

Small Target Detection Algorithm Based on Average Absolute Difference Maximum and Background Forecast

Zhenxue Chen · Guoyou Wang · Jianguo Liu · Chengyun Liu

Received: 15 September 2006 / Accepted: 24 October 2006 /
Published online: 23 January 2007
© Springer Science + Business Media, LLC 2007

Abstract Detecting small targets in clutter scene and low SNR (Signal Noise Ratio) is an important and challenging problem in infrared (IR) images. In order to solve this problem, we should do works from two sides: enhancing targets and suppressing background. Firstly, in this paper, the system utilizes the average absolute difference maximum (AADM) as the dissimilarity measurement between targets and background region to enhance targets. Secondly, it uses a predictor to suppress the background clutter. Finally, our approach extracts the interested small target with segment threshold. Experimental results show that the algorithm proposed has better performance with respect to probability of detection and less computation complexity. It is an effective small infrared target detection algorithm against complex background.

Keywords Small targets · Complex background · Average absolute difference maximum · Background forecast · SNR · Infrared images

1 Introduction

In space scene, when a target is far from the image sensor, it often takes small image size (usually only several pixels). This target is referred as small target. Its image intensity is very weak. Complex background means the space distribution of background image is not smooth, and the statistics of mean value and variance have not continuous invariance. The small target is apt to be submerged by noise because of low SNR. It has been regarded as a

Z. Chen (✉) · G. Wang · J. Liu
Institute of Pattern Recognition & Artificial Intelligence,
Huazhong University of Science and Technology, Luoyu Road 1037,
Wuhan 430074, People's Republic of China
e-mail: chenchenxue@163.com

C. Liu
College of Science,
Wuhan University of Science and Technology, Wuhan 430081,
People's Republic of China

difficult task in most cases to detect or trace a small target under the complex background [1–3]. Denney and de Figuriro presented a small target detection method based on the predication under the adaptive auto-regressive background [4]. Tom *et al.* put forward the morphologic operators for small target detection according to the prior knowledge of targets [5]. Peng and Zhou designed a 5×5 highpass template filter for real-time target detection [6]. Ye *et al.* provided a small target detection method based on wavelet transform modulus maxima [7]. Based on the analysis of various targets detection methods, Hilliard pointed out that a lowpass IIR filter has a better comprehensive performance for clutter prediction [8]. Ulisses, B.N. used morphological top-hat algorithm which employed the top-hat operator to form an estimate of the background image [9]. However, in practice, these methods have been only partially successful and have been shown to produce probability of false alarm. The reason is that these algorithms solve the problem from single side: enhancing targets or suppress background. In this paper, we present a novel small target detection algorithm based on AADM and background forecast (BF). This method uses AADM to enhance an object, and then it utilizes BF map to suppress the background. Finally, the algorithm can detect automatically the small targets according to the various segment thresholds. Experiments have proved that our method is effective and robust.

The organization of this paper is as follows. We describe first the theory of Average absolute difference maximum in Section 2 where we also analyze the effect of size of external window. In the next section, we introduce the background forecast algorithm in detail. Section 4 describes the procedure of small object detection. Experimental results are reported and discussed in Section 5. Conclusions are drawn in Section 6.

2 Average absolute difference maximum

As originally conceived, the problem addressed in this paper is to detect small objects from natural scenes and to derive the criteria of distinguishing objects from natural scenes. In essence, this involves determining the property of average absolute difference based on their neighboring pixels in natural scenes [10]. In many cases, the only visual brightness of small target in infrared image is greater than that of their surround neighborhood, despite the discrimination between object and the area of the neighborhood is frequently small. This also means that the target is conspicuous in a local region. Taking this fact in account, we use a double-window filter to enhance small target.

