Article

DP–MHT–TBD: A Dynamic Programming and Multiple Hypothesis Testing-Based Infrared Dim Point Target Detection Algorithm

Jinming Du [1,*], Huanzhang Lu [1], Luping Zhang [1], Moufa Hu [1], Yingjie Deng [1], Xinglin Shen [1], Dongyang Li [2] and Yu Zhang [1]

[1] National Key Laboratory of Science and Technology on Automatic Target Recognition, College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China
[2] Department of Military Education, College of Military Basic Education, National University of Defense Technology, Changsha 410073, China
* Correspondence: dujinming16@nudt.edu.cn

Abstract: The detection and tracking of small targets under low signal-to-clutter ratio (SCR) has been a challenging task for infrared search and track (IRST) systems. Track-before-detect (TBD) is a widely-known algorithm which can solve this problem. However, huge computation costs and storage requirements limit its application. To address these issues, a dynamic programming (DP) and multiple hypothesis testing (MHT)-based infrared dim point target detection algorithm (DP–MHT–TBD) is proposed. It consists of three parts. (1) For each pixel in current frame, the second power optimal merit function-based DP is designed and performed in eight search areas to find the target search area that contains the real target trajectory. (2) In the target search area, the parallel MHT model is designed to save the tree-structured trajectory space, and a two-stage strategy is designed to mitigate the contradiction between the redundant trajectories and the requirements of more trajectories under low SCR. After constant false alarm segmentation of the energy accumulation map, the preliminary candidate points can be obtained. (3) The target tracking method is designed to eliminate false alarms. In this work, an efficient second power optimal merit function-based DP is designed to find the target search area for each pixel, which greatly reduces the trajectory search space. A two-stage MHT model, in which pruning for the tree-structured trajectory space is avoided and all trajectories can be processed in parallel, is designed to further reduce the hypothesis space exponentially. This model greatly reduces computational complexity and saves storage space, improving the engineering application of the TBD method. The DP–MHT–TBD not only takes advantage of the small computation amount of DP and high accuracy of an exhaustive search but also utilizes a novel structure. It can detect a single infrared point target when the SCR is 1.5 with detection probability above 90% and a false alarm rate below 0.01%.

Keywords: dim small target detection; dynamic programming; second power optimal merit function; two-stage multiple hypothesis testing; infrared image

1. Introduction

Target detection techniques in infrared (IR) images have played an important role in military and civil applications. IR imaging of long-range space targets, which can be modeled as point targets, contain little information about shape and texture. These point targets are usually buried in strong clutter. Therefore, detecting dim small IR targets remains a challenging problem [1–4]. During the past few decades, scholars have performed many studies on dim small-target detection. Existing IR small-target detection methods can be roughly divided into three categories: TBD methods, detect-before-track (DBT) methods and deep learning (DL)-based methods.

All methods have advantages and disadvantages. DBT methods, such as the classical local contrast measure (LCM) [5], local intensity and gradient (LIG) [6] and IR patch-image model (IPI) [7], concentrate on detecting targets in a single frame based on prior information. These methods are usually based on the image data structure and are designed to enhance the target, suppress the background, expand the contrast between the target and the background to improve the recognizability of the target, and then realize the detection of the target. Generally, this kind of method has the advantages of low complexity, high efficiency and easy hardware implementation. However, when the target SCR is low (SCR $\leq$ 1.5), many false alarms are produced, and the detection accuracy is decreased. Under low SCR conditions, TBD methods, such as 3-D matched filtering [8], Hough transform (HT) [9–11], DP [12] and MHT [13], are more stable and demonstrate better detection performance because they usually utilize spatial and temporal information by capturing the target trajectory. However, the computational cost and storage requirements of this kind of method inevitably increase, making practical applications difficult. Recently, DL methods, such as the single shot multibox detector (SSD) [14], you only look once (YOLO) [15], and Faster-RCNN [16], have already achieved remarkable progress in the target detection field. Relying on their strong feature extraction and generalization ability, neural network-based infrared target detection methods, such as the spatial–temporal feature-based detection framework [1], target-oriented shallow–deep feature (TSDF)-based detection method [2] and TBC-Net [17], have attracted increasing attention. In the previous work [1,2], some DL-based IR target detection algorithms were analyzed in detail. When the target has certain shape or texture information (size greater than 2 $\times$ 2) and the characteristics of background clutter are spatially nonstationary but temporally stationary, as with the images in Figure 1a, DL methods could play a role in the IR target detection field by assisting some techniques, such as shallow–deep feature fusion, spatial–temporal information extraction, and a reasonable sample selection strategy. However, when the target occupies only one pixel and the background clutter is temporally nonstationary, as with the images in Figure 1b, it is difficult for the DBT and DL methods to detect targets mainly for the following reasons. (1) Because the target SCR is very low, there are many pixels similar to the target in a single frame image, and the contrast between the target and the background is very low. Even the human eye cannot distinguish the target from the background by relying only on gray information. (2) The spatial–temporal information extraction method applied to the temporally stationary background has difficulty playing a role in the temporally nonstationary background, resulting in the neural network not learning the useful features. Taking the spatial–temporal information extraction method in [2] as an example, to suppress the background and extract the moving features of the target, the operations of frame subtraction and addition are used. When applied to an infrared sequence with a temporally nonstationary background, this kind of operation will not suppress the background but will reduce the SCR of the target and increase the difficulty of detection. This work focuses on the dim point target detection task. The problem analysis, methods and experiment part in this work focus on the 1 $\times$ 1 point target. The key to solving the problem of infrared point target detection under low SCR conditions is to adopt an appropriate TBD method; that is, first accumulate the energy of the target along the target trajectory to improve the SCR, and then detect the target.

In recent years, DP-TBD has become a hot research direction. The core idea of DP is to transform a complex multistep decision process into multiple single-stage decision processes and then optimize each decision to obtain a global optimal solution. The algorithm has clear principles, and the operation is based on the pixel level. This method is relatively small in computation and storage and easy to implement in hardware. There have been several improved DP-TBD methods in recent years. Direction information-based DP methods [18–20] have been proposed to reduce the number of pseudo trajectories under a strong clutter background and reduce the search range to decrease the diffusion effect of the target energy in the accumulation process. As the coverage of the target maneuvering range is limited to the fixed transition step, ISTS-DP-TBD (TBD algorithm with an improved

state transition set) [21] was proposed. The state search efficiency of the maneuvering target is improved with optimization of the state transition strategy. PC-DP-TBD (DP-TBD method using parallel computing) [22] was proposed to solve the problem that the adjacent targets may interfere with each other and the computational complexity is increased with the number of targets. Although some progress has been made in DP-TBD, there are several problems to be solved [23,24]:
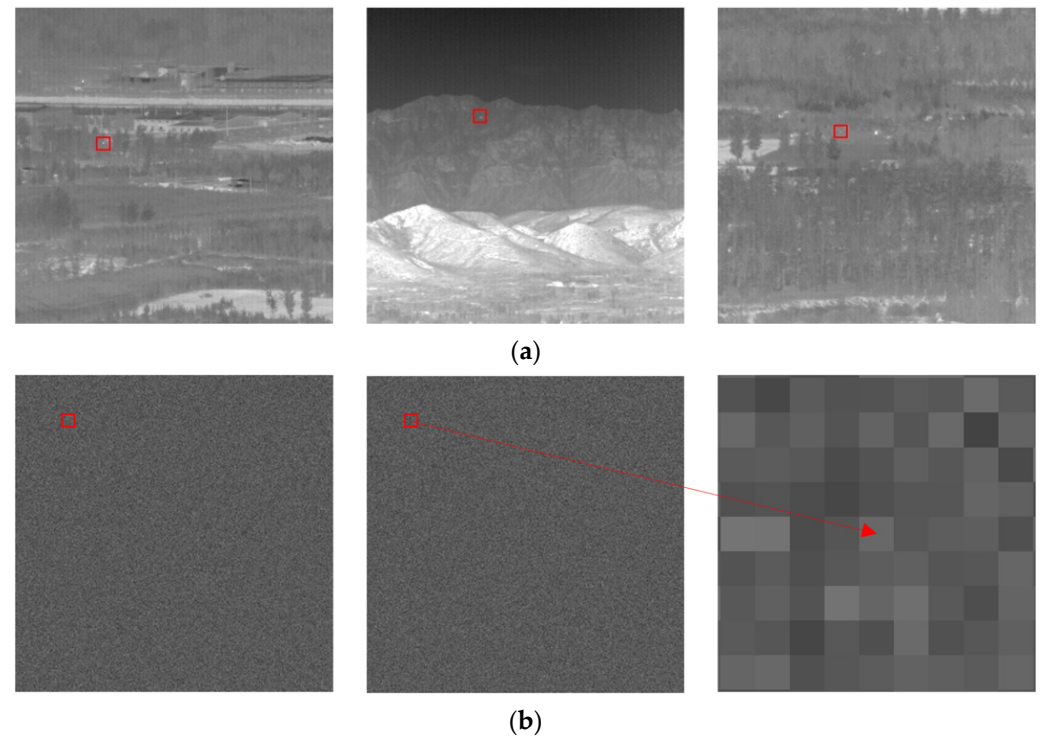

Figure 1. Representative infrared images with various backgrounds. (**a**) Background clutter is spatially nonstationary but temporally stationary; (**b**) Background clutter is temporally nonstationary, the area in red frame is the target and its neighbor pixels.

(1) The algorithm is affected by some parameters, such as the number of state transitions, the number of accumulation frames and the type of merit function value function. The selection of parameters has a great impact on the performance of the algorithm. It is difficult to find a set of parameters that is suitable for different backgrounds.

(2) Agglomeration effect. When tracking and detecting a target, the energy is accumulated along the possible trajectories. After K frame accumulation, the merit value of the target state increases, and in most cases this value is the largest. However, during the accumulation process, the energy of the target in the previous observation frame will diffuse to the next observation frame, forming a cluster of observation areas with similar energy. This is also known as the 'agglomeration effect'. This effect becomes increasingly serious with increasing accumulation frames. When the SCR is low, after accumulation, the agglomeration effect will occur not only in the target but also in noise with strong energy, and even the diffusion of noise will be greater than that of the target. In this case, the state of this kind of noise is stronger than that of the target, which is not conducive to target detection or tracking. In addition, when there are multiple adjacent targets, there will be many trajectories with similar energy after energy accumulation. At this time, the agglomeration areas of different targets may overlap, causing difficulties in differentiating different targets. This is also the reason why most algorithms require the targets to be neither adjacent nor intersected.

Due to the above problems, compared with exhaustive search methods, the DP-TBD method performance can be reduced by 3 dB [12]. When the energy of the detected

target drops to a certain extent, even if the number of accumulated observation frames is increased, the detection performance is poor. Even for some improved methods [21,25–27], the detection performance will be greatly reduced when the SCR is lower than 1.8.

