

An Infrared Dim Target Detection Algorithm based on Density Peak Search and Region Consistency

Baohua Zhang

Inner Mongolia University of Science and Technology

Jinhui Zhu (✉ zjh_9277@163.com)

Inner Mongolia University of Science and Technology <https://orcid.org/0000-0002-9494-9182>

XIAOQI Lu

Inner Mongolia Industrial University: Inner Mongolia University of Technology

Yu Gu

Inner Mongolia University of Science and Technology

Jianjun Li

Inner Mongolia University of Science and Technology

Xin Liu

Inner Mongolia University of Science and Technology

Ming Zhang

Inner Mongolia University of Science and Technology

Research Article

Keywords: Small target detection, density peak search, local mosaic model, infrared images.

Posted Date: April 5th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-313437/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Optical and Quantum Electronics on July 14th, 2021. See the published version at <https://doi.org/10.1007/s11082-021-03056-x>.

An infrared dim target detection algorithm based on density peak search and region consistency

BAOHUA ZHANG,^{1,3} JINHUI ZHU,^{1*} XIAOQI LU,^{2,3} YU GU,^{1,3} JIANJUN LI,^{1,3} XIN LIU,^{1,3} AND MING ZHANG^{1,3}

¹*School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou, Inner Mongolia, 014010, China*

²*School of Information Engineering, Mongolia Industrial University, Huhehaote, Inner Mongolia, 010051, China*

³*Inner Mongolia Key Laboratory of Pattern Recognition and Intelligent Image Processing, Baotou, Inner Mongolia 014010, China*

* zjh_9277@163.com

Abstract: To suppress background clutter and improve detection accuracy, this paper propose a dim target detection algorithm based on density peak search and region consistency. Firstly the density peak search algorithm is used to extract the candidate targets. And then the candidate targets are classified and marked according to the local mosaic probability factor, which is important to suppress the background clutter and accurately strip the candidate target region from the background. Considering the regional stability of the dim targets, local mosaic gradient factors are used to screen real targets from the candidate targets, and then facet kernel filter is used to extract the irregular contours of the dim targets, and as a result, the targets can be enhanced. The experimental results show that compared with the existing algorithms, the proposed method has better detection accuracy and stronger robustness in different complex scenarios.

Keywords: Small target detection; density peak search; local mosaic model; infrared images.

1. Introduction

An infrared search and tracking system (IRST) has the advantages of good concealment and robust anti-interference ability, which is widely used in many aspects such as military early warning, precision guidance and remote sensing (Zeng et al. 2006; Wang et al. 2019). The infrared target detection method is one of key technologies (Chen et al. 2014). However, the dim target detection faces many challenges, such as the target only contains dozens of pixels in each frame, the dim targets miss texture and shape information, and furthermore the

signal to clutter ratio of each frame is very low (Ye et al. 2017; Huang et al. 2019) . So the dim targets are often submerged in complex backgrounds (such as cloud edges, ocean waves, and high-brightness noise).

The infrared dim target detection methods can be divided into single-frame-based detection methods (Dai et al. 2017) and multi-frame-based detection methods (Lv et al. 2018; Wan et al. 2016; Zhang et al. 2019). The single-frame-based methods have many advantages such as low complexity, high execution efficiency and strong real-time performance. Therefore, the single-frame-based methods have attracted more attention.

Most of the traditional single-frame-based methods, such as morphological filtering based method (Zeng et al. 2006) and probability statistical filtering based methods (Deshpande et al. 1999; Soni et al. 1993; Tomasi et al. 1998; Barnett et al. 1989), detected the targets by suppressing the background. Although the above methods are easy to implement in real time, the performance is greatly reduced by the clutter interference in a complex background. Considering the characteristics of the human visual system (HVS), Chen proposed a local contrast measure (LCM), which detected the target based on the brightness difference between the target and the adjacent areas (Chen et al. 2014). However, this method was not suitable for the scenes with high background brightness, and a large number of the improved methods have subsequently appeared, such as the multiscale patch-based contrast measure-based methods (Wei et al. 2016; Nie et al. 2018). Because the above methods were insufficient for extracting the features of the target, and had high false alarm rates caused by high-intensity clutter and high-brightness noise.

From the perspective of separating the targets, Huang proposed a dim target detection method based on maximum gray region growing, which segmented the target and background by selecting the seed growing points (Huang et al. 2019). However, this method was only suitable for scenarios where the number of the targets was much smaller than the number of the candidate targets. Xia combined the random walk algorithm with local contrast features to achieve efficient target detection, because the selected seed points in the algorithm should be labeled and classified, which was inefficient (Xia et al. 2018). Qin introduced the facet kernel filtering to suppress the clutter and enhance the targets (Qin et al. 2019), as a result, the selection of the candidate points was simple, but due to the dependence on a single target characteristic, the false alarm rate of the method was high in complex backgrounds such as strong clutter and bright edge clutter.

In order to robustly detect the dim targets in complex backgrounds, this paper proposed an infrared dim target detection algorithm based on density peak search and region consistency. The main contributions of the algorithm in this paper are as follows:

The dim target detection is regarded as an abnormal point detection problem in this paper, and the candidate target points are initially screened out using an unsupervised method; and then the local mosaic template regions are constructed with the candidate target point as the

center, and the candidate targets are classified and labeled locally to suppress the clutter. According to the gradient difference and direction consistency of dim infrared targets, the local mosaic probability factor (LMPF) is used to separate the target from the noise and the clutter, and the targets are enhanced with local mosaic gradient factor (LMGF). The experimental results show that compared with existing algorithms, the proposed detection algorithm has better performance in complex scenes.

2. Proposed Method

The details of the proposed method are as follows. Firstly, the candidate targets are selected according to the abnormal attributes based on the density peak search algorithm. And then the candidate targets are selected as the center, and the improved RW algorithm is used to construct the local mosaic template region. As a result, the feature search range can be reduced. To suppress the clutter, the LMPF is used to classify and mark the candidate targets, which is a critical step to separate the targets from the clutter. On this basis, according to the direction consistency of the dim targets, the LMGF is used to enhance the edge of the targets, and the targets can be extracted from the background. Fig.1 is the flow chart of the proposed method.

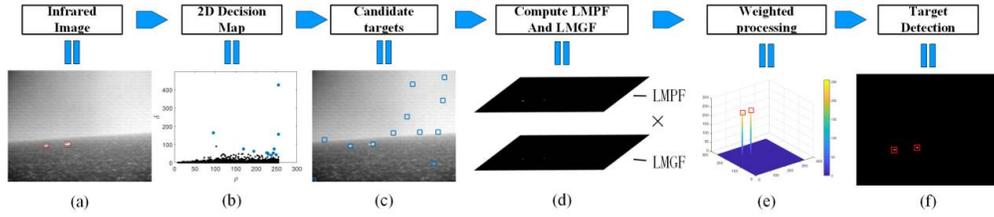


Fig. 1. The entire flow chart of the proposed method.

2.1 Density peak search algorithm

Density peak search algorithm (DPeak) is a granular computing model (Wang et al. 2017), whose theoretical model is based on two assumptions: 1) the center point is surrounded by the neighboring data points with lower local density; 2) there are no points with higher density around any center point.

For any pixel i , the DPeak model needs to calculate the local density ρ_i and the distance to the nearest correlation distance δ_i

$$\rho_i = f_i, \quad (1)$$

$$\delta_i = \min_{i:\rho_j > \rho_i} d_{ij}, \quad (2)$$

where f_i represents the gray value of the pixel, the highest density point $\delta_i = \max_j(d_{ij})$,

$$d_{ij} = |x_i - x_j| + |y_i - y_j|, \quad (3)$$

$$\Upsilon = \rho \times \delta, \quad (4)$$

where d_{ij} is the Manhattan distance, and then calculate the joint feature factor Υ pixel by pixel, and put them in the queue Q in descending order, and extract the first m pixels as candidate target points.

The gray value of the target is normally higher than that of the local neighboring pixels. Thus, the two elements of the DPeak model just meet the characteristics of dim targets. And the DPeak model can reduce the search range of the target and reduce the calculation time.

2.2 Local mosaic probability factor (LMPF) and local mosaic gradient factor (LMGF)

The texture structure of the dim target is essentially different from the clutter. Fig.2 compares the gradient distribution of the two. Fig.2 (a) is an infrared image, Fig.2 (b) is an enlarged view of the target area in Fig.2 (a). The upper image shows the gradient distribution of background clutter, and the lower image shows the gradient distribution of the target and its neighborhood. Obviously, the gradient direction of the target neighborhood is consistent, and the amplitude is uniform; relatively, the gradient direction under the background clutter is messy, and the amplitude difference is large. Fig.2 (c) and Fig.2 (d) are the three-dimensional gradient distribution map of the background clutter area and the target area. In view of the fact that the dim target has the characteristics of higher gray value, high gradient value and high density than the adjacent background area. The essence of candidate point screening is to judge whether it is a target based on local attributes.

