Infrared Small Target Detection via Adaptive M-Estimator Ring Top-Hat Transformation

Lizhen Deng[a], Jieke Zhang[b], Guoxia Xu[c], Hu Zhu[b,*]

[a] *National Engineering Research Center of Communication and Network Technology, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China.*
[b] *Jiangsu Province Key Lab on Image Processing and Image Communication, Nanjing University of Posts and Telecommunications, Nanjing 210003, China.*
[c] *Department of Computer Science, Norwegian University of Science and Technology, 2815 Gjovik, Norway.*

Abstract

Top-Hat transformation is an essential technology in the field of infrared small target detection. Many modified Top-Hat transformation methods have been proposed based on the different structure of structural elements. However, these methods are still hard to handle the dim targets and complex background. It can be summarized as two reasons, one is that the structural elements cannot suppress the background adaptively due to the fixed value of structural elements in image. Another is that simple structural element cannot utilize the local feature for target enhancement. To overcome these two limitations, a special ring Top-Hat transformation based on M-estimator and local entropy is proposed in this paper. First, an adaptive ring structural element based on M-estimator is used to suppress the complex background. Second, a novel local entropy is proposed to weight structural element for capturing local feature and target enhancement. Finally, a comparison experiment based on massive infrared image data (more than 500 infrared target images) is done. And the results demonstrate that the proposed algorithm acquires better performance compared with some recent methods.

Keywords: Infrared Small Target Detection, Top-Hat Transformation,

*Corresponding author
Email address: peter.hu.zhu@gmail.com (Hu Zhu)

M-estimator, Local Entropy

1. Introduction

The automatic target detection and tracking is one of most significant technique in remote warning [1], space surveillance [2], etc [3]. However, the task of detection is difficult due to the complex background and small shape of target [4]. In this case, the detection methods will cause false alarm in result [5]. As a consequence, small target detection under complex background still remains an open problem [6].

Over the past decades, many robust methods for infrared (IR) small target detection have been proposed, which highlight the works about background suppression, spatial-temporal information, adaptive contrast measure and etc [7, 8]. It is noticeable that the Top-Hat transformation has been a vital tool [9]. And numerous modified methods are proposed for Top-Hat transformation improvement [10, 11, 12]. In general, the processing before or after Top-Hat transformation is a common way to improve the SCRG (signal to clutter ratio) of originally image, such as energy accumulation[11, 10]. However, these method can not adaptively handle the heavy noises.

Another way is turning to redefine the Top-Hat transformation through structural elements construction. In [12], a method called CB Top-Hat transformation was proposed. This CB Top-Hat transformation weakens the impact of the structural element size on the results. However, the outline of the large structures still leads to noise amplification, and the smaller structure contour results in noise filtering ineffectively. Then, a toggle contrast operator [13] was further proposed for Top-Hat transformation to enhance the target under noise. Similarly, Bai et al. [14] combined hit-or-miss transformation with Top-Hat transformation to utilize contrast information between target and surrounding background. Besides, a new ring Top-Hat transformation was proposed to incorporate the contrast information into dilation and erosion operation through two ring shape structural elements [15]. However, as Guo et al. [16] discussed,

the result was sensitive to the value of structural elements.

In [15], a ring Top-Hat transformation was proposed to utilize the contrast information for erosion and dilation operation in Top-Hat transformation. However, there are two limitations existing in the traditional methods. The one is that the value of structural elements should be set before and can not be adaptive to the position in image. Structural elements do not suppress the background sufficiently. Another is that the simple structural elements can not capture the local feature to enhance the target in result.In order to handle these two limitations, we propose a ring Top-Hat transformation based on M-estimator and a novel local entropy. Our motivation can be summarized as follows: (1) It is noticeable that M-estimator is widely applied in numerous machine learning and computer vision fields, such as image registration [17], multidimensional scaling [12], tensor recovery [18] due to the robustness for data prediction. The key step of Top-Hat transformation is to obtain the background using opening operation, which also can be regarded as background recovery. The feature of background is definitely unique in the different position of image. Thus, we construct the structural elements through M-estimator. In this way, the structural elements can suppress the background adaptively according to the position of image. (2) Since a target always occupies several pixels in image, the target would rise up the change of local pixel intensity [19]. The entropy is effective to capture this local feature, which is also widely used in small target detection [20, 21]. Unlike the traditional methods, to capture the local feature surrounding target for target enhancement, a novel entropy is proposed to incorporate the structural elements.

The contributions of this paper are as follows: (1)An adaptive structural element based on M-estimator is proposed, which is able to suppress the image background adaptively. (2) A novel local entropy is proposed for applying Top-Hat transformation to capture the local feature for target enhancement.(3) Extensive experiments were done based on massive infrared images. The comparison results between our method and recent popular methods show that our method can acquire the better result.