Therefore, to achieve high-precision point target detection under low SCR (SCR $\leq$ 1.5), an exhaustive search method needs to be adopted. The real target trajectory can be found by searching as many trajectories as possible. The classical representative algorithm is MHT. However, the number of trajectories increases exponentially with the number of accumulation frames, causing tremendous computation and storage costs, so it is difficult to realize the accumulation of many frames. Therefore, the key to exhaustive search methods is to design a reasonable strategy to reduce the number of redundant trajectories.

In summary, to detect point targets under very low SCR (SCR $\leq$ 1.5), it is necessary to design a TBD algorithm that can reduce the trajectory hypothesis space. Based on this core idea, the DP–MHT–TBD is proposed. The main contributions are summarized as follows:

(i) A second power optimal merit function-based DP method is designed. It can find the target search area whose range is 90° (all range is 360°) with high confidence and can reduce the trajectory hypothesis space by $\frac{3}{4}$ for each pixel.

(ii) A two-stage MHT model is designed. It can reduce the trajectory hypothesis space exponentially, be operated in parallel, avoid pruning for the tree-structured trajectory space, greatly reduce the computational cost and save the storage space.

(iii) The proposed DP–MHT–TBD improves the engineering application of the TBD method. It takes advantage of the DP and exhaustive search, utilizes a novel structure, and can detect point targets when the SCR is 1.5 with a probability of more than 90% and a false alarm rate of less than 0.01%.

The remainder of this article is organized as follows. In Section 2, the methodology is described in detail. An overview of the proposed DP–MHT–TBD detection framework in IR images is given. In Section 3, simulation experiments and an analysis of the results are presented. In Section 4, discussions are given. In Section 5, conclusions are given.

2. Methodology

The proposed DP–MHT–TBD, which consists of a second power optimal merit function-based DP, a two-stage MHT model and a target tracking method, is described in detail.

In Section 2.1, through qualitative and quantitative analysis, the agglomeration effect caused by the diffusion of energy in the process of accumulation is studied. It is concluded that using DP to find the target search area is more reliable than directly detecting the target. Based on this idea and the property of the power function, the second power optimal merit function-based DP is designed to obtain a 90° target search area for each point, reducing the trajectory hypothesis space by $\frac{3}{4}$.

In Section 2.2, a novel parallel MHT is first designed. For each point, via the proposed MHT, the tree-structured trajectory space can be quickly obtained, and all trajectories can be processed in parallel. The final accumulated energy can be obtained by using only one testing. Compared with the classical MHT, a one-by-one search of root nodes in all trajectories is avoided, multistage testing is avoided, and pruning for the tree-structured trajectory space is avoided, greatly reducing the computational complexity. In addition, to mitigate the contradiction between the redundant trajectories and the requirements of more trajectories under low SCR, a two-stage strategy is designed. This not only ensures that the target energy is accumulated to a certain extent under low SCR but also reduces the number of trajectories exponentially. The two-stage MHT not only reduces the amount of calculation but also saves on the storage requirements.

In Section 2.3, a target tracking method that can eliminate false alarms by deleting discontinuous trajectories is introduced.

In Section 2.4, the DP–MHT–TBD detection framework is shown; it is a sequential detection process.

2.1. Second Power Optimal Merit Function-Based DP

2.1.1. Basic DP Model

In infrared images, from the 1th measurement to kth measurement, the target observation model is $z_{1:k} = \{z_1, z_2 \cdots z_k\}$.

$$z_k = x_k + \eta_k \tag{1}$$

where k denotes the kth measurement and η denotes observation noise, which obeys a Gaussian distribution $N \sim (\mu, \sigma)$. The mean value is μ, and the standard deviation is σ. x_k is the state vector of the target at the kth measurement frame, which can be described as:

$$x_k = \left[x_k, \dot{x}_k, y_k, \dot{y}_k, I_k \right] \tag{2}$$

where (x_k, y_k) denotes the position of the discrete target state in the kth infrared frame on the X–Y plane, $\dot{x}_k$ and $\dot{y}_k$ denote the velocity toward the X and Y axes, respectively, and I_k denotes the gray value.

In the IR field, the target energy accumulation is based on the principle of ballistic trajectory integral:

$$E = \int_c ede > \dot{E} = \int_{\dot{c}} ede \tag{3}$$

where e denotes the energy of the target or noise, c and $\dot{c}$ denote the trajectory of the target and the noise, respectively. The above formula means that if the energy is accumulated along the trajectory of the target in the IR sequence, the accumulated energy must be greater than the energy accumulated along any other trajectory. Therefore, when using the DP method to accumulate the energy of IR small targets, the optimization process opt and stage merit function ω are usually taken as the maximum function *max* and gray value I, respectively. At this time, the DP-TBD model is:

$$f_k(x_k, y_k) = \max_{(x_k, y_k) \in D_k} I(x_k, y_k) + f_{k-1}(x_{k-1}, y_{k-1}) = \max_{D_k}\{I(x_k, y_k) + \max_{D_{k-1}}[I(x_{k-1}, y_{k-1}) + \cdots + \max_{D_2}[I(x_2, y_2) + f_1(x_1, y_1)]]\} \tag{4}$$

$$f_1(x_1, y_1) = I(x_1, y_1) \tag{5}$$

$$D_k(x_k) = \{(x_{k-1}, y_{k-1}) |\ (x_k - l \leq x_{k-1} \leq x_k + l), (y_k - l \leq y_{k-1} \leq y_k + l)\} \tag{6}$$

where I denotes the gray value, (x_k, y_k) denotes image coordinates in the kth frame, D_k denotes the state transition set consisting of possible states at time $k - 1$; it refers to the motion range of the target between frames and is determined by the position and velocity of the target. l denotes the number of pixels corresponding to the velocity.

According to (4) and (5), after k frame energy accumulation, the energy accumulation map **I** can be obtained. Because the accumulated energy of the noise may be greater than that of the target, the threshold Th is selected according to a certain false alarm rate to segment **I** to obtain the candidate point set X.

$$X = \{(x, y) | \mathbf{I}(x, y) Th\} \tag{7}$$

After the candidate points are obtained, the noise points can be further eliminated by track-association detection or other methods, and the final reserved point is taken as the target.

2.1.2. Agglomeration Effect

From the above analysis, it can be seen that the type of merit function and the size of the state transition set will affect the performance of DP methods. In practical applications, due to the lack of prior information about the target's moving direction, the hypothesis is usually that: the transition of the target state from the $k - 1$th to the kth frame is usually

of equal probability. This means that the target energy in the $k-1$th frame diffuses to the neighborhood of the corresponding position in the kth frame with equal weight. The other pixels in this neighborhood are noise points, resulting in a trajectory containing both target and noise points. In addition, if the target SCR is very low, then the energy of the noise with strong energy will have a similar diffusion phenomenon. Finally, many bright blocks, which are centered on the target or noise with strong energy, appear on the energy accumulation map. This is called the agglomeration effect.

To solve the problem of the agglomeration effect, direction information-based DP methods [18–20] have been proposed. However, as long as the target is not accumulated in strict accordance with the real trajectory, there will be diffusion of the target energy. Therefore, due to the lack of prior information about the target moving direction, the energy diffusion problem can only be alleviated.

To determine whether there is regularity that can be used, qualitative and quantitative analyses of the energy diffusion were carried out.

1. *Qualitative Analysis*

As in Figure 2a, the target position in the current frame is $O(x_o, y_o)$, and the target trajectory is the red curve in the XOY area with O as the origin point. When O is the point to be detected, energy accumulation based on the DP method is carried out in the XOY area. The energy accumulation and diffusion process of the target is shown in Figure 2a, where the red dot indicates the result of energy accumulation along the target track during the accumulation process, and the blue arrow indicates the diffusion of the target energy. The larger the dot is, the greater the energy accumulated. The wider the arrow is, the more energy diffused, and the black triangular points A, B and C indicate the noise points with strong energy in the current frame. When the number of accumulated frames is very small (as in Figure 2a, $t = 1$), the accumulated target energy and the diffused energy are very small, which can be ignored. With a gradual increase in the accumulated frames (as $t > k$), the accumulated target energy and the diffused energy increase gradually, which cannot be ignored. After the accumulation of previous $n-1$ frames, the accumulated energy of the target is $E(t = n-1)$. From $t = n-1$ to $t = n$, the previous accumulated energy diffuses to the target point O (with energy $E_O(t = n)$) and the nearby noise point set $\{A, B, C \cdots\} \in N$ (with energy $E_N(t = n)$). At $t = n$, the energy of the target and the neighborhood noise is:

$$\begin{cases} E(target) = E(t = n-1) + E_O(t = n) \\ E(noise) = E(t = n-1) + E_{\{A, B, C\cdots\}\in N}(t = n) \end{cases} \quad (8)$$
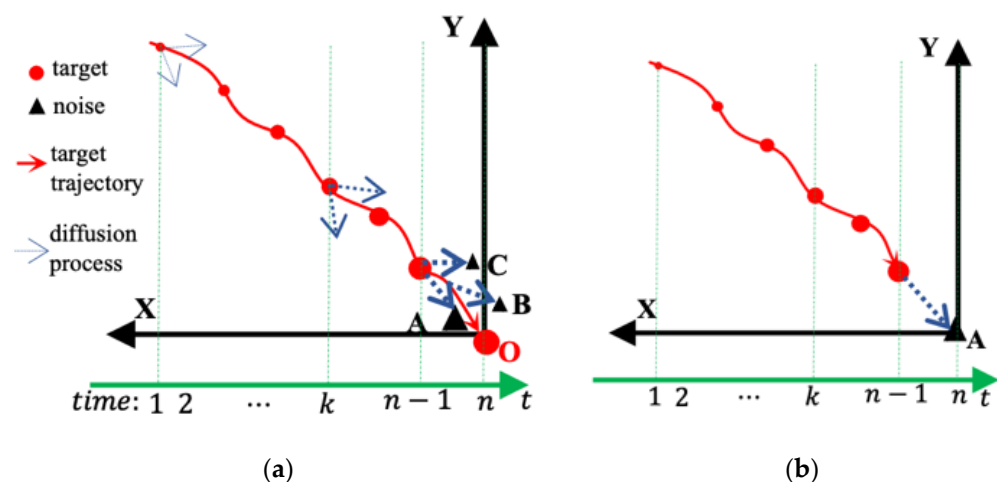

Figure 2. Target energy accumulation and diffusion in the DP process. (**a**) Energy accumulation and diffusion process of target; (**b**) Diffusion effect for noise.