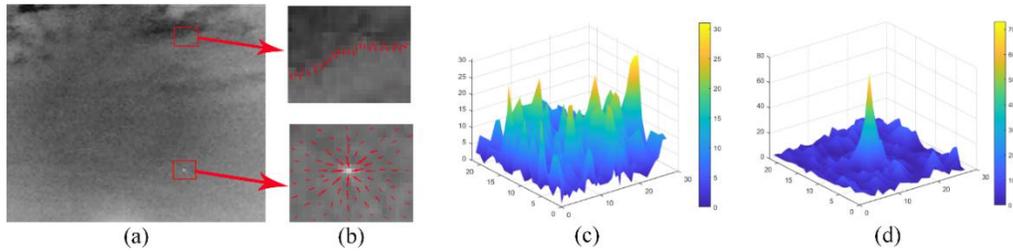


Fig. 2. Background clutter, target and its three-dimensional gradient map in infrared image.

The roaming window of the Random Walker algorithm is a rectangular frame, and the outermost pixels are usually judged as the background. In a multi-target scene, if the targets are adjacent or stuck together, the neighboring targets will be judged as the background (Qin et al. 2019), thus, part of the targets are missed. In order to accurately extract the target from the background, this paper improved the Random Walker (RW) algorithm and proposed an improved marking model. As shown in Fig.3 (a), the red frame region is defined as the central region, and the region between the yellow frame and the red frame is the transition region. The outer area is between the blue frame and the yellow frame. The candidate target is used as

the center pixel of the local mosaic model and is marked as 1-class pixel. And the pixels in the outer area of the mosaic model are marked as 0-class pixel.

Since most dim targets are less than 9×9 pixel, the template size is set to 11×11 pixel. The shape of the marked pixels in the outer region is close to the target contour, which can accurately approximate the true distribution of the target. The transition region and the outer region are defined as follows:

$$MR = (CR \oplus \text{shape}) - CR, \quad (5)$$

$$OR = (MR \oplus \text{shape}) - MR, \quad (6)$$

CR, MR and OR refer to the central region, the transition region and the outer region respectively. \oplus refers to the expansion operation, and shape refers to the diamond-shaped structural element.

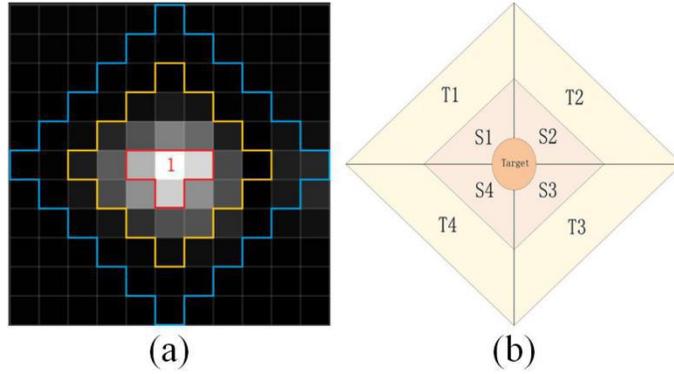


Fig. 3. Local mosaic model. (a) The marking strategy in the local mosaic model; (b) Area division diagram of local mosaic model.

In the local mosaic model, only the central pixel is marked as 1-class. And the pixels close to the outer region are classified into 0-class. The pixels in the transition region is classified according to its closeness to the gray level of the central pixel.

To build a new local mosaic region, the transition region and the outer region are updated based on Eq. (5) and Eq. (6) after obtaining the candidate target points. This paper selects the outer area as the background, and screens the targets by using the LMPF, which is an indicator that compares the probability value of the candidate target region with the probability value of the background.

$$LMPF = \begin{cases} \frac{MVP(\text{pixels})}{MVP(OR)}, & \text{if } \text{pixels} \not\subseteq Q \text{ and } \text{lable} = 1 \\ 0, & \text{otherwise} \end{cases}, \quad (7)$$

In Eq. (7), MVP (mean value of probability) is the average of pixel probabilities, and pixels are pixels expanded by candidate target pixels in the central area. LMPF is the gray value of the central area.

It is difficult to use the LMPF to robustly detect dim targets in highlight scenes, especially in low-contrast scenes. So the proposed method introduced the LMGF to expand the difference between the target and the background, and improved the accuracy of the method. As shown in Fig.3(b), the template is divided into eight parts based on a three-level nested model, namely S1, S2, S3, S4 (the middle area is referred to as S) and T1, T2, T3, T4 (outer area is referred to as T). Here, the LMGF is defined as follows:

$$g_{si} = \frac{1}{N_{si}} \sum_{k=1}^{N_{si}} \|grad_{si}\|^2, \quad (8)$$

$$G_{s\max} = \max_{1 \leq i \leq 4} g_{si}, \quad (9)$$

$$G_{s\min} = \min_{1 \leq i \leq 4} g_{si}, \quad (10)$$

in Eq. (8), $grad_{si}$ is the gradient element in the area of S_i ($i=1,2,3,4$), N_{si} is the number of gradients in S_i , g_{si} is the average magnitude of the gradient in S_i . $G_{s\max}$ and $G_{s\min}$ are the maximum and minimum values of the average gradient of S_i ($i=1,2,3,4$).

$$g_{ti} = \frac{1}{N_{ti}} \sum_{j=1}^{N_{ti}} \|grad_{ti}\|^2, \quad (11)$$

$$G_{t\max} = \max_{1 \leq i \leq 4} g_{ti}, \quad (12)$$

$$G_{t\min} = \min_{1 \leq i \leq 4} g_{ti}, \quad (13)$$

where, $grad_{ti}$ is the gradient element in the area of T_i ($i=1,2,3,4$), N_{ti} is the number of gradients in T_i , g_{ti} is the average magnitude of the gradient in T_i . $G_{t\max}$ and $G_{t\min}$ are the maximum and minimum values of the average gradient of T_i ($i=1,2,3,4$).

$$LMGF = \begin{cases} \sum_{i=1}^4 (g_{ti} + g_{si}), & \text{if } \frac{G_{s\max}}{G_{s\min}} < \frac{G_{t\max}}{G_{t\min}} \\ 0, & \text{otherwise} \end{cases}, \quad (14)$$

where, the LMGF is the gray value of the central area. Because the direction of the gradient in the neighborhood of the dim target is consistent, and the amplitude is close; relatively, the gradient direction and amplitude of the clutter are significantly different. So the target can be identified based on the above information.

2.3 Target extraction and enhancement

The input image is filtered using 5×5 kernel matrix K (Qi et al. 2016) to get the enhanced image, P_i is original image, as shown in Eq. (15):

$$K = \begin{pmatrix} -4 & -1 & 0 & -1 & -4 \\ -1 & 2 & 3 & 2 & -1 \\ 0 & 3 & 4 & 3 & 0 \\ -1 & 2 & 3 & 2 & -1 \\ -4 & -1 & 0 & -1 & -4 \end{pmatrix}, \quad P_e = P_i * K, \quad (15)$$

where, the target map P_m is obtained according to Eq. (16), P_m and P_e are calculated by the Hadamard product to get a weighted image P_w , P_w is called “Weighted”,

$$P_m = LMPF \cdot LMGF, \quad (16)$$

$$P_w = P_m \circ P_e, \quad (17)$$

in the enhanced image P_e , both the targets and the clutter are enhanced, while P_m only maps the target area in the proposed method. Therefore, the weighted image P_w is more approximate the target contour, indicating that the proposed algorithm can effectively suppress the clutter and enhance the targets.

2.4 Proposed Detection Method

As is illustrated in Algorithm 1, the proposed algorithm selected the candidate targets by constructing a feature space of local density-closest correlation distance, and then construct the local mosaic template area with the candidate target point as the center in turn. In order to suppress the clutter, the candidate target regions are classified and labeled according to the local mosaic probability factor, and finally the real targets are selected from the candidate targets by the local mosaic gradient factor.

Algorithm 1 The Proposed Detection Method

Input: infrared image P_i , parameters m and μ .

Output: Detection result.

- 1: **for** $i = 1$ to N **do** (the number of pixels is N)
 - 2: For i , obtain ρ_i according to (1).
 - 3: For i , calculate δ_i according to (2).
 - 4: Obtain 2D Decision Map and Compute Υ according to (4).
 - 5: **end for**
 - 6: Store and Sort all Υ in the queue Q in descending order
 - 7: Find the first m pixels as candidate targets and record their positions.
 - 8: **for** $j = 1$ to m **do**
 - 9: Select the candidate target pixel of $Q[j]$ and labeled as 1
 - 10: Construct the region of 11×11 pixels.
 - 11: Compute the LMPF according to (7).
-

12: Compute the LMGF according to (14).
13: Mark pixels segmented as label 1.
14:**end for**
15: Compute the enhanced map P_e according to (15).
16: Compute the weighted map P_m according to (16).
16: Compute the weighted map P_w according to (17).
17: Segment targets from weighted map using adaptive thresholding.

3. EXPERIMENTS AND ANALYSIS

In this section, the experiments are performed to verify the effectiveness of the proposed method. The experimental computing environment in this paper is 3.40GHz Intel i7-3770 CPU processor, 8GB memory, and the test software is MATLAB 2018b.