The structure of paper can be summarized as follows: Some background knowledge is introduced in Section 2. Section 3 shows the proposed algorithm. And experimental results are displayed in Section 4. Section 5 is conclusion.

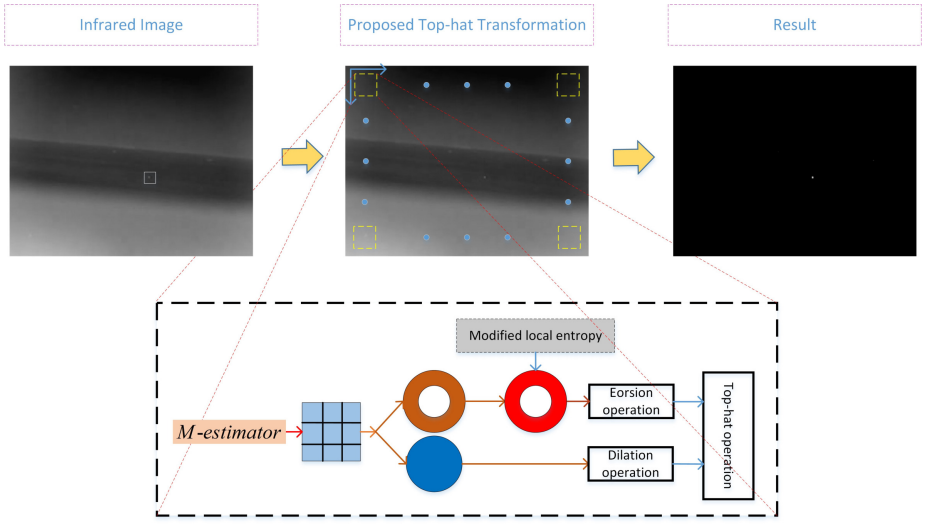

Figure 1: Our proposed algorithm. The main procedure consists of two parts: the ring Top-Hat transformation with local entropy and adaptive structural elements based on M-estimator.

2. Background

2.1. Top-hat Transformation

Morphology is a mathematical tool for analyzing the image, which is based on the structural elements. The basic idea is applying a certain structural element to execute the traversal operation and corresponding calculation in an image. There are four fundamental operations in the mathematical morphology: dilation, erosion, opening and closing. It should be noticed that the structuring element also can be regarded as an image collection.

Dilation and erosion of these two basic operations are first introduced. We denote the gray image I and structural element B as two sets. Then, the dilation

of $I(X,Y)$ by $B(U,V)$ are defined by follow formulations:

$$(I \oplus B)(X,Y) = \max_{U,V}(I(X-U, Y-V) + B(U,V)) \tag{1}$$

The maximum operation would make the gray value of output-image larger than that of input-image.

For gray image, the 'or' operation between the structuring element B and the pixel would be taken. As for erosion operation, the erosion of $I(X,Y)$ by $B(U,V)$ are defined as follows:

$$(I \ominus B)(X,Y) = \min_{U,V}(I(X+U, Y+V) - B(U,V)) \tag{2}$$

Correspondingly, minimum operation would make the gray value of output-image smaller than that of input images. And for a binary image, the 'and' operation would be taken.

The structure elements can be considered as a convolution template, structuring function. It determines the shape and size of window which are used to applied in the morphology operations. As we all know, dilation enhances bright areas while erosion diminishes dark areas.

Two other fundamental morphological operators, opening and closing are defined based on dilation and erosion:

$$(I \circ B)(X,Y) = (I \ominus B) \oplus B \tag{3}$$

$$(I \bullet B)(X,Y) = (I \oplus B) \ominus B \tag{4}$$

Opening is erosion followed by dilation, closing is dilation followed by erosion. And opening smooth bright small regions of image, whereas closing eliminates dark small holes.

Then white Top-Hat transformation and black Top-Hat transformation are defined by:

$$W - TH_{I,B}(X,Y) = I(X,Y) - (I \circ B)(X,Y) \tag{5}$$

$$B - TH_{F,B}(X,Y) = (I \bullet B)(X,Y) - I(X,Y) \tag{6}$$

$W - TH$ can find out bright regions in the image f, and CTH highlights the dark regions. Normally, the target would be remained because the small target region in the image is always a small bright or black region.