According to (8), after energy accumulation, the probability that the energy of the target point is greater than that of the nearby noise point ($P_r\{E(target) > E(noise)\}$) depends on the energy of the target and its nearby points in the current frame. When the target SCR is large (SCR > 3), $P_r\{(E(target) > E(noise)\}$ is large because $P_r\{E_O(\text{t} = \text{n}) > E_{\{A, B, C\cdots\}\in N}(t = n)\}$. However, with a decrease in the SCR, $P_r(E(target) > E(noise)$ decreases because $P_r(E_O(t = n) > E_{\{A,B,C\cdots\}\in N}(t = n))$ decrease, meaning that an increasing number of noise points are enhanced and more false alarms are generated.

2.   *Quantitative Analysis*

The analysis focuses on a popular DP method in which the maximum function *max* is the optimization process and the gray value is the stage merit function. Suppose that before processing, the noise obeys a Gaussian distribution $N \sim (\mu, \sigma)$, the noise point energy is μ, the target point energy is T, and the number of accumulated frames is n.

For each pixel P in the current image, as in Figure 3a, 4 accumulated energy values $I_{xpy}(P)$, $I_{ypz}(P)$, $I_{zpw}(P)$, and $I_{wpx}(P)$ can be obtained after performing the DP method in 4 different areas {XPY, YPZ, ZPW, WPX}. There are three types of pixels in an image (see (9)): pixel P belongs to the target, noise is far from the target (*noise_far*), and noise is near the target (*noise_near*).

$$P \in \{target,\ noise\_near,\ noise\_far\} \tag{9}$$
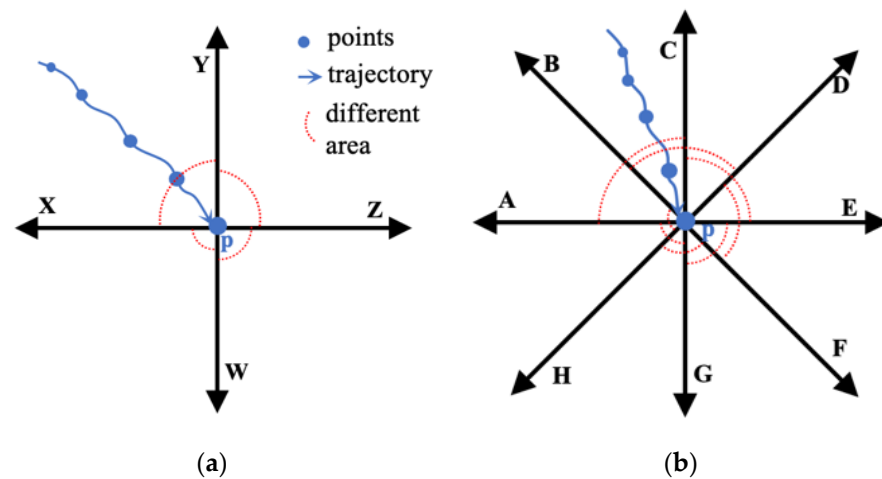

(a)   **(b)**

Figure 3. DP-TBD in different areas. For each pixel, (**a**) DP process in 4 areas; (**b**) DP process in 8 areas.

Assume that when pixel P is the target, as in Figure 3a, the target trajectory C belongs to the XPY area. After energy accumulation, the final energy of P is:

$$I(P) = max\{I_{xpy}(P),\ I_{ypz}(P),\ I_{zpw}(P),\ I_{wpx}(P)\} = \begin{cases} I_{xpy}(P) = nT, & P \in target,\ C \in XPY; \\ I_{xpy}(P) = kT + (n-k)\mu, & P \in noise\_near,\ C \in XPY; \\ n\mu, & P \in noise\_far, C \in 4\ areas\ \text{with}\ equal\ probability. \end{cases} \tag{10}$$

When pixel P belongs to noise that is near the target, in the trajectory of this kind of point, the first k points belong to the targets, while the last $n - k$ points belong to the noise. After accumulation, if SCR > 3, according to the 3σ rule of thumb, then the probability that the target energy is larger than the noise is more than 99%. That is, on the energy accumulation map, the mean value of the probability distribution of the target point is greater than that of the noise point. However, when SCR < 3, this relation might be incorrect. As in Figure 2b, for noise point A with strong energy ($\mu > T$), $k = n - 1$, after accumulation $(kT + (n-k)\mu) > nT$, which means $E(noise) > E(target)$. The lower the SCR is, the more noise points are enhanced after energy accumulation.

Another conclusion can be drawn from (10): If $P_t \in target$, then the corresponding trajectory C_t of P_t contains n target points; if $P_{near} \in noise\_near$, then the corresponding trajectory C_{near} of P_{near} contains k target points and $n - k$ noise points; if $P_{far} \in noise\_far$, then the corresponding trajectory C_{far} of P_{far} contains n noise points. After accumulation, whether the energy of P_t is larger than that of P_{near}, the trajectories C_t and C_{near} are in the same area. As shown in Figure 2, the trajectories of target point O and noise point A belong to areas *XOY* and XAY, respectively. Because the categories of points O and A are not known in advance, O and A are treated as one point P in the current image; thus, *XOY* and XAY are *XPY*. Therefore, as long as the energy of C_t or C_{near} is larger than that of trajectory C_{far}, the target search area can be found via the following backtracking method:

$$max\left(I_{map}\right) \rightarrow P_{max} \in \left\{ \begin{array}{c} target \\ noise\_near \end{array} \right. \rightarrow C \in \left\{ \begin{array}{c} XPY \\ XPY \end{array} \right. \rightarrow XPY \tag{11}$$

First, for the accumulation map I_{map}, the point with the largest energy P_{max} can be found through the maximum function max. When the SCR is not very low, the following premise (12) is true. As above, the point P_{max} belongs to *target* or the target's nearby noise (*noise_near*).

$$\{nT \ or \ (kT + (n - k)\mu)\} > \ n\mu \tag{12}$$

Then, through P_{max} and (10), find the area where the corresponding trajectory C_{max} belongs and take the area as the target search area for each pixel.

Summarizing the above analysis, when using the DP method for energy accumulation under low SCR conditions, (1) the energy of the target will diffuse to the noise points and produce many false alarms; (2) if there is only one target, then the area to which the target trajectory belongs can be found according to the point with the largest energy (as the process in (11)), but this point cannot be taken as the target because it probably belongs to the noise.

2.1.3. Second Power Optimal Merit Function

As in the above analysis, the lower the SCR is, the more similar the probability distribution of the target and noise, and the smaller the probability of $nT > n\mu$ is, the larger the probability that the point with the largest energy P_{max} belongs to the noise far from the target (*noise_far*). According to (10), when $P_{max} \in noise\_far$, it is impossible to find where the trajectory belongs because it belongs to 4 areas with equal probability. This means that the accuracy of finding the target trajectory according to (11) decreases under low SCR conditions.

Therefore, to solve the above problems, the second power optimal merit function-based DP method is proposed. It consists of four steps.

(1) As in Figure 3b, for each pixel P, 8 different search areas are set: $\{APC, BPD, CPE, DPF, EPG, FPH, GPA, HPB\}$, the range of each area is $90°$, all range is $360°$, and the overlap angle between each two areas is $45°$.

(2) In 8 search areas, for each pixel P, the DP method is performed, in which max is the optimization process and the gray value is the stage merit function to obtain 8 accumulated energy values $I(P)$.

$$I(P) = \{I_{APC}, \ I_{BPD}, \ I_{CPE}, \ I_{DPF}, \ I_{EPG}, \ I_{FPH}, \ I_{GPA}, \ I_{HPB}\} \tag{13}$$

(3) For each pixel P, the optimal value is calculated according to the second power optimal merit function (14). Assume that when pixel P is the target, as in Figure 3b, target trajectory C belongs to APC and BPD.

$$I_{opt}(P) = max\{I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8\} = max\{I_{APC} \times I_{BPD}, I_{BPD} \times I_{CPE}$$
$$, I_{CPE} \times I_{DPF}, I_{DPF} \times I_{EPG}, I_{EPG} \times I_{FPH}, I_{FPH} \times I_{GPA}, I_{GPA} \times I_{HPB}, I_{HPB} \times I_{APC}\}$$
$$= \begin{cases} I_1 \approx (nT)^2, & P \in target, C \in \{APC, BPD\}; \\ I_1 \approx (kT + (n-k)\mu)^2, & P \in noise\_near, C \in \{APC, BPD\}; \\ (n\mu)^2, & P \in noise\_far, C \in 8 \ areas \ with \ \text{equal probability.} \end{cases} \tag{14}$$

(4) As in (11), the target search area is found via the following backtracking method.

$$max(I_{map}) \rightarrow P_{max} \in \begin{cases} \{I_i\} \\ i \in \{1, 2 \cdots 8\} \end{cases} \rightarrow i \rightarrow area\ pair \tag{15}$$

where the *area pair* denotes the areas with overlaps.

$$area\ pair = \{1 : APC\&BPD,\ 2 : BPD\&CPE,\ 3 : CPE\&DPF,\ 4 : DPF\&EPG$$
$$5 : EPG\&FPH,\ 6 : FPH\&GPA,\ 7 : GPA\&HPB,\ 8 : HPB\&APC\} \tag{16}$$

where APC & BPD denotes the areas APC and BPD.

The proposed second power optimal merit function essentially transforms the comparison of the target and noise trajectory energy from the first power (as in (12)) to the following second power.

$$\left\{(nT)^2 \ or \ (kT + (n-k)\mu)^2\right\} > (n\mu)^2 \tag{17}$$

As long as premise (17) is true, this method can be used to find the area to which the target trajectory belongs.

According to the property of the power function:

$$y = x^\alpha \tag{18}$$

If $x > 1$, with increasing α ($\alpha > 0$), y increases; if $x > 1$ and $\alpha > 1$, with increasing x, the slope of the curve and the derivative increase, meaning that a small increase in x will lead to a large increase in y.

According to the principle of ballistic trajectory integral (3), after accumulation, the probability (P_r) that the target energy is greater than the noise energy is larger than the probability that the noise energy is greater than the target energy:

$$P_r\{E(target) > E(noise)\} > P_r\{E(noise) > E(target)\} \tag{19}$$

This means

$$P_r\{nT > n\mu\} > P_r\{n\mu > nT\} \tag{20}$$

When the accumulation frame number $n > 1$, $nT, n\mu > 1$, from $\{nT, n\mu\}$ to $\left\{(nT)^2, (n\mu)^2\right\}$, the increase in nT is larger than $n\mu$, meaning there is a larger increase for the y of nT, according to the property of the power function and (20), the following relation can be deduced:

$$P_r\{(nT)^2 > (n\mu)^2\} > P_r\{nT > n\mu\} \tag{21}$$

Similarly, for nT and $kT + (n-k)\mu$:

$$P_r\{(kT + (n-k)\mu)^2 > (n\mu)^2\} > P_r\{(kT + (n-k)\mu) > n\mu\} \tag{22}$$

To find the area to which the target trajectory belongs, premises (12) and (17) must be true. Based on (21) and (22), it can be concluded that the probability of premise (17) being true is larger than that of premise (12), meaning the second power optimal merit function-based method is more reliable than the first power, in theory.