3.1 Experimental Setup

1) Datasets: To verify the robustness and accuracy of the proposed method, six sets of the infrared image sequences under different complex backgrounds are selected. Table 1 gives the detailed information of six groups infrared sequences. Each group of the image sequences have different backgrounds, the Frame1 is the edge of the sky and irregular clouds and the Frame2 is the glare background (Huang et al. 2019); The background of the Frame3 contains a lot of ring-shaped point clutter; While in the Frame4, the target is surrounded by clouds and there is a lot of cloud clutter (Qin et al. 2019). The Frame5 is a dual-target image sequence with high-intensity sky and sea wave clutter as the background, and the Frame6 is a four-target image sequence with low intensity (Xia et al. 2018).

Table 1. Details of Six Infrared Target Datasets

Frame Number	Target Size	Scene	Clutter
1	3×3	cloudy sky	strong edges of cloudy shape
2	2×5	sky	bright and heavy noise
3	3×4	sky	a circular clutter
4	4×5	cloudy sky	complicated cloud clutters
5	4×9 and 4×12	sea	bright noise and lined shape clutters
6	5×7, 5×6, 5×5 and 4×4	Sea and sky	heavy noise

2) Baseline Methods: This paper selects the existing 9 methods as comparison methods, including Top-Hat (Zeng et al. 2006), Fast-Saliency (Qi et al. 2016), VARD (Nasiri et al. 2017), LCM (Chen et al. 2014), MPCM (Wei et al. 2016), AAGD (Deng et al. 2016), NRAM (Zhang et al. 2018), Max-Median (Deshpande et al. 1999) and LIG (Zhang et al. 2018).

Table 2. Parameter Settings of Different Methods

No	Methods	Parameter settings
1	Top-Hat	Local window size: 5×5
2	Fast-Saliency	Gaussian filter size: 5×5 , $\sigma = 1.5$
3	VARD	External, Guard and Internal Patch size: 15×15 , 11×11 , 7×7
4	LCM	Local window size: 8×8 , sliding steps of row and column: 4
5	MPCM	$L = 3$, Local window size: 3×3 , 5×5 , 7×7
6	AAGD	Internal and external window size: 3×3 , 19×19
7	NRAM	Patchsize: 50×50 , slidingstep: 10, $\lambda = 1/\sqrt{\min(m, n)}$, $\mu^0 = \sqrt[3]{\min(m, n)}$, $c = \sqrt{\min(m, n)}/2.5$
8	Max-Median	Local window size: 15×15
9	LIG	$K=0.2, n=19$
10	Proposed	window size: 11×11 , $\mu = 100$

3) Evaluation Metrics:

$$BSF = \frac{\sigma_{in}}{\sigma_{out}}, \quad (18)$$

$$CG = \frac{C_{out}}{C_{in}}, \quad (19)$$

σ_{in} and σ_{out} represent the intensity standard deviation of the source image and the processed image respectively; C_{in} and C_{out} are the contrast between the source image and the processed image. The contrast is defined as Eq. (20):

$$C = |f_t - f_b|, \quad (20)$$

where, f_t is the maximum intensity value of the target area and f_b is the average value of the background area. This paper uses BSF and CG as evaluation indicators. BSF can evaluate the clutter suppression ability of the algorithm globally, and its value is proportional to the clutter suppression effect; CG can evaluate the enhancement effect of the target, and its value is proportional to the target enhancement effect.

$$TPR = \frac{\text{number of true target detection}}{\text{total number of actual targets}}, \quad (21)$$

$$FPR = \frac{\text{number of detected false targets}}{\text{total number of pixels in images}}, \quad (22)$$

TPR and FPR represent True Positive Rate and False Positive Rate, which are used to draw ROC curve and evaluate the algorithm the closer the ROC curve is to the upper left corner, the better the algorithm ability is.

3.2 Qualitative Comparisons

Fig.4 compares the detection results of the source image based on the LMPF, the LMGF and the weighted method, the weighted method represents P_w . It can be clearly seen that in Frame1, Frame2 and Frame4, the detection results of the LMPF still retain the point shadow clutter. So the clutter suppression ability of the LMGF is better than that of the LMPF. But in terms of target edge retention and target enhancement in the LMPF has better performance than the LMGF, except in Frame5 due to the influence of a lot of strong light and sea clutter, the LMPF is slightly lower than that of the LMGF.

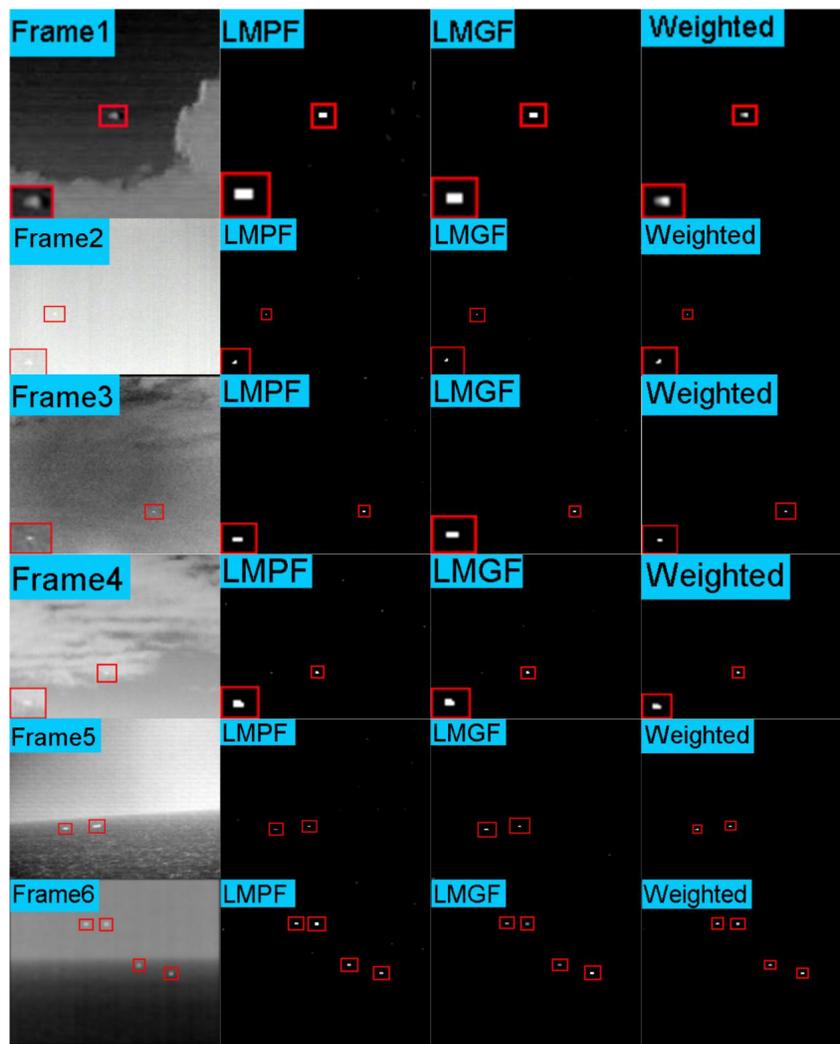


Fig. 4. The detection results of each stage of the algorithm (Each image uses a red rectangular frame to display the real target area and an enlarged view of the target area in the lower left corner of the image).

In order to verify the performance of this algorithm, this paper selects 6 sets of data sets for detection experiments, including single-target and multi-target image sequences. To qualitatively analyze the effect of each step of the proposed method, the source image is processed by LMPF, LMGF and the weighted model respectively, and BSF and CG of the outputs are calculated. Table 3 shows the above results. It can be clearly seen that the weighted model-based results are better than those of the LMPF-based or the LMGF-based, which indicate that the weighted model of the proposed method can effectively enhance the target and suppress background clutter.

Table 3. BSF and CG Values of Each Step of the Method in All Image Frames

Methods	LMPF	LMGF	Weighted	
BSF	Frame 1	5.640	5.754	9.480
	Frame 2	4.955	9.277	12.370
	Frame 3	8.396	9.069	12.003
	Frame 4	4.474	5.328	7.112
	Frame 5	13.748	14.020	19.690
	Frame 6	7.385	9.981	9.983
CG	Frame 1	2.486	2.470	2.486
	Frame 2	8.121	8.121	8.121
	Frame 3	2.671	2.457	2.682
	Frame 4	5.250	5.248	5.257
	Frame 5	1.563	2.013	2.252
	Frame 6	3.344	2.297	3.564

In order to verify the performance of the algorithm in this paper, 6 sets of data sets are selected as the experimental data, including single-target and multi-target image sequences. And 9 methods were selected for comparative experiments, including Top-Hat, Max-Median, Fast-Saliency, LCM, MPCM, VARD, AAGD, NRAM and LIG methods. Fig.5 and Fig.6 are the detection results of 6 data sets. The four data sets in Fig. 5 are all single target sequences, and the background includes high-intensity light, cloud clutter and complex irregular clutter. The two data sets in Fig.6 are two-target and four-target image sequences respectively, which can verify the multi-target detection ability of the algorithm.