2.2. ring Top-Hat transformation

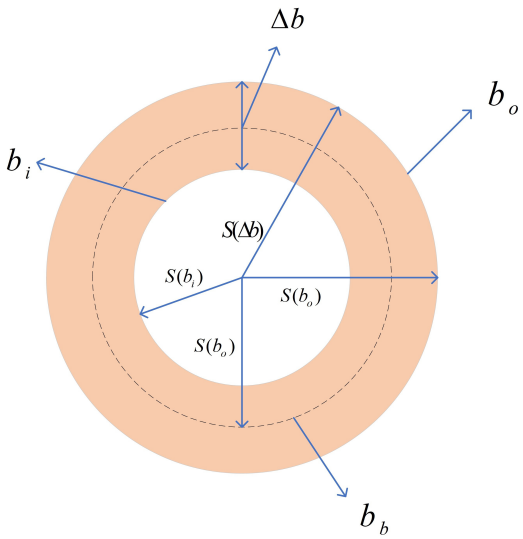

Figure 2: Relationship of the structuring elements

The classical Top-Hat transformation uses two same structuring elements. In this way, it inappropriately considers the differences between the target and surrounding regions.

In order to utilize the difference information between the target region and surrounding background, the method of two different ring shape structuring elements in the transformation is proposed. As [15] defined, b_o and b_i are two flat structuring elements with the same shape and different size. $S(b)$ represents the size of structuring elements b, which is the radial distance from center of b to the contour of b. Δb is the boundary region between b_i and b_o. b_b represents the structuring element whose size is between $S(b_i)$ and $S(b_o)$. The relationship of b_i, b_o, Δb and b_b is demonstrated in Fig. 2.

3. The Proposed Method

In this section, we propose our adaptive ring Top-Hat transformation based on M-estimator and local entropy. The whole algorithm is displayed in Fig. 1

3.1. The Adaptive Structural Element based on M-estimator

M-estimator is a robust estimation method commonly used in statistical neighborhoods, it is also widely used in image processing [22]. Such as image registration [17], multidimensional scaling [12], and tensor recovery [18]. Based on these observations and the robustness of M-estimator for data prediction, we use the M-estimator to obtain the value of structural element adaptively, this structural element would be combined with the local feature of image.

From the view in statistical point, the simplest estimation of the underlying image signal is to estimate the intensity value of each pixel by minimizing the sum of squared residual errors within local windows. One presentative algorithm is least-squares estimation:

$$G_i = \arg\min_{\theta} \sum_{j \in \Omega_i} (\theta - I_j)^2 \tag{7}$$

Here, Ω_i is the neighborhood of pixel i and θ is the pixel ranging in [0,255]. However, the problem of this least-squares estimation is that it is very sensitive to outliers. Thus, pixels on both sides of the edge will affect each other. As a consequence, edges will get blurred.

To increase robustness and reject outliers, the above quadratic function can be replaced with more tolerant function that gives less penalty to outliers. For instance, we can apply the absolute error function like this:

$$G_i = \arg\min_{\theta} \sum_{j \in \Omega_i} |\theta - I_j| \tag{8}$$

More generally, the above two formulations can be extended like this:

$$G_i = \arg\min_{\theta} \sum_{j \in \Omega_i} \rho(\theta - I_j) \tag{9}$$

Here, the $\rho()$ is a loss function. To increase robustness, the loss function should not grow too rapidly. Such as L_1 function($|x|$), Guassian function($\sigma^2(1-\exp(-\frac{x^2}{2\sigma^2}))$) and Lorentzian function($\log(1+\frac{1}{2}(\frac{x}{\sigma})^2)$). The σ is a scale parameter, it is used to control the influence scale. We can observe that the influence function of least-squares estimation is unbounded when $x \to \infty$. Here let θ denotes the set of all possible output pixel values in $[0, 255]$. Then define $D(\theta)$ as a cost function for a window of image in our morphology operations, each pixel p is calculated from input window W as follows:

$$[D(\theta)]_p = \rho(\theta - W_q) \tag{10}$$

Here, W_q is the pixels in local window. Then our M-estimator function can reformulate into structural elements:

$$B_\theta = \arg\min_\theta [D(\theta)]_p \tag{11}$$

The window W has the same size with our structuring elements, and each value of structuring elements can be computed by above formulation. The output for each pixel is approximated by the original pixel of the window in original image, which would keep the property of local area in original image. Here we choose the Gaussian function. Thus, we introduce the M-estimator function to construct the structuring elements directly. Then the Top-Hat transformation by using M-estimator function is defined as:

$$W - TH_{I,B}(X,Y) = I(X,Y) - I \circ B_\theta(X,Y) \tag{12}$$

$$B - TH_{I,B}(X,Y) = I \bullet B_\theta(X,Y) - I(X,Y) \tag{13}$$

3.2. Ring Top-Hat Transformation based on the Proposed Novel Local Entropy

After using the M-estimator, the structuring elements would be adaptive to different location in image by introducing local feature information. In this way, the ability of top-transformation would be improved at background suppression.