Similar to (11), the point with the largest energy P_{max} in (15) can only be used to find the target search area but cannot be taken as the target because it probably belongs to the noise.

As in Figure 3b, point P is the target, and the target trajectory is the blue curve in the APC area with P as the origin point. According to the proposed second power merit function-based DP ((12)–(15)), the area pair $\{APC\&BPD\}$ can be obtained. Both APC and BPD can be used as the target search area because both areas contain the target trajectory.

To accurately locate the position of the target, for each pixel to be detected, the next task is to find the possible trajectories in the target search area. Compared with an exhaustive search, this approach reduces the trajectory hypothesis space by $\frac{3}{4}$ for each pixel.

2.2. Two-Stage MHT Model

2.2.1. The Proposed Parallel MHT Model

After the target search area is obtained, for each pixel in the current frame, the possible trajectories in the target search area need to be found. Since the sampling frequency of the infrared detector cannot completely match the moving speed of the target, if the target moves in a straight line in the real 2D space, then the imaging of the target is probably not a straight line in the discretized 2D image space. As in Figure 4a, in the 2D space the real trajectories that pass through point P' are the red straight lines. If these lines are mapped to the 2D image space, then the trajectories are no longer straight lines but tree-structure-like curves that pass through point P, and the nodes of the curves are different pixels, as shown in Figure 4b. The number of trajectories (m) in the XPY area are related to the length of the sequence n.
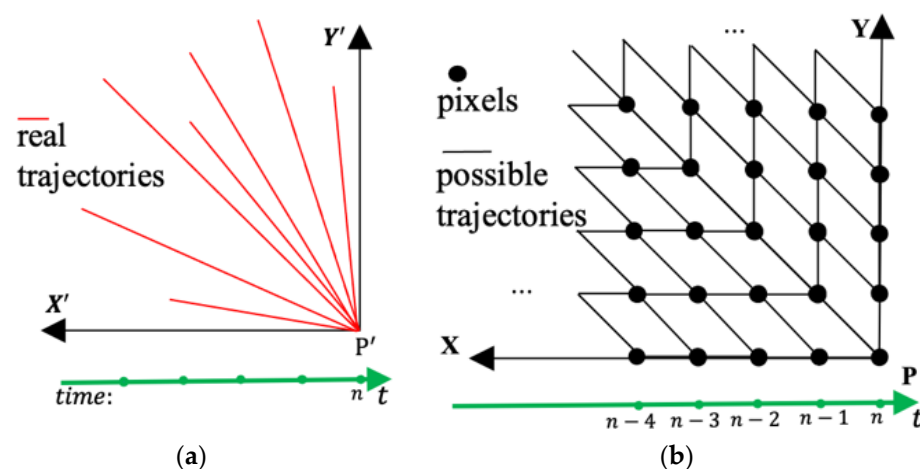
$$m = 2^n - 1 \tag{23}$$

(a) (b)

Figure 4. Possible trajectories in the target search area: (**a**) the trajectories in 2D space and (**b**) the trajectories in 2D image space. It is a tree-structured trajectory space.

These trajectories have the following property. For the points in different positions on the current image, only the coordinates of nodes in each trajectory are different; the shape of the tree-structured trajectory space and the number of all trajectories are the same. This is called the 'trajectory shape similarity property' in this paper. To describe this property, the multiple trajectory hypotheses model ΔH_{tree} is constructed. As in Figure 5, ΔH_{tree} is a $m \times n$ 2-D matrix. The relative coordinates of all trajectories are saved in ΔH_{tree}. As in Figure 5, ΔX_n^m denotes the relative coordinates $(\Delta x, \Delta y)$ of the nth point in the mth trajectory. For the point at position $P(x, y)$, the tree-structured trajectory space can be described by $H_{(x,y)}$.
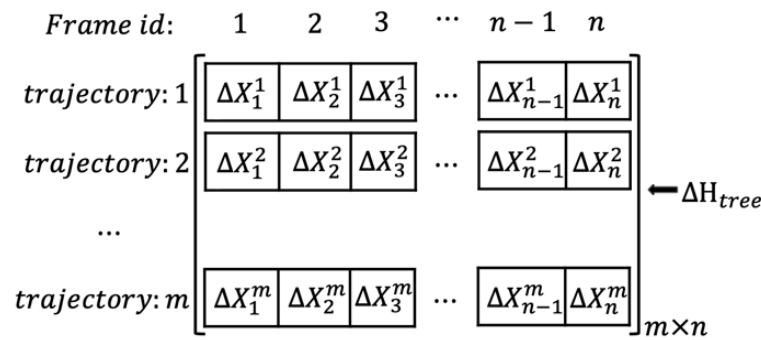
Figure 5. Proposed multiple trajectory hypotheses model $\Delta\mathrm{H}_{tree}$. ΔX_n^m denotes the relative coordinates $(\Delta x, \Delta y)$ of the nth point in the mth trajectory.

$$H_{(x,y)} = \Delta\mathrm{H}_{tree} \oplus (x, y) \tag{24}$$

where $\oplus$ denotes adding the coordinates of the current point (x, y) to each node in $\Delta\mathrm{H}_{tree}$. For point P, after obtaining trajectories, the testing process is:

$$E(P(x,y)) = \max_{1:m}\left\{ S(1:n, \mathrm{M}, \mathrm{N}) \circledast H_{(x,y)} \right\} = \max_{1:m}\{I_1, I_2, I_3, \cdots I_m\} \tag{25}$$

where S denotes the infrared sequence with length n, the size of each frame is $\mathrm{M} \times \mathrm{N}$, and $1 \le x \le \mathrm{M}, 1 \le y \le \mathrm{N}$. $\circledast$ denotes matching the node coordinates on each trajectory with the image gray value and accumulating energy for each trajectory. I_m denotes the energy of the mth trajectory. $\max\limits_{1:m}$ denotes the testing process, which takes the trajectory with the largest energy as the trajectory for point P and the corresponding energy as the accumulated energy $E(P(x,y))$. Equations (24) and (25) constitute the proposed MHT model. There are the following advantages:

(1) The trajectory search problem is simplified. Because of the 'trajectory shape similarity property', the tree-structured trajectory space for each point can be saved in $H_{(x,y)}$ using (24). In this way, a one-by-one search of root nodes in all trajectories is avoided, reducing the calculation cost and improving the efficiency of obtaining trajectory space. The larger the image resolution is, the larger reduction in calculations.

(2) The energy accumulation process can be implemented in parallel. The process of energy accumulation is independent of the trajectory order, so all trajectories can be operated in parallel. In actual operation, all trajectories can be allocated to different threads of different central processing units (CPUs). This means that (25) can be performed in different threads. This is very beneficial for improving the operability of the TBD method.

(3) Compared with the classical MHT method [13], the proposed MHT method considers both accuracy and efficiency. Under low SCR conditions, to ensure the detection probability, more nodes should be reserved in the early stage, and the redundant trajectories should be deleted in the later stage when the accumulated energy reaches a certain extent. The classical MHT is a multistage testing process, and some nodes will be deleted in each stage, causing the loss of accuracy. To avoid this loss, in the first stage of proposed MHT model, all trajectories and nodes are saved. In addition, the whole process of classical MHT includes the establishment of a tree-structured list and the judgment, insertion and deletion of the tree nodes. The computational complexity of the operation of the tree list is high. In the proposed MHT model, the operations involving the tree list are only (24). The only testing process is $\max\limits_{1:m}$. The computational complexity is very low.

2.2.2. Two-Stage MHT Model

In the process of energy accumulation, the lower the target SCR is, the longer the length of the trajectory needed, and the larger the number of trajectories, meaning the larger the trajectory hypothesis space. As in (23), the number of trajectories m increases

exponentially with increasing sequence length n. For example, from $n = 15$ to $n = 30$, m changes from 3.3×10^4 to 1.1×10^9. The trajectory length is only doubled, and the number of trajectories is increased by approximately 3.3×10^4 times, which brings much more computing and storage consumption than the benefits of energy accumulation. The main reason for this problem is the redundancy of the trajectory. As shown in Figure 6a, trajectory OP is one trajectory that accumulates from point O ($t = 1$) to point P ($t = n$). For point P, when $t = n + 1$, there are 3 possible trajectories: OPQ, OPR and OPS. OP is the overlapping trajectory. If OP is the target trajectory and n is relatively large, the energy of the target has been accumulated to a certain extent at point P. The OPQ, OPR and OPS contain the target trajectory OP, so the energy accumulated along the three trajectories may be similar. At this time, OPQ, OPR and OPS are redundant trajectories, and two of them should be removed. When n is small, the energy of the target has not been fully accumulated when reaching point P. In this case, any trajectory cannot be eliminated to ensure that the target trajectory is retained. Therefore, the premise of eliminating redundant trajectories is that the target energy has accumulated to a certain extent.
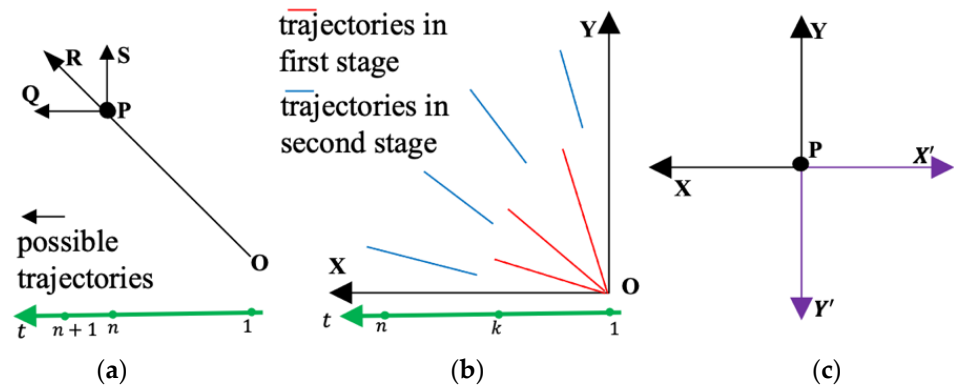

Figure 6. Redundant and sparse trajectories: (**a**) redundant trajectories; (**b**) sparse trajectories; and (**c**) target search area and its opposite area.