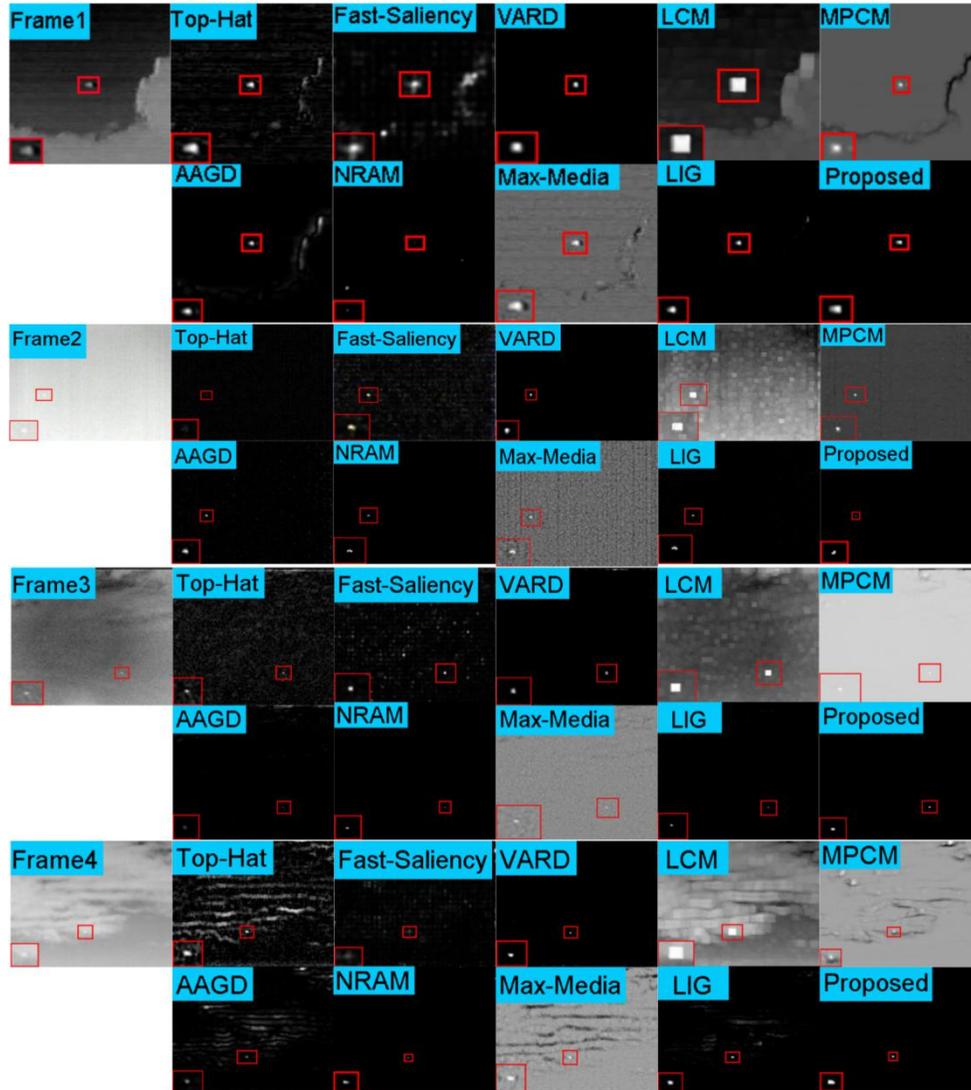


Fig. 5. Four original image sequences and detection results under different methods (Each image uses a red rectangular frame to display the real target area and an enlarged view of the target area in the lower left corner of the image).

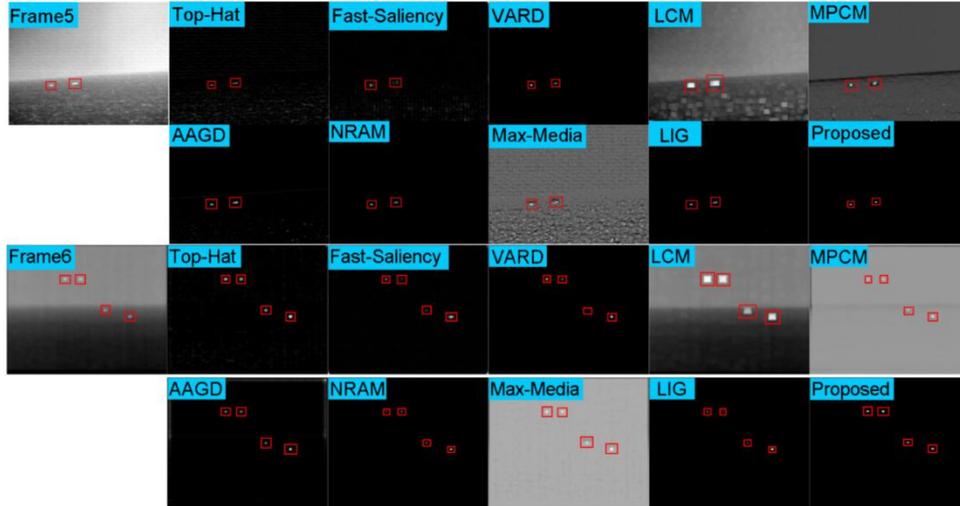


Fig. 6. Two original image sequences and detection results under different methods (Each image shows the real target area with a red rectangular frame).

The source images in Fig.5 are 4 sets of image sequences in the sky. In Frame1, the sky and the irregular cloud edges are the background, and there are a lot of irregular cloud edge clutter; Frame2 is in a strong light background, the background brightness is close to the target; The background of Frame3 has a lot of ring-shaped point clutter; And the target in Frame4 is surrounded by clouds, and there is a lot of cloud clutter interference.

The detection results of the Top-Hat and the Max-Median methods have a lot of the clutter, which indicate that they have poor background suppression effect, and furthermore, the Top-Hat method loses the target in the Frame2; As a comparison, the detection results of the LCM and the MPCM methods not only have background clutter, but also have a high false alarm rate; In contrast, there are a few background clusters in the results of the Fast-Saliency method; The other comparison methods also have different problems, that is the AAGD method has a dim amount of linear clutter in the Frame1 and Frame4; The NRAM method loses its target in the Frame1; The LIG method has a small amount of background clusters in the Frame4. In general, this method effectively suppresses the background clutter and preserves the target contour, and has the best detection performance.

The experimental data set in Fig. 6 includes two-target and four-target image sequences of the sea and sky background, which can verify the multi-target detection performance. The Frame5 is a dual target image sequence with strong sky light and sea wave clutter. The background suppression effects of the Top-Hat method, the Fast-Saliency method, the VARD method, the AAGD method, the NRAM method and the LIG method are good; Because there are a lot of clutter in the detection results of the LCM method, the MPCM method and the Max-Median methods, the background suppression effects are poor. The Frame6 is a four-target image sequence under low brightness. The Fast-Saliency method and the VARD

method lose a target respectively; And the LCM, the Max-Median and the MPCM methods have poor background suppression effects; On the other hand, the targets detected by the NRAM and the LIG methods are blurry, while the contours of the Top-Hat and the AAGD are clear. Compared with the comparative methods, this method has obvious advantages in background clutter suppression and target enhancement, which shows better performance in all data sets.

3.3 Quantitative Comparisons

The BSF, CG, ROC curves and the average running time of the algorithm are used to further evaluate the detection performance of this method. The BSF evaluates the background suppression capability, while the CG evaluates the target enhancement capability of the algorithm. Table 4 shows the BSF, the CG and the average running time of the algorithm.

Table 4. BSF, CG and Average Running Time of Different Algorithms under Each Image Frame. (The red font indicates the best result)

Methods	Proposed	Top-Hat	Fast-Saliency	VARD	LCM	MPCM	AAGD	NRAM	Max-Media	LIG	
BSF	Fra1	9.480	3.754	2.726	8.718	1.356	5.947	5.111	7.432	6.075	8.842
	Fra2	12.370	5.190	1.686	6.229	0.414	3.218	4.090	7.322	1.045	6.107
	Fra3	12.003	1.648	2.271	5.634	1.339	2.514	2.141	24.345	2.313	19.925
	Fra4	10.099	0.723	2.613	7.578	0.709	2.080	1.656	7.112	1.227	3.499
	Fra5	19.690	9.436	9.350	17.215	1.730	7.819	12.314	18.115	4.407	17.257
	Fra6	9.983	7.683	10.346	13.714	1.065	5.487	5.638	17.825	4.364	16.589
CG	Fra1	2.486	2.385	2.268	2.485	1.914	1.634	2.476	0.565	1.502	2.486
	Fra2	8.121	0.948	7.555	8.105	3.368	6.090	8.050	8.089	4.907	8.093
	Fra3	2.682	2.473	2.605	2.325	1.588	0.479	0.698	2.188	1.228	2.682
	Fra4	5.257	4.738	1.334	5.239	1.714	1.752	2.511	3.999	2.088	5.244
	Fra5	2.252	0.962	0.608	1.199	1.264	1.175	2.206	1.943	1.453	2.013
	Fra6	3.564	3.497	1.701	3.005	1.893	0.862	2.444	1.486	1.100	1.571
Time	Fra1	0.580	0.570	0.026	0.067	0.116	0.103	0.049	0.830	0.725	0.694
	Fra2	1.140	0.520	0.093	0.073	0.159	0.137	0.059	3.810	3.234	2.499
	Fra3	1.030	0.430	0.069	0.069	0.150	0.126	0.052	3.110	2.658	2.213
	Fra4	0.820	0.510	0.078	0.080	0.152	0.131	0.066	2.040	1.983	1.716
	Fra5	1.250	0.470	0.106	0.082	0.167	0.153	0.064	5.920	3.472	3.175
	Fra6	0.840	0.420	0.075	0.076	0.146	0.133	0.059	4.090	2.577	2.562