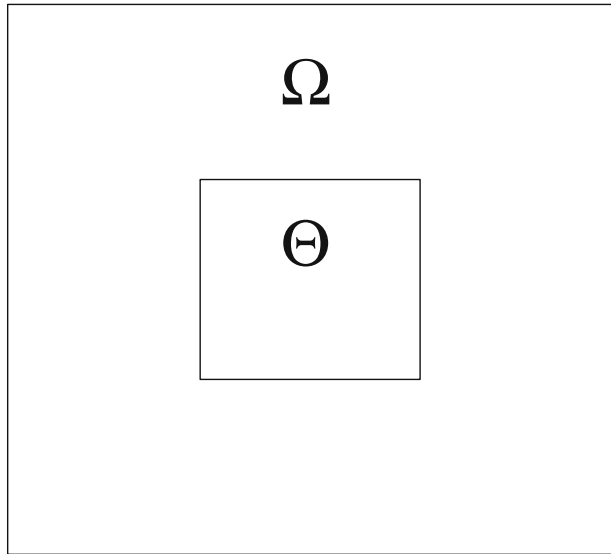
As show in Fig. 1, set Θ and Ω denote the pixels in the internal window and pixels between internal window and external window, respectively. If the size of external window is slightly bigger than that of the object and internal window can be changed, then the AADM can be defined as:

$$\Delta D_{\max} = \max_{\forall \Theta} \left\{ \Delta D | \Delta D = \left| \frac{1}{N_{\Theta}} \sum_{x \in \Theta} I_x - \frac{1}{N_{\Omega}} \sum_{y \in \Omega} I_y \right| \right\} \quad (1)$$

where N_{Θ} and N_{Ω} denote the numbers of pixels in set Θ and Ω , I_x and I_y denote the gray levels of pixel x and y , respectively.

It has been shown that AADM has a low value if set Θ and Ω include only the pixels in a background, and reach the largest response when set Θ and Ω cover only the pixels of an

Fig. 1 Object and scene windows.



object and background, respectively. If the target shape is not square, then Θ includes the pixels of object Θ_o and partial background Θ_b , that is,

$$\begin{aligned} \Delta D &= \left| \frac{1}{N_\Theta} \sum_{x \in \Theta} I_x - \frac{1}{N_\Omega} \sum_{y \in \Theta} I_y \right| = \left| \frac{N_{\Theta_o} \bar{I}_o + (N_\Theta - N_{\Theta_o}) \bar{I}_{b\Theta}}{N_\Theta} - \bar{I}_{b\Omega} \right| \\ &= \left| \frac{N_{\Theta_o}}{N_\Theta} \bar{I}_o + \frac{N_\Theta - N_{\Theta_o}}{N_\Theta} \bar{I}_{b\Theta} - \bar{I}_{b\Omega} \right| = \left| \frac{N_{\Theta_o}}{N_\Theta} (\bar{I}_o - \bar{I}_{b\Theta}) + (\bar{I}_{b\Theta} - \bar{I}_{b\Omega}) \right| \quad (2) \\ &\leq \left| \frac{N_{\Theta_o}}{N_\Theta} (\bar{I}_o - \bar{I}_{b\Theta}) \right| + |(\bar{I}_{b\Theta} - \bar{I}_{b\Omega})| \end{aligned}$$

where N_{Θ_o} , N_{Θ_b} denotes pixels of object and scene in Θ , respectively. \bar{I}_o , $\bar{I}_{b\Theta}$, and $\bar{I}_{b\Omega}$ are the average gray values of the object, the internal window, and external window, respectively. If the nature scene is a stable Gaussian noise, then $\bar{I}_{b\Theta} \approx \bar{I}_{b\Omega}$. ΔD reaches its extremum, when one of the following conditions holds.

$$\frac{N_{\Theta_o}}{N_\Theta} (\bar{I}_o - \bar{I}_{b\Theta}) = (\bar{I}_{b\Theta} - \bar{I}_{b\Omega}) \Delta D = 2|\bar{I}_{b\Theta} - \bar{I}_{b\Omega}| \quad (3)$$

$$\frac{N_{\Theta_o}}{N_\Theta} (\bar{I}_o - \bar{I}_{b\Theta}) = 0 \Delta D = |\bar{I}_{b\Theta} - \bar{I}_{b\Omega}| \quad (4)$$

$$\bar{I}_{b\Theta} - \bar{I}_{b\Omega} = 0 \Delta D = \left| \frac{N_{\Theta_o}}{N_\Theta} (\bar{I}_o - \bar{I}_{b\Theta}) \right| = \frac{N_{\Theta_o}}{N_\Theta} |\bar{I}_o - \bar{I}_{b\Omega}| \quad (5)$$

From (3), if $|\bar{I}_o - \bar{I}_{b\Theta}| \gg 0$ and $N_{\Theta_o} \ll N_\Theta$ for an object with concave contour, then ΔD is small. From (4), if $N_{\Theta_o} = 0$ or $\bar{I}_o = \bar{I}_{b\Theta}$ for no object, $\Delta D \approx 0$. In (5), for $\frac{N_{\Theta_o}}{N_\Theta} \leq 1$, so when $N_{\Theta_o} \approx N_\Theta$, $\Delta D \approx |\bar{I}_o - \bar{I}_{b\Omega}|$ and approaches to the maximum value.