The two-stage search-based energy accumulation method can be used to reduce redundant trajectories, as shown in Figure 6b. In the first stage, from $t = 1$ to $t = k$, O is taken as the starting point, and energy is accumulated along the sparse trajectories. After accumulation, some trajectories are missed because the trajectories are sparse. In the second stage, starting from the missing point in the first stage, the energy is accumulated from $t = k$ to $t = n$. After a two-stage search, not only is the number of redundant trajectories reduced but also the search of multiple trajectories is realized.

Therefore, to mitigate the contradiction between the redundant trajectories and the requirements of more trajectories under low SCR, the two-stage MHT model is proposed. For each point $P(x, y)$ in the current frame, the two-stage MHT model is used to obtain the accumulated energy. It consists of three steps.

(1) In the first stage, from $t = 1$ to $t = k$, in the target search area XPY, as in (24) and (25), the proposed MHT model is used to obtain the tree-structured trajectory space and the accumulated energy $E_1(P(x, y))$.

$$E_1(P(x, y)) = \max_{1:k} \left\{ S(1 : k, \mathrm{M}, \mathrm{N}) \circledast H_{(x,y)} \right\} \tag{26}$$

(2) In the second stage, from $t = k$ to $t = n$, in the opposite area $X'PY'$, as in Figure 6c, the proposed MHT model is used to obtain $E_2(P(x, y))$.

$$E_2(P(x, y)) = \max_{k:n} \left\{ S(n : k, \mathrm{M}, \mathrm{N}) \circledast H'_{(x,y)} \right\} \tag{27}$$

In $H'_{(x,y)}$, the relative coordinates of all trajectories in $X'PY'$ are saved.

(3) For each point P, the final accumulated energy $E(P(x,y))$ is:

$$E(P(x,y)) = E_1(P(x,y)) + E_2(P(x,y)) \qquad (28)$$

If $k = n \div 2$, the trajectory number m_2 in the proposed two-stage MHT model is

$$m_2 = 2 \times (2^{0.5 \times n} - 1) \qquad (29)$$

The ratio of m and m_2 is α:

$$\alpha = \frac{m}{m_2} = \frac{2^n - 1}{2 \times (2^{0.5 \times n} - 1)} \approx \frac{2^n}{2^{0.5 \times n + 1}} = 2^{0.5 \times n - 1} \qquad (30)$$

α shows that with increasing trajectory length, compared with the proposed single-stage MHT, the trajectory hypothesis space of the two-stage MHT decreases exponentially. For example, when $n = 30$, the trajectory number in the single-stage MHT is approximately 1.07×10^9. Therefore, many trajectories make the trajectory space very difficult to store and calculate, and there is basically no operability. However, the trajectory number in the two-stage MHT is approximately 6.5×10^4, making 30-frame-based energy accumulation possible.

2.3. Target Tracking Method

For each pixel of the current frame, after energy accumulation via two-stage MHT, the energy accumulation map can be obtained, and the candidate target points can be obtained by constant false alarm (CFAR) segmentation of the map.

To further eliminate false alarms, a target tracking method is designed. It consists two steps.

(1) All the reserved candidate points in the target search area are tracked.

(2) The discontinuous trajectories and the associated points are deleted.

2.4. Detection Framework

The DP–MHT–TBD detection framework is shown in Figure 7. It is a sequential detection process.

Part 1: Second power optimal merit function-based DP

The continuous n_1 frames are fed into the proposed second power optimal merit function-based DP to obtain the target search area XPY and its opposite area $X'PY'$. These two areas are used in the next part. The flow of this part consists of four steps (the details can be seen from Formula (13) to (16) in Section 2.1.3).

Part 2: Two-stage MHT and CFAR segmentation

(1) The previous n_2 frames are used for energy accumulation. For each pixel P in the current frame, the proposed two-stage MHT model is used to accumulate energy along the trajectories only in areas XPY and $X'PY'$. Once all the points have been processed, the energy accumulation map can be obtained. The flow of this part consists of three steps (the details can be seen from Formula (26) to (28) in Section 2.2.2).

(2) The CFAR segmentation method is used to segment the energy accumulation map to obtain the candidate points of the current frame. In the segmentation process, the threshold Th is set according the pre-set false alarm rate F_a.

$$Th = -lnF_a \qquad (31)$$

Part 3: Target tracking

All the reserved candidate points are tracked, the discontinuous trajectories are found, and the candidate points in the discontinuous trajectories to noise are classified. The flow of this part consists of two steps (the details can be seen in Section 2.3).
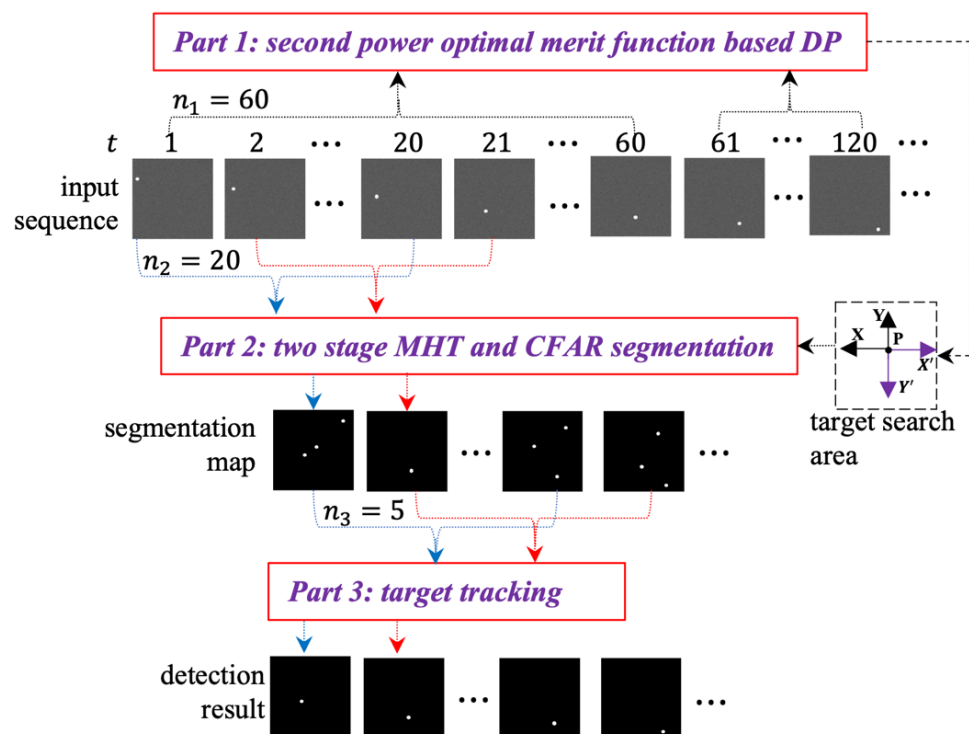
Figure 7. Proposed DP–MHT–TBD detection framework. It consists of the second power optimal merit function-based DP, two stage MHT and CFAR segmentation, target tracking. It is a sequential detection process.

3. Experiments and Analysis

To verify the proposed algorithm, experiments regarding the proposed second power optimal merit function-based DP, two-stage MHT and target tracking were carried out.

3.1. Datasets and Evaluation Setup

In practical applications, it is difficult to obtain the IR data of space point targets, so an IR image with low SCR is simulated first. Some images are shown in Figure 1b. When building the simulation image, the noise function of MATLAB is first used to add Gaussian noise ($N \sim (\mu, \sigma)$) to a blank 256×256 image to obtain the background B. The mean value μ and the standard deviation σ of the background are 90 and 10, respectively. Then, according to an SCR value, as in (32), the target T is only set at one certain position in each image. In the datasets, the target occupies only one pixel in each frame and the target motion speed is 1 pixel per frame.

$$T = B + SCR \times \sigma \tag{32}$$

The targets in the infrared sequence simulated according to the above process obey normal distribution, making the simulated datasets more reasonable and more correspondent with the actual infrared point target scene.

The target search area detection probability P_{area} is used to verify whether the proposed second power DP can correctly find the target search area.

$$P_{area} = \frac{number\ of\ correct\ detections}{number\ of\ experiments} \tag{33}$$

The detection probability P_d and the false alarm rate F_a are used as the evaluation metrics for target detection, and P_d and F_a are the ordinate and the abscissa of the receiver operating characteristic (ROC) curve, respectively [4].

$$P_d = \frac{number\ of\ true\ detections}{number\ of\ actual\ targets} \tag{34}$$

$$F_a = \frac{number\ of\ false\ detections}{number\ of\ images} \tag{35}$$

The detected target is considered true if it simultaneously meets two requirements: (1) the center of the ground truth is detected and (2) the pixel distance between the center of the ground truth and the result is less than 2 pixels (Manhattan distance).

All the experiments were implemented with MATLAB R2019 and C++ in Ubuntu 16.04 on a PC with a 4-core CPU and 16-GB RAM.

3.2. Experiments on Second Power Optimal Merit Function-Based DP

A total of 1000 sets of sequences with different lengths and different SCRs are used to verify the proposed second power optimal merit function-based DP.

3.2.1. The Second Power Optimal Merit Function-Based DP

To determine the advantage and validity of the proposed method, a comparative experiment with respect to the first/second power optimal merit function-based DP is carried out. The SCR of the sequences used in the experiment is 1.5. The comparison results are shown in Table 1, where n denotes the number of frames used in different methods.

Table 1. Quantitative comparison of DP.

Methods	n	P_{area} (%)	P_d (%)	Time
first power DP	30	68.76	44.76	0.4 s
second power DP	30	80.76	61.62	0.8 s
first power DP	45	86.10	56.38	0.6 s
second power DP	45	95.33	73.05	1.2 s
first power DP	60	92.29	62.76	0.8 s
second power DP	60	98.57	78.76	1.6 s
first power DP	75	92.84	57.00	1 s
second power DP	75	98.69	74.34	2 s

From the comparison result (see Table 1), the following conclusions can be drawn:

(1) In a very low SCR condition (SCR = 1.5), no matter how many frames are used, and whether the first or second power function-based DP is used, $P_{area} > P_d$. For example, when using 60 frames of images, the target search area detection probability is 98.57%, while the target detection probability is only 78.76%. As in Table 1, the P_d value of all methods is less than 80%. These data show that using the DP method to find the target search area is more reliable than directly detecting the target position.

(2) P_{area} increases with the number of frames used in the energy accumulation. However, this parameter does not grow linearly but grows more slowly. For example, when n increases from 45 to 60, P_{area} increases approximately 3.3%. However, when n increases from 60 to 75, P_{area} only increases approximately 0.1%. P_d increases with the number of frames, but is not necessarily larger as the number of frames continues to increase. For example, compared with $n = 60$, when $n = 75$, P_d decreases by approximately 5%.

(3) The first power DP and the second power DP accumulate energy in four and eight areas, respectively, so the time cost of the second power DP is twice that of the first power DP. However, even if 60 frames are used, it only takes 1.6 s to find the target search area. In practical applications, the more frames that are used, the greater the amount of calculation

and the worse the real-time performance. After comprehensive consideration of various factors, the appropriate n is 60.