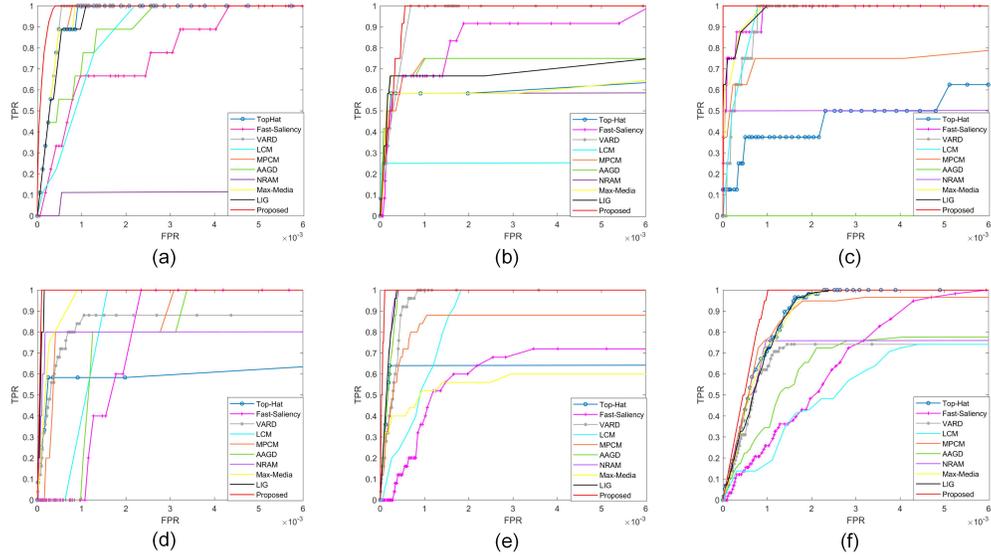


Fig. 7. ROC curves of different image frames.

In Table 4, the BSF of the proposed method performs best in the Frame1, the Frame2, the Frame4, and the Frame5, respectively, and the NRAM method performs best in the Frame3 and the Frame4. This shows that the background suppression ability of the proposed method is higher than the comparison methods.

The CG is the target local contrast gain. The larger the CG value, the better the background suppression and target enhancement performance. The CG of the proposed method is the best in all image sequences.

In terms of the running time comparison, the Fast-Saliency takes the shortest time, and the AAGD takes the shortest time in the remaining five image sequences; relatively, the NRAM method takes the longest time in all image sequences. The running time of this method is in the middle, taking into account the detection accuracy and computing efficiency.

Fig.7 shows the comparison result of the ROC curve of the algorithm. It can be clearly seen that the ROC curve of this method has the largest value among all methods.

3.4 Parameter Sensitivity Analysis

This method regards target detection as a clustering problem of unsupervised learning, so the selection of the number of candidate points is the main parameter that determines the performance of the proposed method. While the weight between adjacent pixels is a key factor in locating the target contour. In this method, the parameter μ is used to determine the weight between the pixels. In this section, the experiments are performed to analyze the influence of the parameters on the test results.

Considering that the dim targets are often submerged in the clutter, in the proposed method, it is assumed that most of the seed points are false targets, and a few seed points are

target points, which makes the target points an abnormal points. Fig.8 shows the TPR diagrams when the number of the candidate points m is different in six groups of image sequences. According to the experimental results, this proposed method sets m to 12.

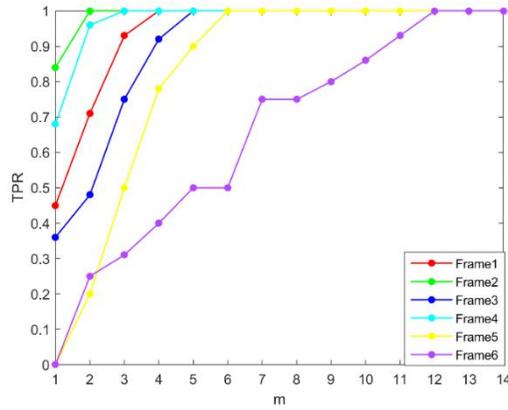


Fig. 8. TPR curves of different parameters m in this algorithm.

Fig.9 shows the distribution of the ROC curve with different values of the parameter μ in the proposed method. It can be clearly seen that if the parameter is set to 100, the ROC curve results are better in all image sequences. If the parameter μ is set to 200, the curve is slightly lower. And if the value is set to 0 or 300, the target may be missed or the false targets may be counted, which cause poor detection. So the parameter is set to 100.

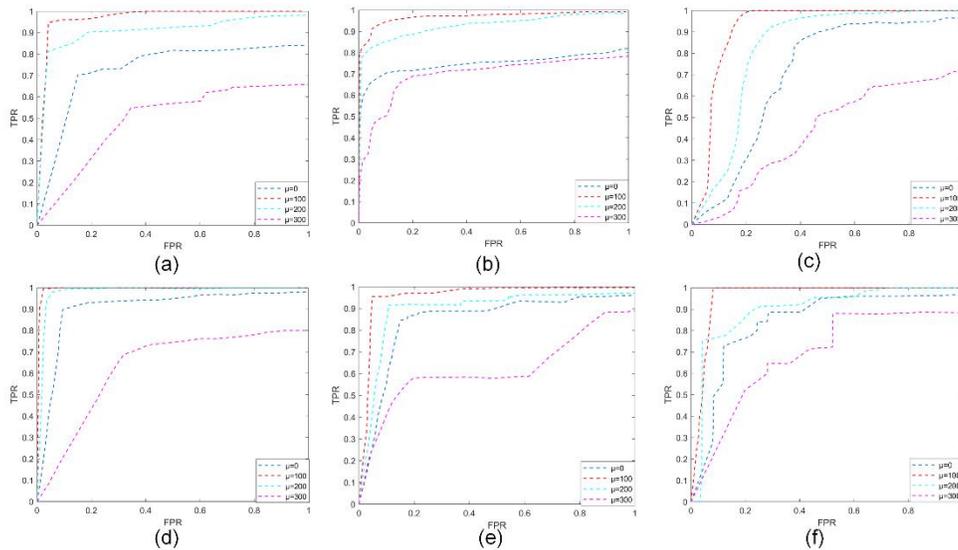


Fig. 9. TPR curves of different parameters μ in this algorithm.

4. Conclusion

In this paper, the density peak algorithm is used to traverse each pixel to establish a density-distance spatial decision map. On this basis, candidate points are selected according to the proportional relationship between noise points and dim targets. Then a local mosaic model is built with the candidate points as the center, and the targets can be screened from the candidate targets based on the mapping relationship between the LMPF and the LMGF, which narrows the target search range and reduces the amount of calculation. The proposed method can also accurately extract the targets for the multi-target image sequences. The experimental results show that the proposed method is better than the comparison method in performance, not only can suppress clutter and enhance the target, but also have a wide range of applications. Next, we will find a new positioning method for the target region, and the running time can be further reduced, and a robust feature space for screening real targets can be built.

Acknowledgments.

The authors thank the anonymous reviewers and editors for the very constructive comments. This work was supported by the National Natural Science Foundation of China(61962046, 61663036, 61841204). Inner Mongolia Outstanding Youth Cultivation Fund (2018JQ02). Inner Mongolia Natural Science Foundation (2015MS0604).

Disclosures. The authors declare no conflicts of interest.