It is obvious that (4) corresponding to the thing that internal window and external window are both background, here ΔD is the smallest; (5) corresponding to the thing that internal window is target and external window is background, here ΔD is the largest.

There, another problem that we must consider is how to decide the size of external window. If the size of external window (we call as “field”) is too small, the potential target may be masked by background. On the other hand, if the size of field is too large, the operation will consume more computation. The small target in infrared image often occurred several pixels, normally on more than 7 pixels, so the appropriate field area should greater 36 pixels and no more than 81 pixels.

Figure 2 shows that the probability of detection Pd varies with size of the external windows.

As shown in Fig. 2, when $L_{max}=9$, the Pd reached its maximum. If we select $L_{max}>9$, the Pd is not vary, but the computation complexity is greatly increased.

3 Background forecast

An ideal clutter suppression algorithm would be capable of transforming an arbitrary inhomogeneous and no stationary background into an approximately stationary, homogeneous Gaussian noise background without reducing the effective target signal power. In practice, preserving the target signal energy is the critical consideration for weak targets.

As to the small targets, it is difficulty that their other features, except gray feature, are found. Therefore we must utilize the gray distributed and fluctuating feature of background to detect the small targets, except their self-feature of gray. Suppose the image background doesn't have strong fluctuation, but the targets always have extrusive change at itself region. So we can consider that if the pixel belongs to the background, its gray must be forecasted by others pixels around it. As for the pixel that belongs to the target has worse relativity with its surrounding background. According to this theory, we propose the background forecast algorithm to suppress the background clutter.

The forecast model can be shown as follow:

$$Y(m, n) = \sum_{l,k \in S_j} W_j(l, k) X_0(m - l, n - k) \tag{6}$$

$$m = 0, 1, \dots, M - 1 ; n = 0, 1, \dots, N - 1$$

where Y and X_0 denote the reference image and the forecast image, respectively. Their sizes are both $M \times N$. W_j is the j^{th} power matrix. S_j is a region around certain pixel.

Fig. 2 Probability of detection $Pd \sim L_{max}$.

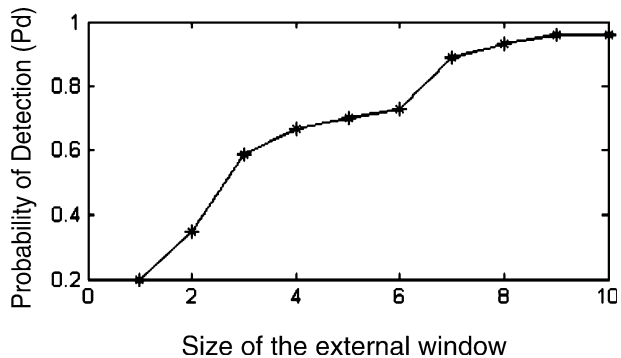
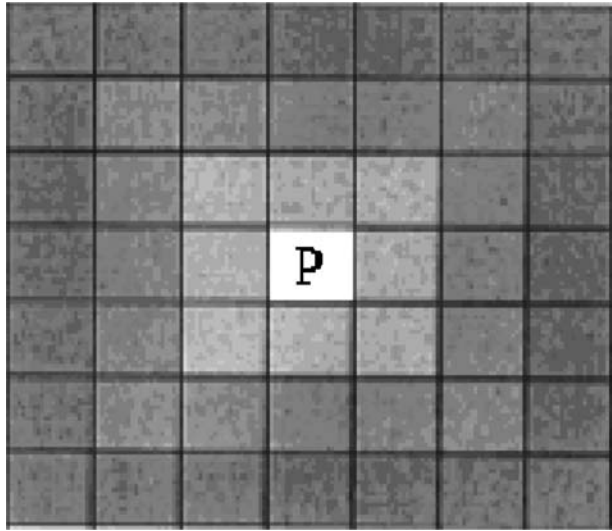


Fig. 3 Power selecting sketch map.