3.2.2. The SCR of the Infrared Sequence

In each different SCR (from 1 to 1.5), 1000 sets of sequences are used to investigate the performance of the proposed method. The length n of every sequence is 60.

The P_{area} and P_d performances are shown in Table 2 and Figure 8. The following conclusions can be drawn from the results:

Table 2. Performance of the proposed DP with different SCRs.

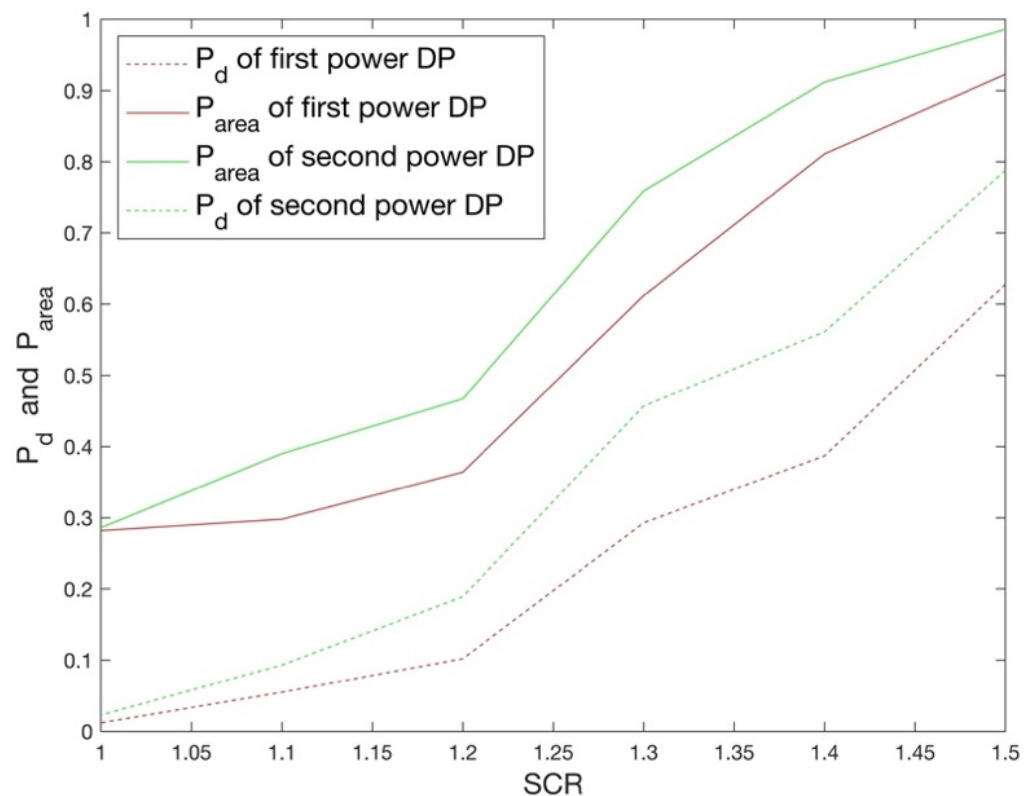| SCR: Methods	1.5	1.4	1.3	1.2	1.1	1
P_{area} of first power DP	92.3	81.1	61.2	36.4	29.8	28.2
P_d of first power DP	62.8	38.7	29.3	10.2	5.5	1.2
P_{area} of second power DP	98.6	91.2	75.9	46.7	39.0	28.6
P_d of second power DP	78.8	56.1	45.7	18.9	9.3	2.3

Figure 8. P_{area} and P_d of different method at different SCRs.

(1) In different SCRs, the performance of the proposed second power DP is better than that of the original first power DP. As shown in Figure 8, the P_{area} and P_d curves of the second power DP are always above the curve of the first power DP, showing the advantages of the second power DP. In addition, the P_{area} curves of different methods are above the PD curves, further showing that using the DP method to find the target search area is more reliable than directly detecting the target position.

(2) As the SCR decreases, P_{area} decreases rapidly. For the proposed second power DP, when $SCR \leq 1.3$, P_{area} is less than 80%, and when $SCR \leq 1.2$, P_{area} is less than 50%, indicating that the effect of the proposed method is not good when the $SCR < 1.4$. The subsequent target detection part is based on the correct detection of the target search area.

If P_{area} is less than 90%, even if the P_d of the subsequent MHT part is very high, then the final detection probability will not be very high.

3.3. Experiments on Two-Stage MHT

The condition of the proposed MHT is that the target search area is correctly determined by the proposed DP. Event A is defined as the target search area correctly found by the proposed DP, while event B is defined as the target correctly detected by the proposed MHT. The target detection probability of the DP–MHT–TBD is $P(dp - mht - tbd)$. In this work, $P_{area} = P(A)$ and $P(dp - mht - tbd) = P(AB)$ According to the conditional probability formula, $P(AB)$ is:

$$P(AB) = P(A) \times P(B|A) \tag{36}$$

where $P(B|A)$ denotes the conditional probability. Therefore, to obtain $P(B|A)$, experiments were carried out on the basis of event A being true.

According to above analysis, when $SCR \geq 1.4$, the proposed DP can find the target search area with probability greater than 90. The MHT-based target detection part is based on the target search area. Therefore, this section only investigates the performance of the proposed two-stage MHT when SCR $\geq$ 1.4. A total of 1000 sets of sequences with different lengths n are used to verify the performance of the proposed MHT, and the target SCR of each sequence is 1.5 or 1.4. The ROC curves are shown in Figure 9, and the P_d and time cost are shown in Tables 3 and 4, respectively. In this section, $P_d = P(B|A)$.
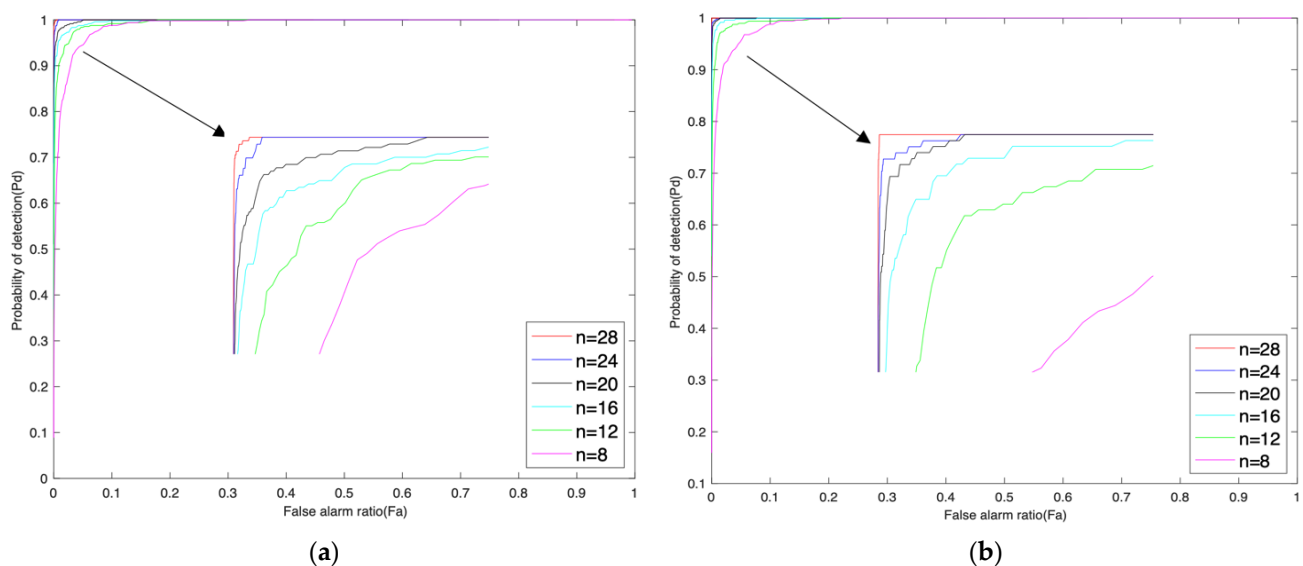

(**a**)   (**b**)

Figure 9. ROC curves of two stage MHT. n denotes the number of used images. (**a**) ROC curves when SCR is 1.4; (**b**) ROC curves when SCR is 1.5.

Table 3. $P(B|A)$ when SCR is 1.4.

F_a	*n* 8	12	16	20	24	28
0.1	97.69	98.18	98.79	99.13	99.69	99.89
0.01	77.23	89.57	94.63	96.88	97.62	98.79
0.001	41.33	67.01	83.79	90.10	95.77	98.12
0.0001	19.35	35.66	58.79	76.54	86.07	94.83
time	0.01 s	0.03 s	0.15 s	0.74 s	3.5 s	16 s

Table 4. $P(B|A)$ when SCR is 1.5.

F_a \ n	8	12	16	20	24	28
0.1	97.79	98.39	99.72	99.72	99.98	100
0.01	81.46	94.29	97.54	98.58	99.10	99.12
0.001	52.42	73.80	89.91	94.76	97.64	98.97
0.0001	28.43	49.80	72.08	86.65	93.40	97.92
time	0.01 s	0.03 s	0.15 s	0.74 s	3.5 s	16 s

From the results, the following conclusions can be drawn:

(1) The more images used, the better the performance but the greater the amount of calculation. As shown in Figure 9, the larger n is, the closer the ROC curve is to the top left corner, indicating that the more images used, the greater the energy accumulated in the energy accumulation process, the higher the target detection probability and the lower the false alarm rate. With increasing n, the amount of calculation also increases. For example, when $n = 30$, the time cost is 33 s, which means that it takes approximately 9 h to perform 1000 experiments. The amount of calculation is too large, so n is less than 28 in the experiment.

(2) The performance decreases with decreasing SCR. As in Tables 3 and 4, to meet the dual requirements of $F_a < 0.1\%$ and $P_d > 90\%$, when SCR is 1.5 and 1.4, the minimum number of images required is approximately 16 and 20, respectively.

After obtaining $P(A)$ and $P(B|A)$, according (36), the final target detection probability of the DP–MHT–TBD $P(dp - mht - tbd)$ can be obtained. When $F_a = 0.1\%$, $P(dp - mht - tbd)$ is shown in Table 5. When SCR is 1.5, the proposed DP–MHT–TBD can detect the target in the image sequence with a detection probability of more than 90%.

Table 5. $P(dp - mht - tbd)$ with $F_a = 0.001$.

SCR \ n	8	12	16	20	24	28
1.4	37.69	61.11	76.42	82.17	87.34	89.49
1.5	51.69	72.77	88.65	93.43	96.27	97.58
time	0.01 s	0.03 s	0.15 s	0.74 s	3.5 s	16 s

3.4. Experiments on Target Tracking

After CFAR segmentation of the energy accumulation map, candidate points can be obtained. However, the candidate points contain both targets and noise.