Since a target always occupies several pixels in an image, the definition of foreground and background lies in the change of local pixel intensity. And in the target area, the complex degree is different compared to other local areas. However, the ring Top-Hat transformation proposed in [15] can not consider this local feature. Consequently, target will not be enhanced sufficiently after Top-Hat transformation. Since the information entropy can represent the complex degree of gray value in the image. We proposed a novel modified local entropy here for target enhancement.

For a small block area W with a size of $R_1 \times R_2$, $F(X,Y)$ is defined as the gray value at the center pixel point (X,Y). If the block contains m various of gray values f_i, $i = 1, 2, ..., m$, our novel local entropy is formulated as:

$$E(x,y) = -\sum_{i=1}^{m} \left(F_i - \overline{F}\right)^2 P_i \log_2 P_i \tag{14}$$

Here, P_i is the probability density function of the ith gray level and $\overline{F} = \sum_{i=1}^{m} F_i/m$. The Δb has the difference information between target and surrounding regions. Thus, we calculate the local entropy S in the area of Δb, and the structuring elements of Δb which is weighted by E. Then the structural elements will include more difference information between target and surrounding regions. Therefore, our Top-Hat transformation can be redefined as follows:

$$W - TH(X,Y) = I(X,Y) - (I \ominus \Delta b_\theta \cdot E) \oplus b_b \tag{15}$$

$$B - TH(X,Y) = (I \oplus \Delta b_\theta \cdot E) \ominus b_b - I(X,Y) \tag{16}$$

The size of structural element is usually same as the size of target. If the window is located at the target, there is the least difference information between target and background. Thus, our local entropy in the window is close to 1 (the maximum value). Let (X_o, Y_o) be center pixel of a target. Then:

$$E(X_o, Y_o) \approx 1 \geq E(X_i, Y_j) \tag{17}$$

Here, (i,j) is the index of pixel in other positions. The intensity of the erosion

operation in the position of target would almost not change.

$$(I \ominus \Delta b_\theta \cdot E)(X_o, Y_o) \approx (I \ominus \Delta b_\theta)(X_o, Y_o) \tag{18}$$

If the center of window is locating at the edge of the target or background, whose surrounding exists different degrees of clutters and noise. Then the gray value in this block is uneven and the modified entropy in this block would be close to zero.

$$E(X_i, Y_j) \ll 1 \tag{19}$$

Correspondingly, the intensity of the erosion operation in the position of target will very small. And after erosion operation we will get:

$$(E \ominus \Delta b_\theta \cdot E)(X_i, Y_j) \geq (I \ominus \Delta b_\theta)(X_i, Y_j) \tag{20}$$

Then after the dilation operation,

$$(I \ominus \Delta b_\theta \cdot E) \oplus b_b(X_i, Y_j) \geq (I \ominus \Delta b_\theta) \oplus b_b(X_i, Y_j) \tag{21}$$

and in the next Top-Hat operation, the background with noise and clutter will be reduced more:

$$(I - (I \ominus \Delta b_\theta \cdot E) \oplus b_b)(X_i, Y_j) < (I - (I \ominus \Delta b_\theta) \oplus b_b)(X_i, Y_j) \tag{22}$$

3.3. Our Proposed Small Target Detection Method

The target and background will be enhanced and suppressed respectively after the processing of our method. As a result, the most salient point of image is likely to be the target. Then we can detect it by the following threshold:

$$T = \overline{I}_{in} + L\sigma_{I_{in}} \tag{23}$$

Among them, the setting of the value of L refers to the literature [20, 21]. For the output image, $\overline{I}_{in}$ is the mean and $\sigma_{I_{in}}$ is the standard deviation. Normally, the L is range from 5 to 6.

Algorithm 1 Overall algorithm for target detection

The size of image is $M \times N$
  for $x = 1$ to M do
    for $y = 1$ to N do
      Calculate the local entropy according to Eq. (3.2) .
      Calculate ring structural elements according to Eq. (3.1)
      Use b_b and $E \cdot \Delta b$ to do dilate and erosion operations.
    end
  end
Use 3.3 to detect target.

4. Experimental Results

4.1. Experiment Setting

In this section, the recent methods, evaluation metrics and infrared image data are used for experimental comparison. The experiments are conducted on a laptop with 8-GB of random access memory and an Intel(R) Core(TM) i5-7440HQ CPU with 2.80GHz processor, and the implementation is presented in MATLAB 2016a.