52	51	52	50	53	53	53	43	52	42	51	50	53	54	54
51	51	53	42	53	52	53	43	41	52	61	50	63	64	54
53	52	51	50	52	52	53	44	62	62	51	52	54	53	56
72	51	12	10	13	13	13	13	12	12	11	10	13	44	34
62	51	12	10	13	13	15	15	12	12	11	10	13	54	74
72	61	12	10	13	13	16	18	12	12	11	10	13	55	65
72	51	12	10	13	13	13	13	12	12	11	10	13	44	34
62	51	12	10	13	13	55	55	12	12	11	10	13	54	74
72	61	12	10	13	13	56	58	12	12	11	10	13	55	65
72	51	12	10	13	13	13	13	12	12	11	10	13	44	34
62	51	12	10	13	13	15	15	12	12	11	10	13	54	74
72	61	12	10	13	13	16	18	12	12	11	10	13	55	65
51	51	53	42	53	52	53	43	41	52	61	50	63	64	54
52	51	52	50	53	53	53	43	52	42	51	50	53	54	54
51	51	53	42	53	52	53	43	41	52	61	50	63	64	54

Fig. 4 A diagram with a target and the fluctuating background.

50	46	45	48	51	53	51	46	52	47	51	47	50	54	51
49	51	45	42	45	52	50	48	47	52	55	50	56	60	54
50	52	49	47	46	46	50	46	57	57	51	52	52	53	52
67	50	31	27	27	28	27	28	29	25	25	26	26	30	31
55	50	32	20	23	22	22	23	21	20	21	20	19	44	50
56	50	19	18	18	18	17	18	15	15	14	15	17	55	60
66	49	17	15	15	16	16	16	16	17	18	17	17	35	31
54	47	17	17	16	18	21	20	18	19	17	18	19	50	61
66	57	19	18	16	19	24	24	19	19	16	17	19	51	54
65	44	18	17	17	16	18	18	19	17	16	17	17	32	28
54	43	16	16	16	17	17	18	15	15	14	15	15	45	65
54	51	17	16	16	17	19	20	18	16	17	17	19	41	56
41	41	42	34	35	36	41	31	32	41	41	38	42	43	45
44	47	51	46	52	50	45	35	50	47	46	50	54	56	54
51	48	51	40	53	52	54	41	40	50	54	49	57	58	50

Fig. 5 A diagram with background forecast.

The difference between forecast image and reference image is

$$E(m, n) = |X(m, n) - Y(m, n)| \tag{7}$$

where $E(m, n)$ denotes the difference image.

In order to express the adaptability, W_j can be defined as

$$W_j(l, k) = \frac{(r(l, k))^n}{\sum_{l, k \in S_j} (r(l, k))^n} \tag{8}$$

where $r(l, k)$ is the geometry distance from region background pixel to the forecasted pixel, $n=1, 2, \dots$. It is obvious that the less distance between the forecast pixel and the environment pixel, the smaller the power $W_j(l, k)$ is. It is shown as Fig. 3. There, the color of the black is deeper means that the power is bigger.

For example, a diagram containing a target and the fluctuating background is shown as Fig. 4.

The target lies in the thick black pane. It is clearly seen that the small target is merged by the fluctuating background. We utilize the forecast model as mask to do convolution with the diagram. After convolution operation, the diagram becomes another diagram which is shown as Fig. 5.

If we subtract the forecasted background from the original image, we can obtain the candidate target.

But in the actual scene, the background is more complex and presents some block form. As to this reason, we present the segmentation-block background forecast algorithm. The theory of this algorithm is the background region will be divided up four parts centering the