The energy accumulation maps are obtained by the two-stage MHT in which 20 frames are used for energy accumulation. Then, the CFAR segmentation method is used to segment 450 continuous energy accumulation maps. In the target tracking part, the trajectories that are discontinuous in five continuous segmentation maps are deleted, and the associated points are classified as noise. For each operation, the time cost is 0.003 s. The change in F_a before and after target tracking is shown in Table 6. It can be seen that the target tracking method can eliminate false alarms by an order of magnitude.

Table 6. F_a after target tracking.

SCR \ F_a'	0.1	0.01	0.001	0.0001
1.5	0.033	0.0010	0.00007	0.0000062
1.4	0.046	0.0015	0.0001	0.0000078

3.5. Comparison Experiments

To verify the superiority of the proposed DP–MHT–TBD, three kinds of methods should be compared: TBD, DBT and DL Methods. However, this work focuses on the

detection of the 1×1 point target under low SCR condition (SCR = 1.5). There is no shape and texture information in the 1×1 point target. It is difficult to train the neural networks with a sample in which there is only grey information of one pixel. Many deep learning methods were tested in the experiment, such as Faster-RCNN, the previous work's spatial–temporal based method [1], TSDF [2] and so on. In these methods, the neural network could not learn the useful features and the network could not achieve coverage during the training process. To the best of our knowledge, there is no deep learning method that could detect the infrared point target under low SCR (SCR = 1.5). Thus, there are no experiments about deep learning methods in this work. In addition, the point target detection task under low SCR is usually related to the military field; it is difficult to obtain relevant code due to confidentiality rules or technical reasons. Considering the above factors, the following representative DBT and TBD methods are selected: DBT: LIG [6], multiscale patch-based contrast measure (MPCM) [28], absolute average gray difference (AAGD) [29], absolute average difference weighted by cumulative directional derivatives (AADCDD) [30] and LCM utilizing a tri-layer (TLCM) [31]; TBD: DP [12], second power optimal merit function-based DP, and facet derivative-based multidirectional edge awareness with spatial–temporal tensor (FDMDEA-STT) [32]. The experimental parameters and time cost of all methods are listed in Table 7. The ROC curves are shown in Figure 10. Some CFAR segmentation maps are shown in Figure 11.

Table 7. Parameter setting and time cost of different methods.

| Methods	Parameter Setting	Time Cost
AADCDD [30]	local window size: $N = 3, 5, 7$	0.01 s
AAGD [29]	internal and external window size $N = 3, 5, 7$	0.005 s
LIG [6]	window size: $N = 3$; ratio $\kappa = 0.2$	0.6 s
MPCM [28]	cell size: $N = 3, 5, 7$	0.04 s
TLCM [31]	window size: $N = 3, 5, 7$	0.56 s
FDMDEA-STT [32]	$\lambda = 0.1$, $t = 3$, $\rho = 1.05$, $\epsilon = 10^{-7}$	1.5 s
DP [12]	20 frames are used to accumulate energy	0.07 s
second power DP	20 frames are used to accumulate energy	0.07 s
Proposed	60 frames are used to find the target search area 20 frames are used to accumulate energy 5 segmentation maps are used to eliminate false alarms	1.6 s 0.74 s 0.003 s
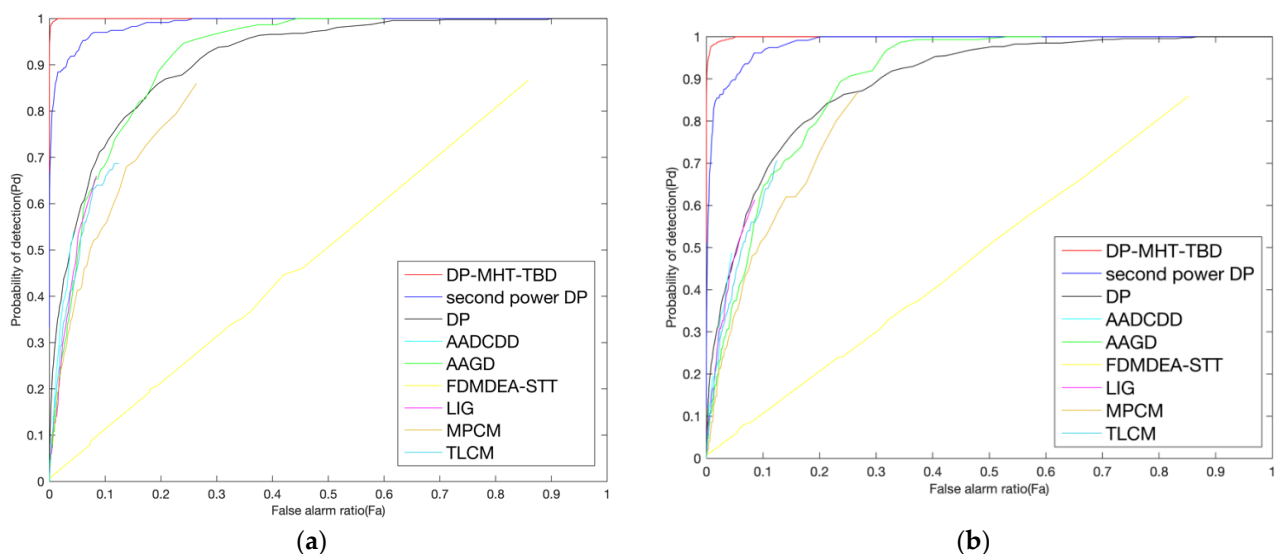
Figure 10. ROC curves of different methods under different data. (**a**) SCR is 1.5; (**b**) SCR is 1.4.
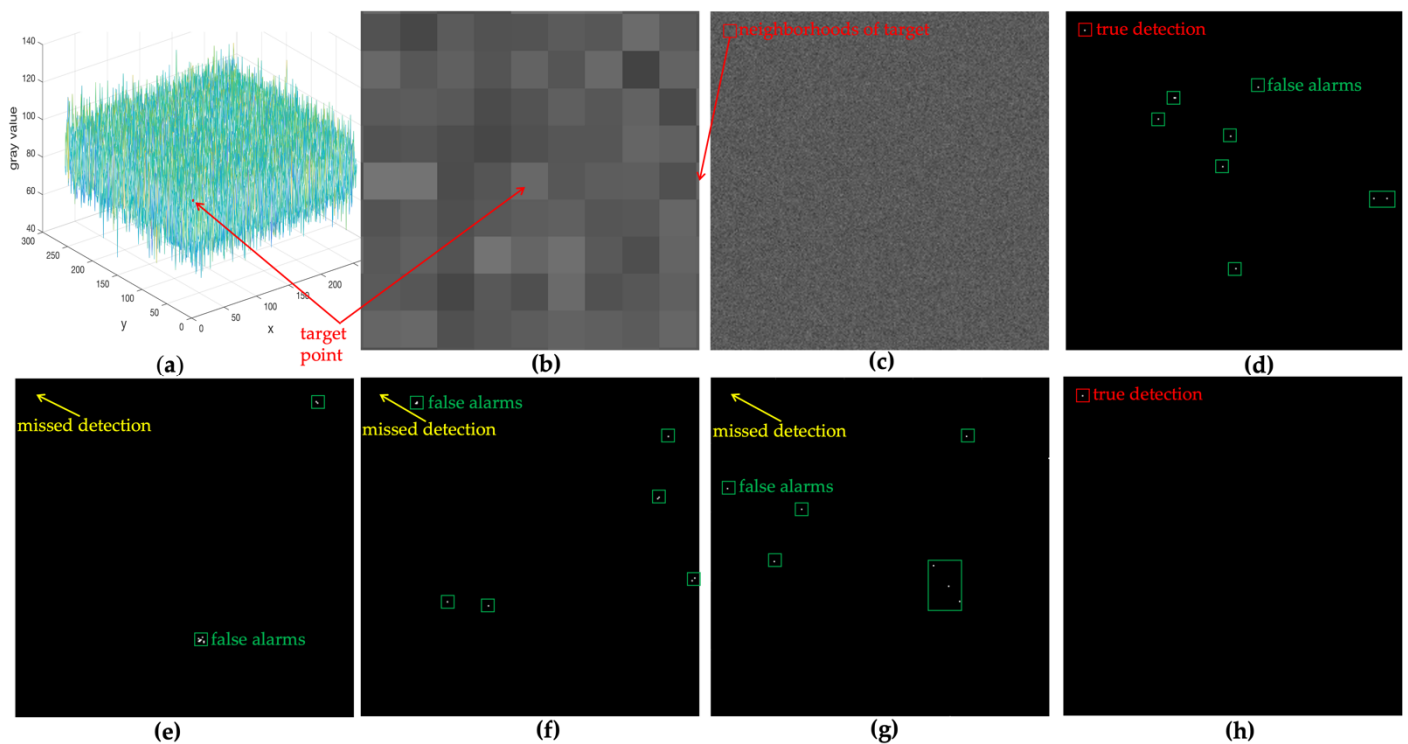
Figure 11. Detection results of the different methods. (**a**) 3-D display of the input image; (**b**) 9 × 9 neighborhoods of the target; (**c**) input image, SCR is 1.5; (**d**–**g**) are CFAR segmentation maps of different methods: (**d**) two stage MHT; (**e**) second power DP; (**f**) DP; (**g**) AAGD; (**h**) final result of DP–MHT–TBD. In the CFAR segmentation process, F_a = 0.01%.

From the results, it can be seen that:

(1) The proposed DP–MHT–TBD is superior to other methods with respect to the detection probability and false alarm rate. The proposed second power DP has better performance than the original DP.

(2) The proper energy accumulation method is the key to the detection of dim point targets. Only the proposed DP–MHT–TBD can detect the target because the energy of the point target is accumulated through the right way. In other methods, the target energy is not accumulated (in single-frame detection methods and FDMDEA-STT), or there is a serious diffusion of energy during the accumulation process (in DP and second power DP); the target is missed after segmentation because the energy of the target point is not superior to that of the noise point.

(3) The consumption time of the DP–MHT–TBD is greater than that of the other methods because there are more images to be processed. The proposed method is a sequential detection process and the image processing speed can be accelerated by using GPU or other ways, having high practical application values.

In brief, in Section 3, experiments on the proposed second power optimal merit function-based DP, two-stage MHT, target tracking and comparison methods are carried out. The results of each part verify the superiority of the corresponding part. When SCR is 1.5, 60 images (image size is 256 × 256) are used in the second power optimal merit function-based DP to find the target search area, 20 images are used in the two-stage MHT to accumulate energy, and 5 segmentation maps are used in target tracking to further eliminate false alarms. According to Tables 2, 5 and 6, the proposed DP–MHT–TBD can detect a single point target in the image sequence with a detection probability above 90% and a false alarm rate of below 0.01%.