References

- Zeng, M., Li, J., Peng, Z.: The design of top-hat morphological filter and application to infrared target detection. *Infrared Phys. Technol.* **48**, 67-76 (2006)
- Wang, X., Yang, L. T., Li, H., Lin, M., Han, J., Apduhan, B. O.: NQA: A nested anti-collision algorithm for RFID systems. *ACM Trans. Embed. Comput. Syst.* **18**, 1-21 (2019)
- Chen, C.L.P., Li, H., Wei, Y.T., Xia, T., Tang, Y.Y.: A local contrast method for small infrared target detection. *IEEE Trans. Geosci. Remote Sens.* **52**, 574-581 (2014)
- Ye, Y., Shan, J., Bruzzone, L., Shen, L.: Robust registration of multimodal remote sensing images based on structural similarity. *IEEE Trans. Geosci. Remote Sens.* **55**, 2941-2958 (2017)
- Huang, Z., Chen, L., Zhang, Y., Yu, Z., Fang, H., Zhang, T.: Robust contact-point detection from pantograph-catenary infrared images by employing horizontal-vertical enhancement operator. *Infrared Phys. Technol.* **101**, 146-155 (2019)
- Dai, Y., Wu, Y.: Reweighted infrared patch-tensor model with both nonlocal and local priors

- for single-frame small target detection. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **10**, 3752-3767 (2017)
- Lv, P.-Y., Sun, S.-L., Lin, C.-Q., Liu, G.-R.: Space moving target detection and tracking method in complex background. *Infrared Phys. Technol.* **91**, 107-118 (2018)
- Wan, M., Gu, G., Cao, E., Hu, X., Qian, W., Ren, K.: In-frame and inter-frame information based infrared moving small target detection under complex cloud backgrounds. *Infrared Phys. Technol.* **76**, 455-467 (2016)
- Zhang, P., Wang, X., Wang, X., Fei, C., Guo, Z.: Infrared Small Target Detection Based on Spatial-Temporal Enhancement Using Quaternion Discrete Cosine Transform. *IEEE Access* **7**, 54712-54723 (2019)
- Deshpande, S. D., Er, M. H., Venkateswarlu, R., Chan, P.: Max-mean and max-median filters for detection of small targets. in *Signal and Data Processing of Small Targets 1999*, (International Society for Optics and Photonics, 1999), 74-83.
- Soni, T., Zeidler, J. R., Ku, W. H.: Performance evaluation of 2-D adaptive prediction filters for detection of small objects in image data. *IEEE Trans. Image Process.* **2**, 327-340 (1993)
- Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. in *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*, (IEEE, 1998), 839-846
- Barnett, J.: Statistical analysis of median subtraction filtering with application to point target detection in infrared backgrounds. in *Infrared Systems and Components III*, (International Society for Optics and Photonics, 1989), 10-18
- Wei, Y., You, X., Li, H.: Multiscale patch-based contrast measure for small infrared target detection. *Pattern Recognit.* **58**, 216-226 (2016)
- Nie, J., Qu, S., Wei, Y., Zhang, L., Deng, L.: An infrared small target detection method based on multiscale local homogeneity measure. *Infrared Phys. Technol.* **90**, 186-194 (2018)
- Huang, S., Peng, Z., Wang, Z., Wang, X., Li, M.: Infrared small target detection by density peaks searching and maximum-gray region growing. *IEEE Geosci. Remote Sens. Lett.* **16**, 1919-1923 (2019)
- Xia, C., Li, X., Zhao, L.: Infrared small target detection via modified random walks. *Remote Sens.* **10**, 2004 (2018)
- Qin, Y., Bruzzone, L., Gao, C., Li, B.: Infrared small target detection based on facet kernel and random walker. *IEEE Trans. Geosci. Remote Sensing.* **57**, 7104-7118 (2019)
- Wang, G., Yang, J., Xu, J.: Granular computing: from granularity optimization to multi-granularity joint problem solving. *Granular Computing* **2**, 105-120 (2017)
- Qi, S., Xu, G., Mou, Z., Huang, D., Zheng, X.: A fast-saliency method for real-time infrared small target detection. *Infrared Phys. Technol.* **77**, 440-450 (2016)
- Nasiri, M., Chehresa, S.: Infrared small target enhancement based on variance difference. *Infrared Phys. Technol.* **82**, 107-119 (2017)

- 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.* **52**, 60-72 (2016)
- Zhang, L., Peng, L., Zhang, T., Cao, S., Peng, Z.: Infrared small target detection via non-convex rank approximation minimization joint $l_2, 1$ norm. *Remote Sens.* **10**, 1821 (2018)
- Zhang, H., Zhang, L., Yuan, D., Chen, H.: Infrared small target detection based on local intensity and gradient properties. *Infrared Phys. Technol.* **89**, 88-96 (2018)

Figures

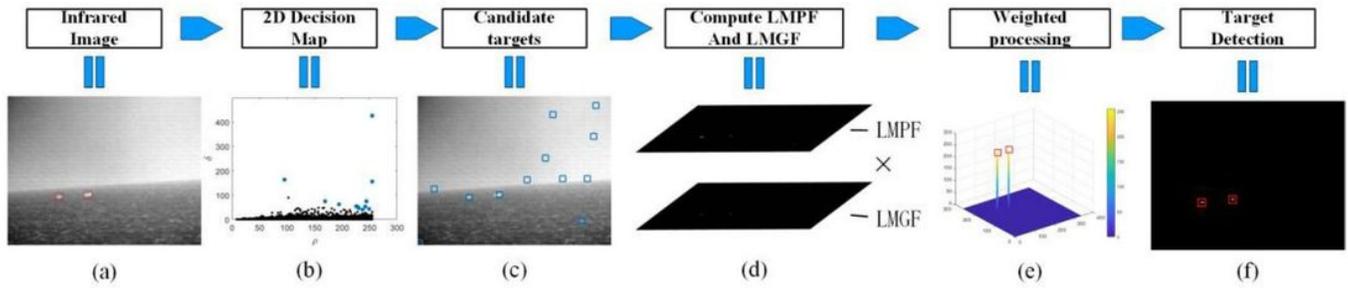


Figure 1

The entire flow chart of the proposed method.

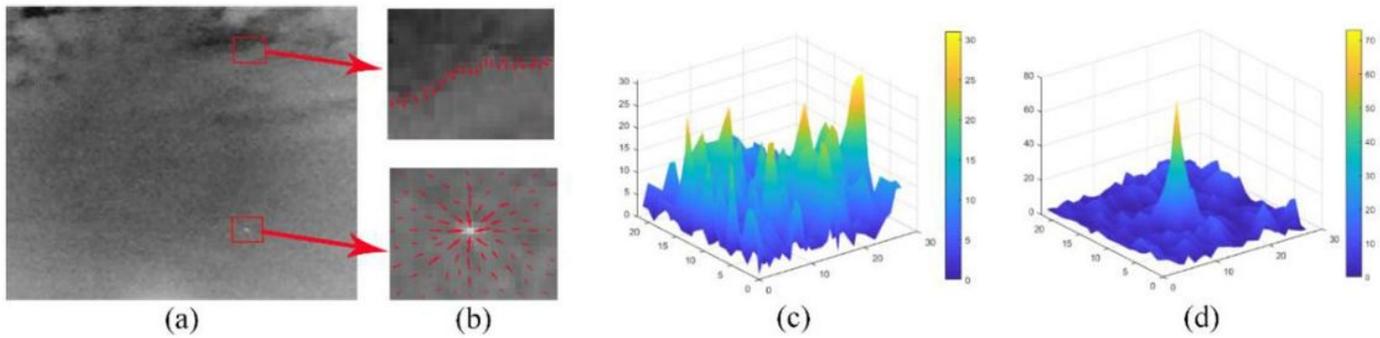
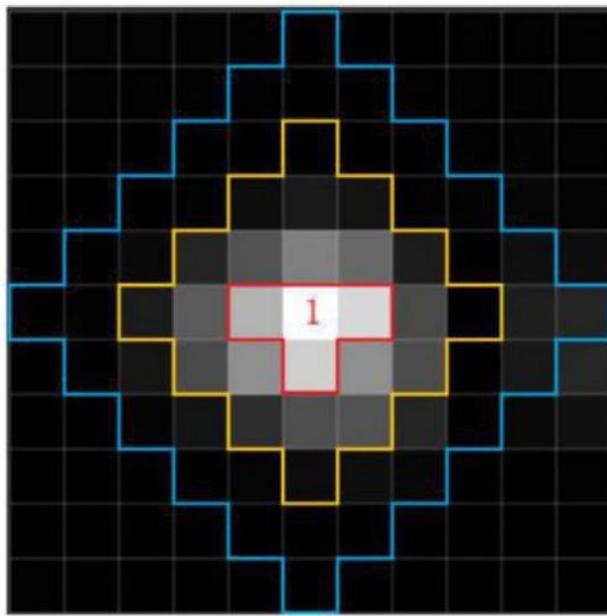
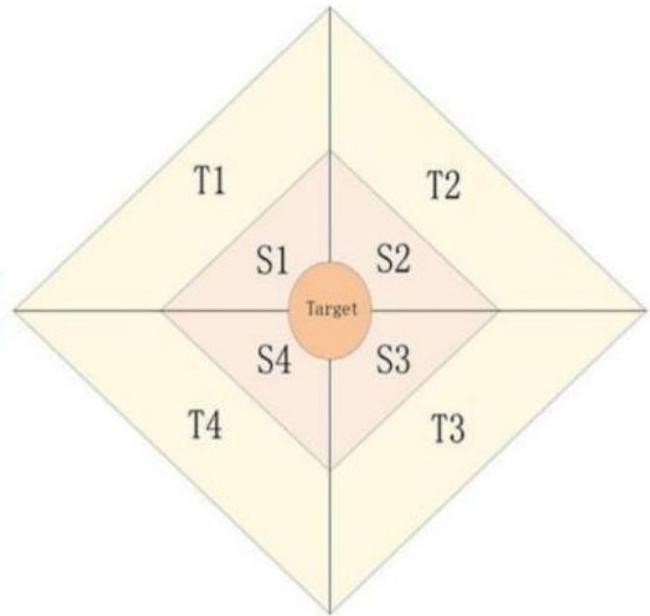


Figure 2

Background clutter, target and its three-dimensional gradient map in infrared image.



(a)



(b)

Figure 3

Local mosaic model. (a) The marking strategy in the local mosaic model; (b) Area division diagram of local mosaic model.



