using two-dimensional least mean square (TDLMS) filter based on neighborhood analysis,” J. Infr., Millim., THz Waves, vol. 29, no. 2, pp. 188–200, 2008.

Z. Z. Li and N. L. Dong, “Adaptive filtering-based detection of weak targets in strong undulating backgrounds,” J. Instrum., vol. 25, no. 4, pp. 663–665, 2004.

J. C. Li and Z. K. Sheng, “Detection method for moving point targets in infrared undulating background,” Inf. Laser Eng., vol. 26, no. 6, pp. 8–13, 1997.

Z. J. Wang, Z. J. Yu, K. Ma, J. Wu, and J. C. Zhuge, “An adaptive median gradient cepstral weighted image filtering algorithm,” Laser Optoelectron. Prog., vol. 54, no. 121001, pp. 1210011–1210014, 2017.

Y. Yong, “Research on small target detection and recognition technology,” Ph.D. dissertation, Inst. Opt. Electron., Chin. Acad. Sci., Beijing, China, 2005, pp. 51–60.

X. Y. Jang and S. Z. Zhou, “Hidden Markov models for automatic detection and tracking of multiple targets,” Acoust. Electre. Eng., vol. 124, no. 4, pp. 29–33, 2016.

B. Y. Xu, “Research on motion target detection method based on Markov random field,” Ph.D. dissertation, Hefei Univ. Technol., Hefei, China, 2011, pp. 4–20.

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.

D. B. Wang, S. Q. Liu, X. M. Kou, and M. Hong, “MRF-based adaptive regularized infrared background clutter suppression algorithm,” J. Infr. Millim., THz Waves, vol. 28, no. 6, pp. 440–442, 2009.

Y. X. Jiang, W. X. Jin, X. T. Wang, and H. Huang, “Markov random field based motion target detection,” Comput. Electre. Eng., vol. 41, no. 15, pp. 182–184, 2010.

L. Yang, “Infrared small target detection and tracking under complex background conditions tracking algorithm research,” Ph.D. dissertation, Inst. Image Process. Pattern Recognit., Shanghai Jiao Tong Univ., Shanghai, China, 2006, pp. 41–46.

Y. He, M. Li, J. Zhang, and Q. An, “Small infrared target detection based on low-rank and sparse representation,” Inf. Phys. Technol., vol. 68, pp. 98–109, Jan. 2015.

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.

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.

J. L. Fan, Y. M. Gao, Z. H. Wu, and L. Li, “Infrared dim small target detection technology based on RPCA,” in Proc. Int. Conf. Electron. Inf. Technol. Intellectualization, vol. 3, 2017, pp. 748–752.

C. Wang and S. Qin, “Adaptive detection method of infrared small target based on target-background separation via robust principal component analysis,” Inf. Phys. Technol., vol. 69, pp. 123–135, Mar. 2015.

Q. Hou, Z. Wang, F. Tan, Y. Zhao, H. Zheng, and W. Zhang, “RISTDNNet: Robust infrared small target detection network,” IEEE Geosci. Remote Sens. Lett., vol. 19, pp. 1–5, 2021.

J. Bai, H. Zhang, and Z. 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.

C. H. Liu, Random Process. Hubei, China: Huzhong Univ. of Science and Technology Press, 2017, p. 180.