Furthermore, to measure the saliency of target and background, the signal-to-clutter ratio gain (SCRG) and the background suppression factor (BSF) these two metrics are used. They are defined as follows:

$$SCRG = (S^{out}/N^{out})/(S^{in}/N^{in}) \tag{24}$$

$$BSF = C^{in}/C^{out} \tag{25}$$

S and C represent average value of the pixels in image and the clutter standard deviation respectively. Then C^{in} and C^{out} represent the standard deviation of input image and output image respectively.

Moreover, the probability of detection P^d and the rate of false alarm P^f are introduced to measure the performance of detection. Their formulations are listed as follows:

$$P^d = N_T/N \times 100\% \tag{26}$$

$$P^f = N_F/N \times 100\% \tag{27}$$

Here, the N_T, N_F and N denote the number of correctly detect targets, the number of incorrectly targets and the number of totally real target. Here some recent method about local descriptors and mathematical morphology are introduced for comparison. Including ELDM [20], LCM [23], NWTH [15], MPCM [24], DSBM [25]. Then six image sequences with different target and background are used. In Fig. 3, six representative images of the sequence are displayed, and the detail of them are listed in Table 1.

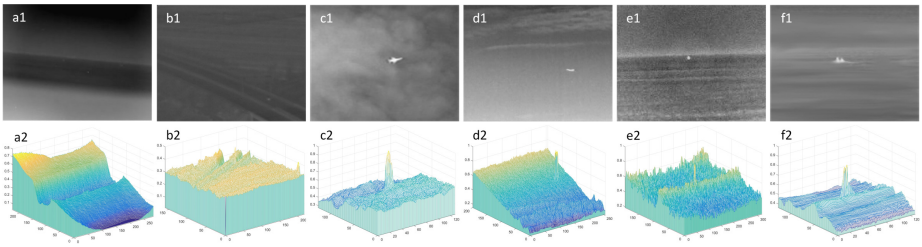

Figure 3: (a1)-(f1): Representative images of six real infrared images sequences. (a2)-(f2): Three-dimensional gray distribution of (a1)-(f1).

Table 1: The details of real image sequences

	Image size	Target size	Target type	Target details	Background details
sequence1	286×225	3×3	Small ship	Small and moveless	Sea-sky background with little change
sequence2	220×160	5×7	Car	A large size and keep moving	Highway background
sequence3	128×128	16×10	Aircraft	A large size and keep moving	Sky background with heavy cloudy. Keep changing.
sequence4	256×200	15×6	Airplane	A change size in a big range	Sea-sky background with cloudy.
sequence5	320×196	6×5	Small ship	Round size and keep moving	Sea-level background with much clutter
sequence6	128×128	8×3	Ship	Keep moving with the change of size	Sea background with noise

4.2. Qualitatively Analysis

In Fig. 4, each row shows the enhanced 3D gray results from different methods, including ELDM, LCM, NWTH, MPCM, DSBM and our method. The detailed features of these algorithms will be explained one by one below:

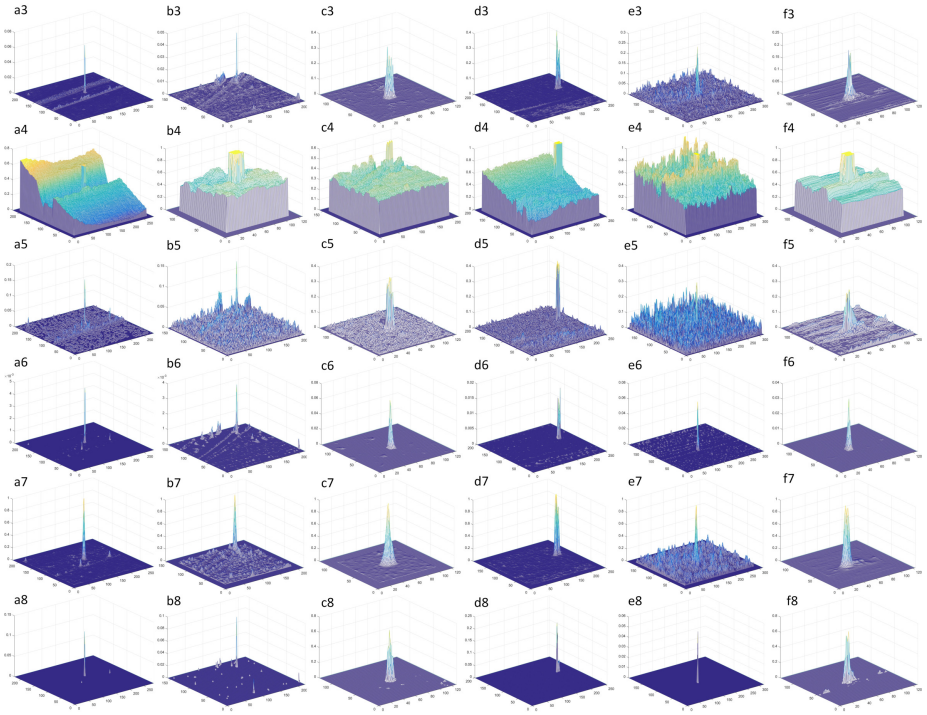

Figure 4: The contrast between our method and other methods. (a3)-(f3),(a4)-(f4), (a5)-(f5), (a6)-(f6), (a7)-(f7), and (a8)-(f8) are the enhanced result obtained through the TGRS, LCM, NWTH, MPCM, DSBM and our methods, respectively.

- Entropy-based window selection for detecting dim and small infrared targets (ELDM)[20]. This method adopt an adaptive entropy-based window selection technique to construct a novel local difference measure map of an input image to detect the targets.

- A local contrast method for small infrared target detection(LCM)[23]. This method used the proposed local contrast measure which measures the dissimilarity between the current location and its neighborhoods.

- Analysis of new top-hat transformation and the application for infrared dim small target detection(NWTH)[15]. This method used two different but correlated structuring elements to reorganize the classical top-hat transformation.

- Multiscale patch-based contrast measure for small infrared target detection (MPCM)[24]. This method increased the contrast between target and background, which makes it easy to segment small target by simple adaptive thresholding method.

- A robust directional saliency-based method for infrared small-target detection under various complex backgrounds(DSBM)[25]. This method formulated this detection as salient region detection, which is inspired by the fact that a small target can often attract attention of human eyes in infrared images.

It is clear that our method can almost get better result than other methods since the background clutter are less in the results from our method. The other algorithms also work well when the background is flat, but once the background cloud or fog is severe, false alarms are prone to occur. For example, in the second set of infrared image sequences, algorithms other than LCM can perform well, at least the target can be found. NWTH and MPCM will produce some noise, which affects the detection result to a certain extent. In the fifth set of sequences, the differences between all algorithms can be clearly seen. In the sea level background with a lot of clouds, our algorithm can accurately locate the target, and the smoothing effect on the background area is also excellent. Our method exploits the novel entropy method with weight structural elements which can dramatically improve the local feature and target enhancement.

4.3. Quantitatively Analysis

Furthermore, to demonstrate the performance of our algorithm objectively, the SCRG and BSF are measured for five random images in each sequence. The higher the SCRG and BSF are obtained, the better performance of the algorithm is. The result of each sequence is listed from Table 2 to Table 7. From the result, our method also almost get the higher SCRG and BSF than other methods. Fig. 5 shows the receiver operating characteristic (ROC) curves of six real image sequences. Our method achieves better performance than other

methods, especially for sequence 2, 4, and 6. The proposed method achieves the higher detection probability and lower false alarm.

Table 2: The indexes acquired from different methods (Image Sequence 1)

	1		2		3		4		5	
	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
ELDM	61.45	337.32	59.19	399.01	57.96	335.15	50.61	347.78	55.46	359.10
LCM	1.53	0.95	1.59	0.97	1.55	1.96	1.59	0.93	1.25	1.91
NWTH	29.86	67.95	34.86	73.15	26.15	60.65	24.43	62.51	21.05	62.16
MPCM	88.68	71.73	87.23	71.72	87.27	83.39	83.59	74.21	68.05	112.21
DSBM	18.60	11.84	18.86	11.46	17.57	11.08	16.91	11.89	11.01	11.15
Our Method	**97.21**	**382.27**	**109.7**	287.8	**91.38**	138.19	**99.76**	292.87	**103.6**	341.1

Table 3: The indexes acquired from different methods (Image Sequence 2)

	1		2		3		4		5	
	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
ELDM	6.25	448.73	4.12	405.46	9.34	394.40	4.25	412.55	7.81	431.89
LCM	0.89	0.84	0.88	0.84	0.87	0.84	0.90	0.84	0.89	0.84
NWTH	2.27	21.28	2.22	19.12	2.97	18.85	2.26	20.70	2.75	21.38
MPCM	17.22	540.10	16.53	581.8	17.53	447.5	16.53	461.1	19.05	472.70
DSBM	3.17	2.16	3.12	1.82	7.16	3.37	4.24	4.02	2.33	2.59
Our Method	**23.87**	152.81	13.59	542.26	**41.45**	391.69	**27.84**	**500.56**	**35.56**	**753.02**

Table 4: The indexes acquired from different methods (Image Sequence 3)