4. Discussion

Several observations from experimental and quantitative analysis are discussed.

First, theoretical and experimental studies have shown that the proposed second power optimal merit function-based DP outperforms the original DP. The proposed second power DP is used to find the target search area instead of directly detecting the target, avoiding the influence of agglomeration effect and reducing the trajectory hypothesis space.

Second, the proposed two stage parallel MHT is the key to detecting a point target under low SCR. This can be attributed to the fact that the novel parallel MHT structure and the two-stage strategy simplifies the trajectory search problem and reduces the hypothesis space exponentially. This enables the energy accumulation process of dim targets to be very efficient.

Third, as the SCR decreases, the point detection task becomes more and more difficult. The gain is finite by only using infrared data or improving TBD methods. Some potential ideas should be considered. (1) The idea of adaptive spatial–temporal context [33] can be used to improve the robustness of the tracking part in the TBD. (2)The multi-sensor data fusion [34,35] may be a valuable project for theoretical investigation and practical application in the dim point detection field. With the development of detector technology, it will be easy to obtain radar, infrared, hyperspectral and other data. If these data can be fused to take advantage of their respective advantages, it will be of great help in reducing the trajectory hypothesis space, improving processing efficiency and detection capability.

5. Conclusions

In this paper, a novel and accurate DP–MHT–TBD (dynamic programming–multiple hypothesis testing–track before detect) algorithm is proposed for infrared dim point target detection. The method consists of three parts: the second power optimal merit function-based DP, two-stage MHT and target tracking. In particular, first, for each point to be detected, the second power optimal merit function-based DP is used to find the target search area to reduce the trajectory hypothesis space. Next, the two-stage MHT, which can further reduce the trajectory hypothesis space exponentially and mitigate the contradiction between the redundant trajectories and the requirements of more trajectories under low SCR (signal-to-clutter ratio), is used to save the tree-structured trajectory space and accumulate energy in parallel. Finally, after CFAR (constant false alarm) segmentation of the accumulation map, the target tracking method is used to further eliminate false alarms. Via experiments on each part and comparison methods, the proposed DP–MHT–TBD algorithm, which takes advantage of the small computation cost of DP and high accuracy of exhaustive search, greatly reduces the computational cost and storage requirements, showing superior ability in detecting point targets under low SCR conditions. The point target detection task under low SCR is difficult but significant. Although the proposed method has engineering applicability and good detection performance, it can only detect a single point target. In future work, the studies could be conducted from the following aspects: the combination of TBD and deep learning methods, the use of adaptive spatial–temporal context information and the fusion of multi-sensor data.

References

1. Du, J.M.; Lu, H.Z.; Zhang, L.P.; Hu, M.F.; Chen, S.; Deng, Y.J.; Shen, X.; Zhang, Y. A Spatial-Temporal Feature-Based Detection Framework for Infrared Dim Small Target. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–12. [CrossRef]
2. Du, J.M.; Lu, H.Z.; Hu, M.F.; Zhang, L.P.; Shen, X.L. CNN-based infrared dim small target detection algorithm using target-oriented shallow-deep features and effective small anchor. *IET Image Process.* **2021**, *15*, 1–15. [CrossRef]
3. Ren, X.; Wang, J.; Ma, T.; Zhu, X.; Bai, K.; Wang, J. Review on Infrared Dim and Small Target Detection Technology. *J. Zhengzhou Univ.* **2020**, *52*, 1–21. [CrossRef]
4. Zhang, P.; Zhang, L.; Wang, X.; Shen, F.; Pu, T.; Fei, C. Edge and Corner Awareness-Based Spatial–Temporal Tensor Model for Infrared Small-Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 10708–10724. [CrossRef]
5. Chen, C.L.P.; Li, H.; Wei, Y.; Xia, T.; Tang, Y.Y. A Local Contrast Method for Small Infrared Target Detection. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 574–581. [CrossRef]
6. Zhang, H.; Zhang, L.; Yuan, D.; Chen, H. Infrared small target detection based on local intensity and gradient properties. *Infrared Phys. Technol.* **2018**, *89*, 88–96. [CrossRef]
7. Gao, C.; Meng, D.; Yang, Y.; Wang, Y.; Zhou, X.; Hauptmann, A.G. Infrared Patch-Image Model for Small Target Detection in a Single Image. *IEEE Trans. Image Process.* **2013**, *22*, 4996–5009. [CrossRef]
8. Reed, I.S.; Gagliardi, R.M.; Shao, H.M. Application of Three-Dimensional Filtering to Moving Target Detection. *IEEE Trans. Aerosp. Electron. Syst.* **1983**, *19*, 898–905. [CrossRef]
9. Carson, B.D.; Evens, E.D.; Wilson, S.L. Search radar detection and track with the Hough transform part I: System concept. *IEEE Trans. Aerosp. Electron. Syst.* **1994**, *30*, 102–108. [CrossRef]
10. Carson, B.D.; Evens, E.D.; Wilson, S.L. Search radar detection and track with the Hough transform Part II: Detection statistics. *IEEE Trans. Aerosp. Electron. Syst.* **1994**, *30*, 109–115. [CrossRef]
11. Carson, B.D.; Evens, E.D.; Wilson, S.L. Search radar detection and track with the Hough transform Part III: Detection performance with binary integration. *IEEE Trans. Aerosp. Electron. Syst.* **1994**, *30*, 116–125. [CrossRef]
12. Barniv, Y. Dynamic Programming Solution for Detecting Dim Moving Targets. *IEEE Trans. Aerosp. Electron. Syst.* **1985**, *21*, 144–156. [CrossRef]
13. Blostein, S.D.; Huang, T.S. Detecting small, moving objects in image sequences using sequential hypothesis testing. *IEEE Trans. Signal Process.* **1991**, *39*, 1611–1629. [CrossRef]
14. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, C. SSD: Single shot multibox detector. In Proceedings of the 14th European Conference on Computer Vision (ECCV 2016), Amsterdam, The Netherlands, 8–16 October 2016. [CrossRef]
15. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, 26–30 June 2016. [CrossRef]
16. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [CrossRef] [PubMed]
17. Zhao, M.; Li, C.; Xu, Y.; Peng, F.; Liu, L.; Wu, N. TBC-Net: A real-time detector for infrared small target detection using semantic constraint. *arXiv* **2019**, arXiv:2001.05852. [CrossRef]
18. Lin, H.U.; Wang, S.Y.; Wan, Y. Improvement on Track-Before-Detection Algorithm Based on Dynamic Programming. *J. Air Force Radar Acad.* **2010**, *24*, 79–82. [CrossRef]
19. Gao, F.; Zhang, F.; Zhu, H.; Sun, j.; Wang, J. An improved TBD algorithm based on dynamic programming for dim SAR target detection. In Proceedings of the 12th International Conference on Signal Processing (ICSP 2014), Bangkok, Thailand, 19–21 November 2014. [CrossRef]
20. Lyu, T.J.; Yuan, Z.Q.; Lei, G.; Ren, Z. A DP-TBD Algorithm under Complex Clutter Environment. *Fire Control. Radar Technol.* **2021**, *50*, 78–81. [CrossRef]
21. Xing, H.; Suo, J.D.; Sun, B. Dynamic programming track-before-detect algorithm with improved state transition set. *Mod. Electron. Tech.* **2020**, *43*, 1–5. [CrossRef]
22. Guo, Q.; Li, Z.; Song, W.; Fu, W. Parallel Computing Based Dynamic Programming Algorithm of Track-before-Detect. *Symmetry* **2018**, *11*, 29. [CrossRef]
23. Wang, Y. The Study of Weak Target Track before Detect Algorithm Basing on Dynamic Programming. Master's Thesis, Yanshan University, Qinhuangdao, China, 2017.
24. Dong, J. Tracking before Detection Based on Dynamic Programming (DP-TBD) Algorithm. Master's Thesis, Dalian Maritime University, Dalian, China, 2017.
25. Chen, H.; Sun, G.F.; Lu, H.Z.; Chen, S.F. Detection Algorithm for Small Moving Targets Based on Dynamic Programming and Confidence Test. *Syst. Eng. Electron.* **2003**, *25*, 472–475. [CrossRef]
26. Chen, S.F.; Xiao, S.Z.; Lu, H.Z. Dim targets detection based on multi-regions dynamic programming and track matching. *Infrared Laser Eng.* **2007**, *36*, 738–741. [CrossRef]
27. Chen, S.F.; Xiao, S.Z.; Lu, H.Z. An algorithm of low SNR small targets real-time detection in imagery. *Signal Process.* **2009**, *25*, 601–606. [CrossRef]
28. Wei, Y.T.; You, X.; Li, H. Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recognit.* **2016**, *58*, 216–226. [CrossRef]

29. Deng, H.; Sun, X.; Liu, M.; Ye, C.; Zhou, X. Infrared small-target detection using multiscale gray difference weighted image entropy. *IEEE Trans. Aerosp. Electron. Syst.* **2016**, *52*, 60–72. [CrossRef]
30. Aghaziyarati, S.; Moradi, S.; Talebi, H. Small infrared target detection using absolute average difference weighted by cumulative directional derivatives. *Infrared Phys. Technol.* **2019**, *101*, 78–87. [CrossRef]
31. Han, J.; Moradi, S.; Faramarzi, I.; Liu, C.; Zhang, H.; Zhao, Q. A Local Contrast Method for Infrared Small-Target Detection Utilizing a Tri-Layer Window. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1822–1826. [CrossRef]
32. Pang, D.; Shan, T.; Li, W.; Ma, P.; Tao, R.; Ma, Y. Facet Derivative-Based Multidirectional Edge Awareness and Spatial–Temporal Tensor Model for Infrared Small Target Detection. *IEEE Geosci. Remote Sens.* **2022**, *60*, 1–15. [CrossRef]
33. Mehmood, K.; Jalil, A.; Ali, A.; Khan, B.; Murad, M.; Khan, W.U.; He, Y.G. Context-Aware and Occlusion Handling Mechanism for Online Visual Object Tracking. *Electronics* **2021**, *10*, 43. [CrossRef]
34. Stateczny, A.; Kazimierski, W. Multisensor Tracking of Marine Targets-Decentralized Fusion of Kalman and Neural Filters ehmood. *Int. J. Electron. Telecommun.* **2011**, *57*, 65–70. [CrossRef]
35. Liu, W.; Liu, Y.; Bucknall, R. Filtering based multi-sensor data fusion algorithm for a reliable unmanned surface vehicle navigation. *J. Mar. Eng. Technol.* **2022**, 1–17, *Early Access*. [CrossRef]