Figure 4

The detection results of each stage of the algorithm (Each image uses a red rectangular frame to display the real target area and an enlarged view of the target area in the lower left corner of

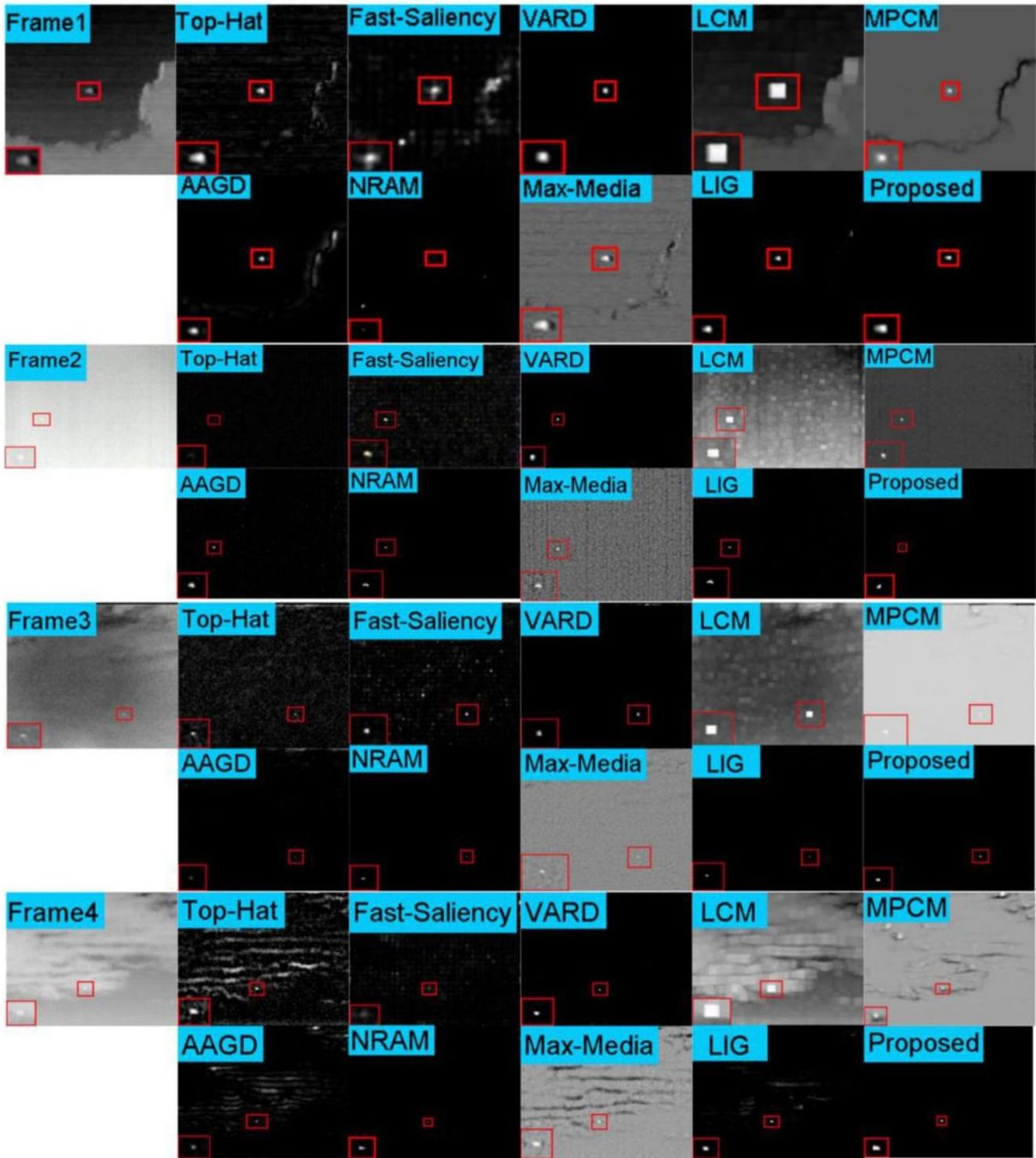


Figure 5

Four original image sequences and detection results under different methods (Each image uses a red rectangular frame to display the real target area and an enlarged view of the target area in the lower left corner of the image).

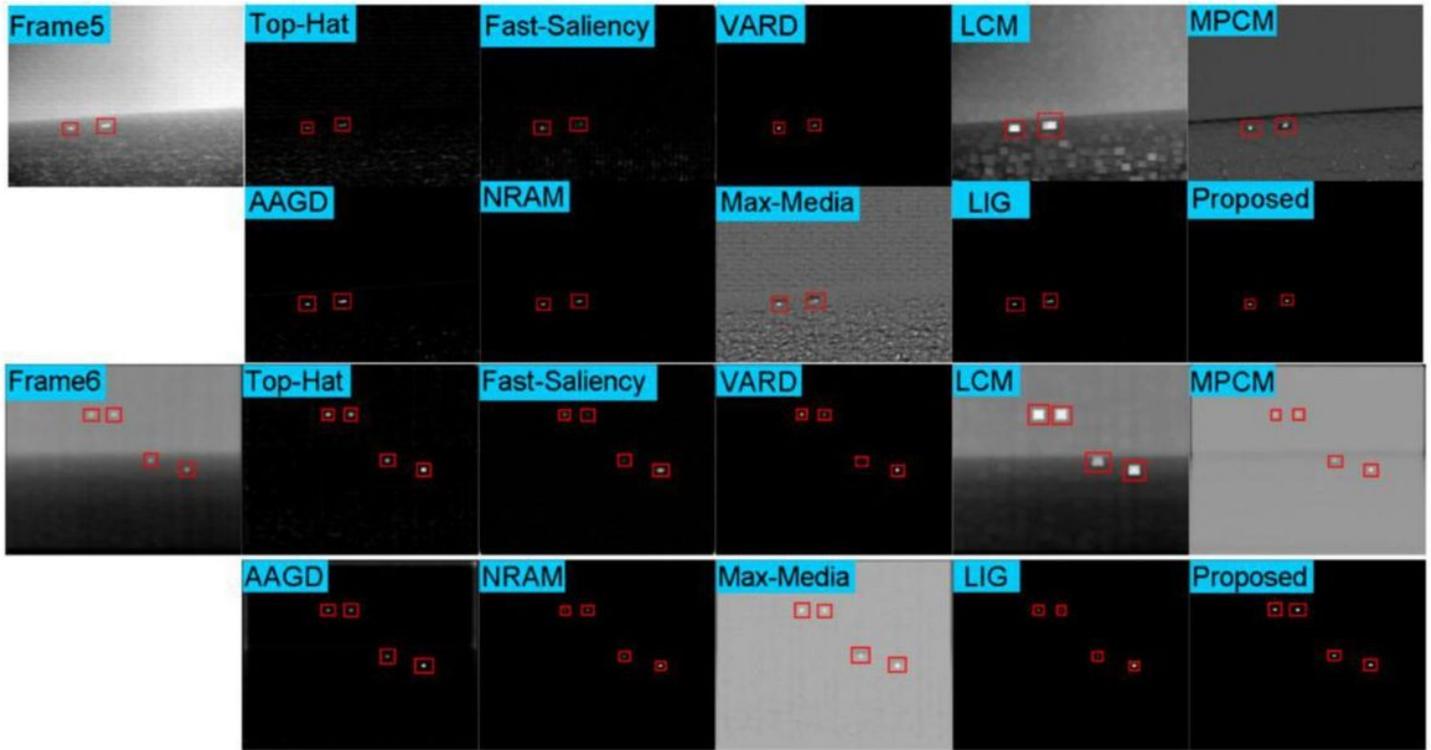


Figure 6

Two original image sequences and detection results under different methods (Each image shows the real target area with a red rectangular frame)