	1		2		3		4		5	
	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
ELDM	4.26	41.92	2.66	44.52	3.06	43.10	3.31	50.25	3.06	50.18
LCM	0.29	0.13	0.23	0.13	0.26	0.13	0.20	0.13	0.21	0.13
NWTH	1.11	3.33	0.97	3.40	0.99	3.29	0.96	3.46	0.97	3.42
MPCM	4.41	562.52	4.23	613.07	4.21	585.45	4.53	634.41	4.99	615.65
DSBM	3.40	1.65	3.27	1.09	3.34	0.93	3.61	0.99	3.42	1.04
Our Method	**9.15**	45.76	**7.17**	50.19	**7.33**	45.38	**6.78**	47.03	**9.40**	70.89

4.4. The Contribution of the Adaptive Structural Elements and the Proposed Local Entropy

Our adaptive structural elements based on M-estimator makes our ring Top-Hat transformation suppress the complex background and the proposed local

Table 5: The indexes acquired from different methods (Image Sequence 4)

	1		2		3		4		5	
	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
ELDM	2.15	7.65	1.69	5.10	1.16	4.29	1.35	3.89	2.44	5.93
LCM	0.52	0.26	0.30	0.17	0.25	0.13	0.26	0.13	0.54	0.29
NWTH	2.20	5.94	1.45	3.66	0.96	3.28	1.13	2.62	2.36	4.91
MPCM	2.18	57.46	1.40	36.33	2.16	34.42	1.61	27.14	1.13	61.51
DSBM	2.31	2.01	1.98	1.43	1.24	1.07	1.35	1.00	3.22	2.30
Our Method	**3.60**	3.20	**2.45**	2.21	2.12	1.43	**2.07**	1.55	**4.36**	3.52

Table 6: The indexes acquired from different methods (Image Sequence 5)

	1		2		3		4		5	
	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
ELDM	7.72	20.57	7.19	31.68	8.04	33.71	7.61	28.50	9.91	29.45
LCM	1.18	0.62	0.82	0.55	0.80	0.51	0.53	0.51	0.64	0.58
NWTH	5.72	12.68	4.87	13.90	5.43	14.63	5.56	15.11	7.12	17.57
MPCM	3.66	491.57	4.86	846.22	4.19	813.96	5.22	594.12	1.25	732.51
DSBM	4.63	5.15	10.72	5.34	12.01	6.40	3.38	5.58	11.60	5.77
Our Method	**9.59**	32.96	9.94	61.13	**19.41**	168.78	**10.60**	97.20	**13.34**	85.14

Table 7: The indexes acquired from different methods (Image Sequence 6)

	1		2		3		4		5	
	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
ELDM	1.22	7.94	1.55	6.07	2.03	6.24	2.09	6.06	1.60	6.78
LCM	0.26	0.17	0.28	0.19	0.27	0.19	0.30	0.19	0.27	0.19
NWTH	1.19	3.83	1.08	3.43	1.14	2.60	1.08	2.66	1.14	2.71
MPCM	2.47	6.07	1.83	5.06	1.36	0.50	1.80	1.86	1.59	2.18
DSBM	2.05	1.04	1.72	1.33	1.28	1.24	1.72	1.28	1.88	1.15
Our Method	1.93	2.42	**1.94**	1.50	1.51	1.61	1.74	1.50	**1.97**	1.60

entropy makes our algorithm robust for target enhancement. Here we compare the result between ring Top-Hat transformation, adaptive ring Top-Hat based on M-estimator based and adaptive ring Top-Hat transformation based on M-estimator and local entropy.

We select three representative image sequence (named A, B and C) with complex background and clutters. The size of target all ranges from 5 to 40

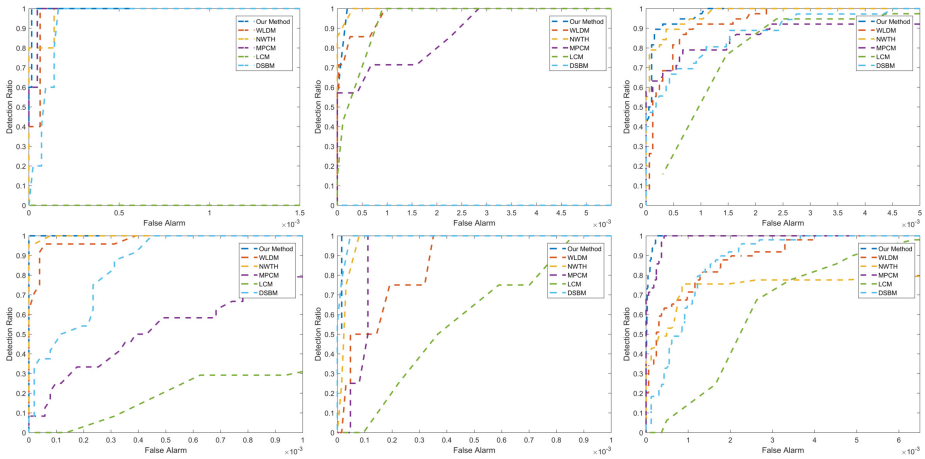

Figure 5: ROC curves of six image sequences. The horizontal axis shows the false alarm and vertical axis shows the detection probability.