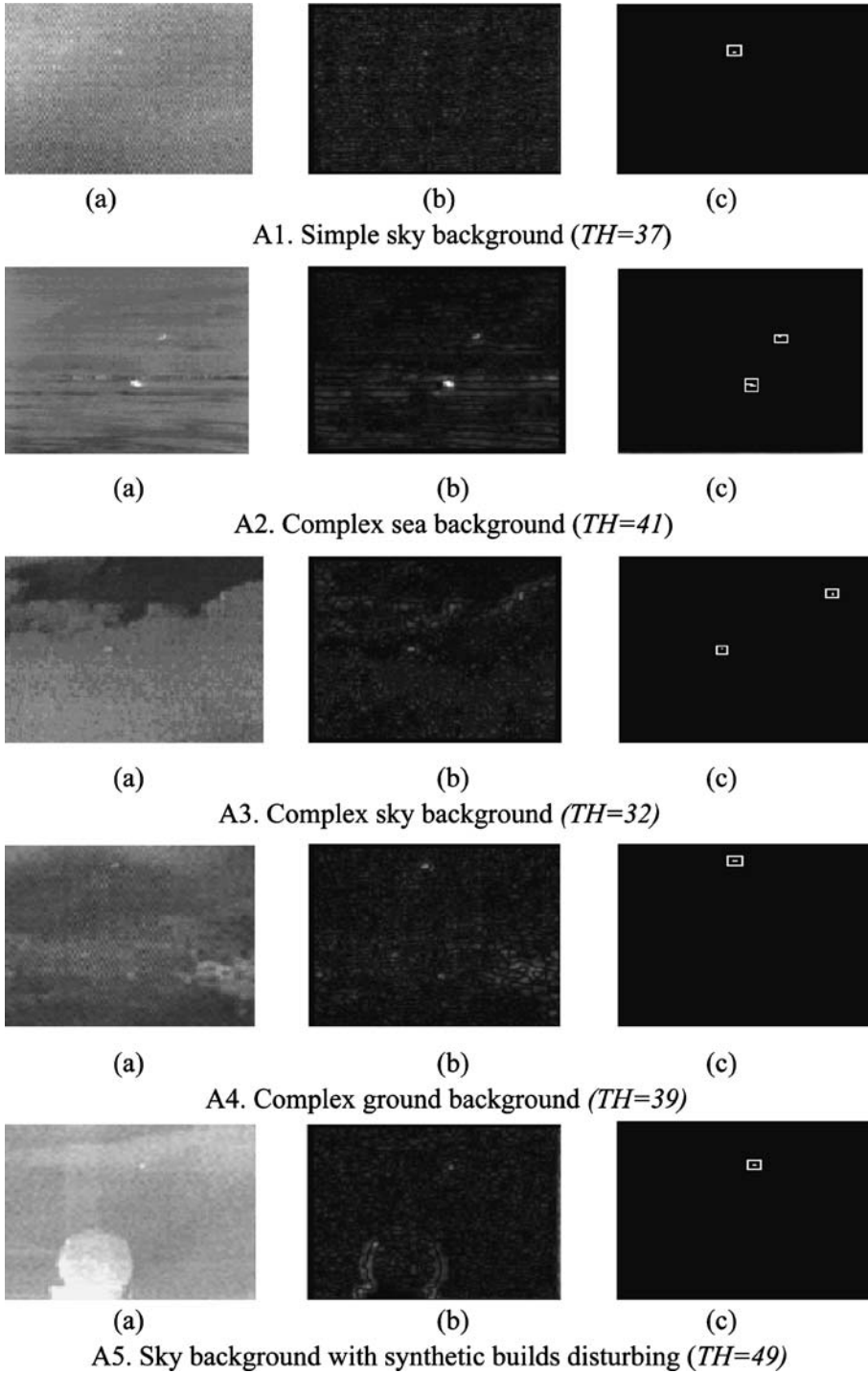


Fig. 6 Results of object detection. (a) Original image; (b) Results of AADM filter; (c) Typical detection results after BF.

forecast pixel, and the each part corresponds to the each quadrant in the Euclid space. The four forecast values are

$$\begin{aligned}
 Y_1(m, n) &= \sum_{l=0}^{M-1} \sum_{k=0}^{N-1} W_j(l, k) X_0(m + l, n + k) \\
 Y_2(m, n) &= \sum_{l=0}^{M-1} \sum_{k=0}^{N-1} W_j(l, k) X_0(m - l, n + k) \\
 Y_3(m, n) &= \sum_{l=0}^{M-1} \sum_{k=0}^{N-1} W_j(l, k) X_0(m - l, n - k) \\
 Y_4(m, n) &= \sum_{l=0}^{M-1} \sum_{k=0}^{N-1} W_j(l, k) X_0(m + l, n - k)
 \end{aligned}$$

We can forecast the background by the four quadrant pixels, respectively. And then we select the value that is a power average of four quadrants as the final forecast value, that is

$$Y(m, n) = \left\{ \left(P \cdot Y_L(m, n) + \sum_{\substack{k=1 \\ k \neq L}}^4 Y_k(m, n) \right) / (P + 3) \right\} \quad (P2) \quad (9)$$

where $Y_L(m, n) = Arg \left\{ \min_{k=1}^4 |Y_k(m, n) - X(m, n)| \right\}$, P is a power coefficient.