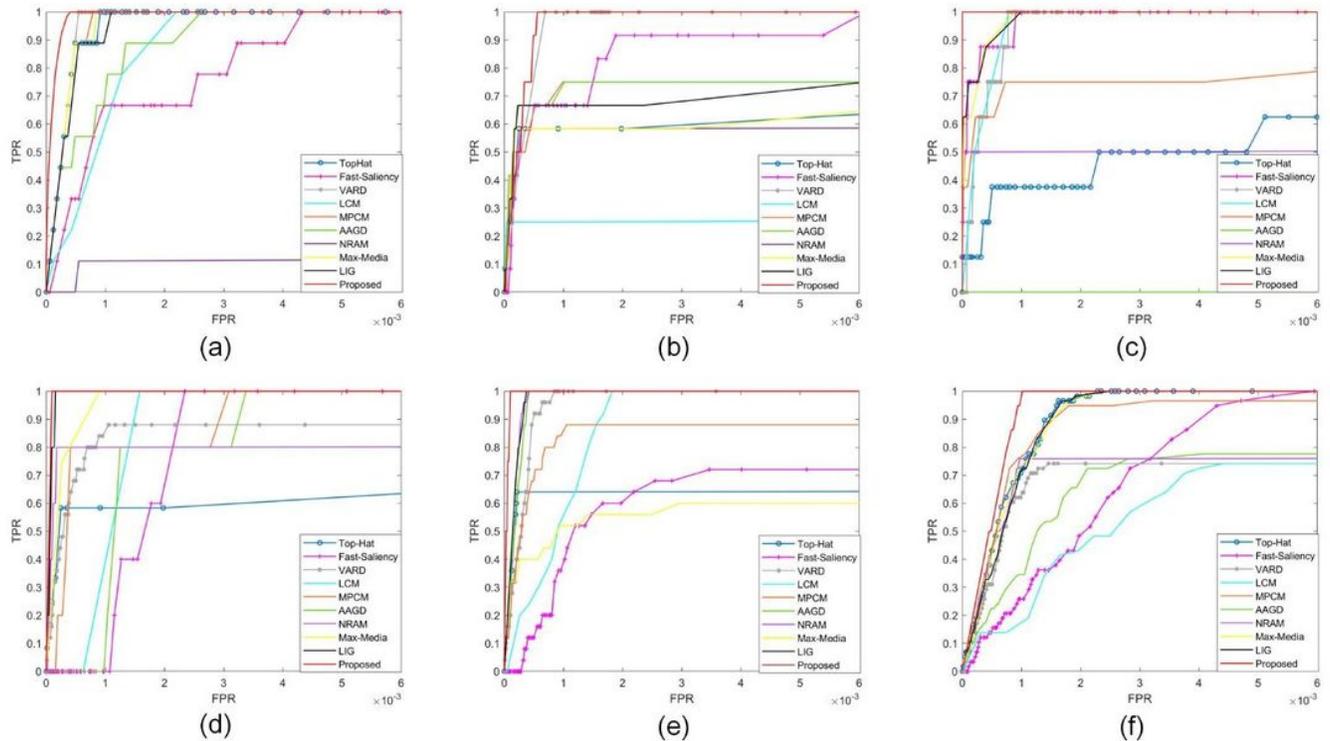


Figure 7

ROC curves of different image frames

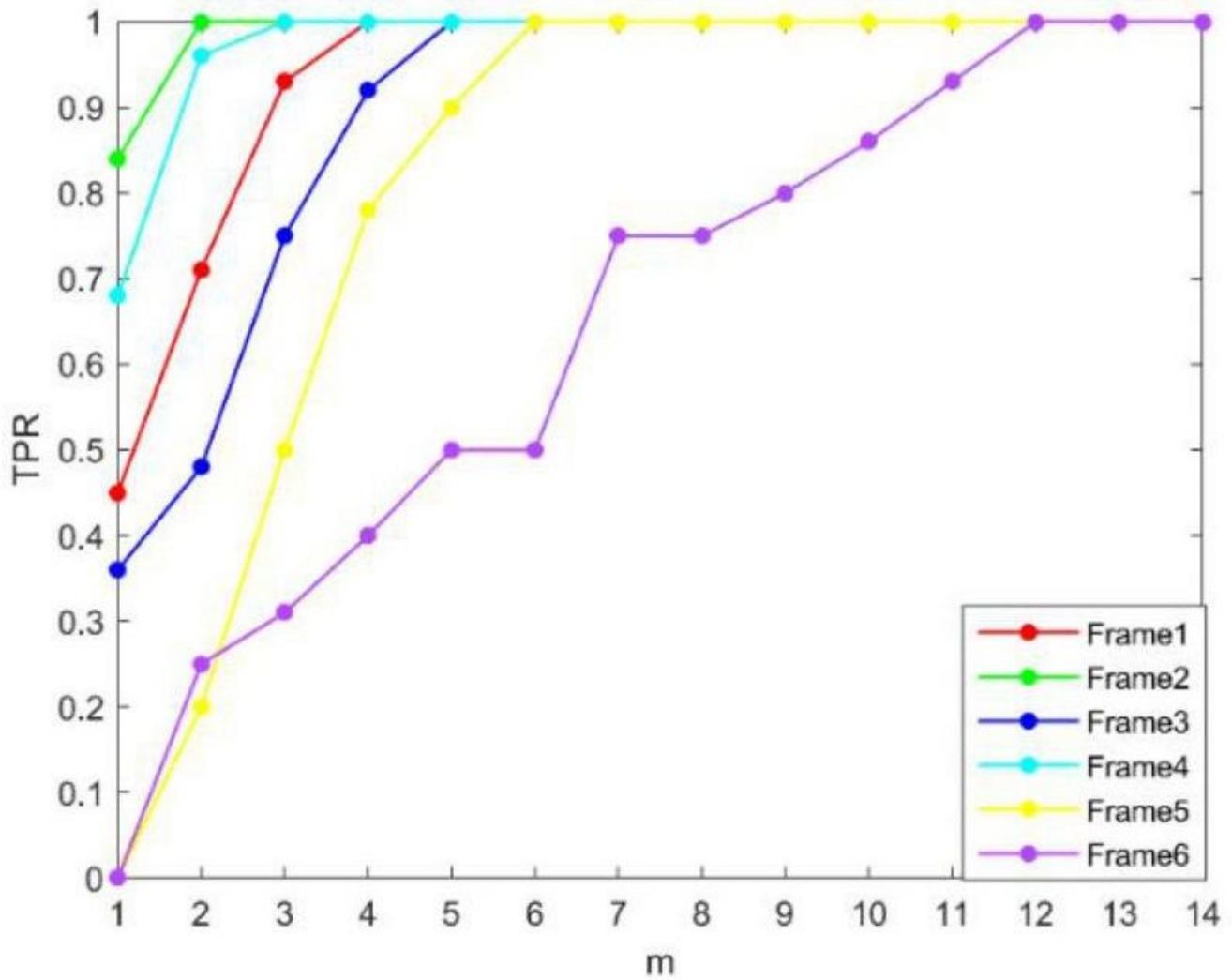


Figure 8

TPR curves of different parameters m in this algorithm

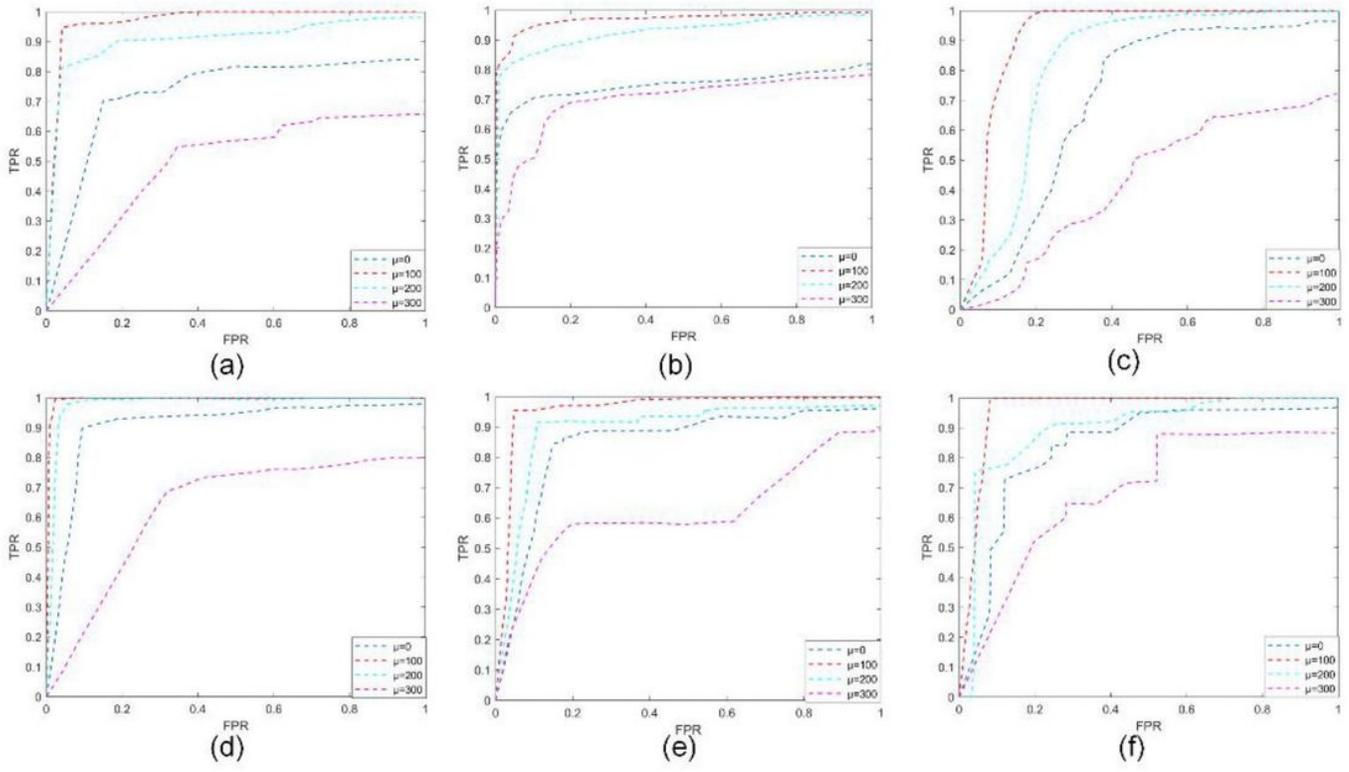


Figure 9

TPR curves of different μ in this algorithm.