The prediction model of Eq. 9 ensures that the pixel at any directional edge can be forecasted in the higher luminance region or lower luminance region. This method can greatly reduce the probability of false alarm at the boundary between light and dark.

4 Algorithms of small object detection

When the size of the target is changing, to determine the correct deviation between target and background, the inner neighbor set Θ and the external neighbor set Ω of point (x, y) must be adaptable to target size; that is, first, we can give the enough scale of the external neighbor set Ω of point (x, y) ; second, we change the scale of the inner neighbor set Θ , and finally, we compute the maximum of average absolute difference. Therefore, our algorithm can be described in detail as follows:

- Step1: For any pixel (x_0, y_0) in the image, set an external window with the fixed size $L_{\max} \times L_{\max}$ around the pixel.
- Step2: Around the pixel (x_0, y_0) , we chose a series of small internal windows with the sizes of $(2 \cdot l + 1) \times (2 \cdot l + 1), l \in L, L = \{l | (2 \cdot l + 1) < L_{\max}\}$ and compute the gray average absolute difference ΔD_l , respectively.
- Step3: Define the max value $\Delta D_{\max} = \max_{l \in L} \{\Delta D_l\}$ and make ΔD_{\max} replace the value of (x_0, y_0) .
- Step4: Process all pixels in whole image in turn and obtain the gray average absolute difference map.
- Step5: Adopt the background forecast algorithm and obtain the background image.
- Step6: Subtract between the original image and background image and obtain the difference image.
- Step7: According to the setting threshold, segment targets from the background.
- Step8: The end.

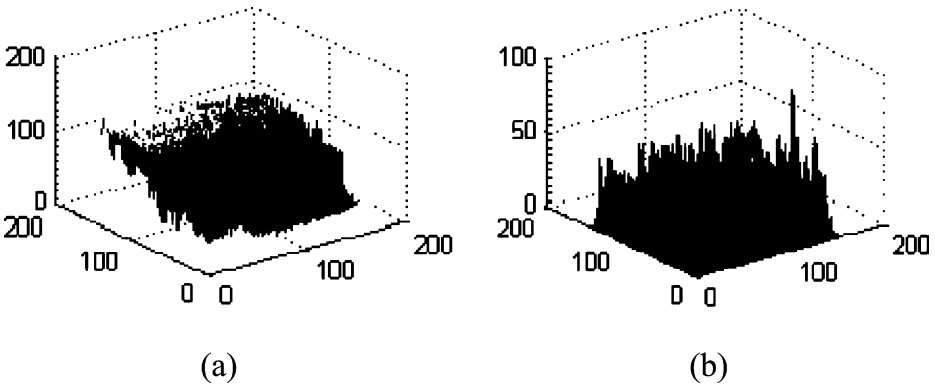


Fig. 7 3-D projection map. (a) Projection map before filter (SNR=1.4); (b) Projection map after filter (SNR=2.1).

5 Experimental results and analysis

The imaging sensor for this experiment is an infrared thermal sight that is cooling type and installed on a vehicle or flying. The environment is natural field. The targets embedded in the images contain vehicles, naval ships and fire balloons. The imaging sensor was located at a distance of about 500 m~2000 m from the target. The digitized image size is 228*280 pixels. The available dynamic image range of the displayed “black and white” images is 256 gray levels or 8 bits/pixel. The thermal condition of the target is warm (engine working) or cold (engine off).

880 (176×5) frame gray images are used for testing. Figure 6 shows the detection effects of five typical images under different clutter background in AADM and BF. In experiments, we set $L_{max}=9$ and the size of forecast window is 7×7 .

Segment threshold TH is selected by

$$TH = Avg + k * Std \tag{10}$$

where *Avg* is the image mean, *Std* is the image variance, *k* is a coefficient. Usually, *k* samples value from in 1~3. In the five class figures, we also use a white rectangular to mark the detected targets, respectively.

As shown in Fig. 6, the targets signal has been enhanced and the background signal has been reduced. Targets are detected accurately from the background and the probability of false alarm is lower. Figure 6 (c) show the processing results of our algorithm.

Table 1 Comparison of several target detection methods.

Detection Method	Median	5×5 Highpass template	Wavelet modulus	Morphological top-hat	AADM and BF
A1 (%)	96.1	96.6	97.1	96.2	98.1
A2 (%)	93.5	93.9	94.1	93.3	97.9
A3 (%)	93.3	93.0	92.5	93.1	97.2
A4 (%)	89.4	92.5	92.0	92.6	96.1
A5 (%)	91.1	90.6	92.9	92.1	95.8

A1, A2, A3, A4, A5: original infrared images shown in Fig. 6.

In order to reflect concretely the increase of SNR, Fig. 7 shows the 3-Dimension projection map of complex ground background in Fig. 6. As shown in Fig. 7, the target becomes outstanding in the background and the SNR is improved actually.

To compare our method with others, probability of detection Pd is used to evaluate the performance of the algorithms. It is defined as follows:

$$Pd = \frac{Nt}{Nd} \times 100\%$$

where Nt denotes the number of detection reports that correspond to a true target, Nd represents the number of database targets in the corresponding sequence.

In order to compare performance with our algorithm, four other target detection techniques were also implemented on the same images.

The experimental data are listed in Table 1, which shows that the detection performances of several algorithms are very close with each other when they are used for the simple background of infrared images. However, once the images are influenced by clutter, the detection algorithm AADM and BF is better than the others. Because the algorithm not only enhances the targets but also suppresses the background, its detection performance will be improved. It is obvious that the AADM and BF maintain better performance for small target detection under the different background.

6 Conclusions and future work

Automatic target detection based on infrared imaging system has the advantage of large dynamic range, long operating distance and all-day-around use, *etc.* however, because infrared band is susceptible to background scatter and atmospheric absorption, the structural information and texture of infrared targets are blur and the image contrast is very low. To improved ability of the small targets detection in single frame image and reduce the computation complexity. In this paper, we have presented an efficient algorithm for infrared small target detection from natural scenes, established the criterion of distinguishing the targets from natural scenes by AADM, and adopted BF to estimate the background. The results of theoretical and experimental with infrared images have shown the validity and efficiency of proposed method for small target detection.

To improve the Pd , future work should develop a transferable method for small targets detection based on a multi-features fusion model such as gray, motion, color, signature of discontinuity.

Acknowledgment This work has been supported by the National Defense Science Foundation of P.R. China (51401020201JW0521).

References

1. J. Y. Wang, and F. S. Chen. 3-D object recognition and shape estimation from image contours using B-splines, shape invariant matching and neural network. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(1), 13–23 (1994).
2. P. Mulassano, and L. Lo Presti. Object detection on the sea surface, based on texture analysis. *The 6th IEEE International Conference on Electronics, Circuits and Systems* **2**, 855–858 (1999).
3. A. Shashua. Projective structure from uncalibrated images: structure from motion and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**(8), 778–790 (1994).

4. B. S. Denney, and R. J. P. de Figueiredo. Optimal point target detection using adaptive auto regressive background predictive. *Signal Data Process. Small Targets*, Orlando, FL, USA. **4048**, 46–57 (2000).
5. V. T. Tom, *et al.* Morphology-based algorithm for point target detection in infrared backgrounds. *Signal Data Process. Small Targets*, Orlando, FL, USA. **1954**, 2–11 (1993).
6. J. X. Peng, and W. L. Zhou. Infrared background suppression for segmenting and detecting small target. *Acta Electron. Sin.* **27**(12), 47–51 (1999).
7. Z. J. Ye, *et al.* Detection algorithm of weak infrared point targets under complicated background of sea and sky. *J. Infrared Millim. Waves* **19**(2), 121–124 (2000).
8. C. I. Hilliard. Selection of a clutter rejection algorithm for real-time target detection from an airborne platform. *Signal Data Process. Small Targets*, Orlando, FL, USA **4048**, 74–84 (2000).
9. B. N. Ulisses, C. Manish, and G. John. Automatic target detection and tracking in forward-looking infrared image sequences using morphological connected operators. *J. Electron. Imaging* **13**(4), 802–813 (2004).
10. G. Wang, T. Zhang, L. Wei. Efficient method for multiscale small target detection from a natural scene. *Opt. Eng.* **35**(3), 761–768 (1996).