Table 8: The contribution of different parts in our method

	Sequence A SCRG/BSF	Sequence B SCRG/BSF	Sequence C SCRG/BSF	Average Runing Time
Ring Top-hat	2.19/2.63	2.16/1.95	1.26/1.08	0.91s
Introducing M-estimator	6.17/5.13	4.12/3.65	2.15/1.56	1.45s
Introducing local entropy	12.89/9.87	5.17/4.63	4.40/3.98	1.89s

pixels. As is shown in Fig. 6, First and second column are the representative three input image and corresponded 3D gray distribution image. It is obvious that there are heavy clutters in background area. However, the ring Top-Hat transformation can not suppress the background thoroughly due to the simple structural elements. After using our adaptive structural elements, the ring Top-Hat transformation can suppress more background clutters. Finally, after introducing the proposed local entropy, the background is almost suppressed and target is enhanced sufficiently. Moreover, the testing average values of the SCRG and BSF are listed in the Table 8, which shows the contribution of these two parts in our algorithm.

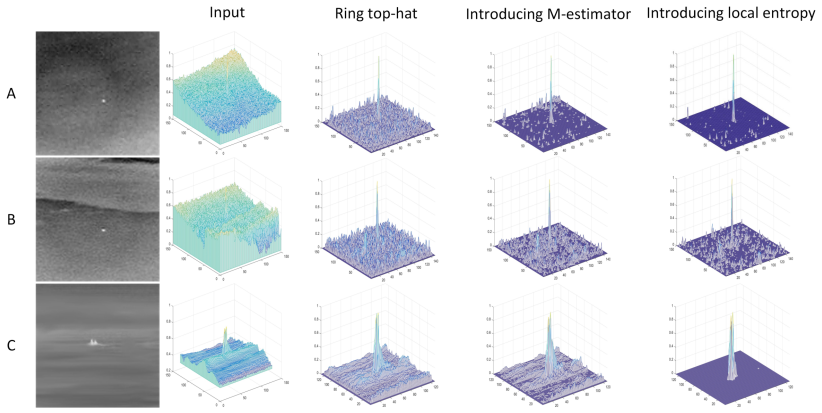

Figure 6: The contribution of different part in our method. First and second column show the input image and corresponded 3D-distribution. The third column shows the result obtained by ring Top-Hat transformation. The fourth column shows the result obtained by adaptive ring Top-Hat transformation with M-estimator. The fifth column shows the result obtained by adaptive ring Top-Hat transformation with M-estimator and local entropy.

4.5. Complexity Analysis

Here, we briefly discuss the complexity of our method. As shown in Fig. 1 , it is clear that the computation of our algorithm mainly consists of two parts:1) the construction of structural element 2) the entropy operation.

We define that the size of image is $M \times N$. For the local entropy operation, the computational complexity is $O(M \times N \times R^2)$. For structural element construction, the complexity is also $O(M \times N \times R^2)$. Then for dilation and erosion operation, the total computational complexity is $O(M^2 \times N^2 \times R^2)$. The algorithm complexity and computational time for Fig. 3 with different methods are shown in Table 9.

Table 9: The Computational Complexity and Time Consume Through Different Methods

	ELDM	LCM	NWTH	MPCM	DSBM	Our Method
Complexity	$O\left(R^3MN^2\right)$	$O(M \times N \times R^2)$	$O(M^2 \times N^2)$	$O(l \times M \times N \times R^2)$	$O\left(MN\right)$	$O(M \times N \times R^2)$
Time/s	2.24	0.77	0.89	0.65	0.24	1.25

In this way, based on the above analysis, the computational complexity is $O(M \times N \times R^2 + 2 \times M^2 \times N^2 \times R^2)$.

4.6. Conclusion

In this paper, an adaptive ring Top-Hat transformation is proposed based on M-estimator and a novel local entropy for small infrared target detection. In this algorithm, the construction of structural element is adaptively introduced the method of M-estimator, which can contain the local feature information of image. And the background can be suppressed efficiently by this structural element. Additionally, to utilize the local information for target enhancement, two different shapes are used for structuring elements and a modified entropy is proposed to weight the Top-Hat operation at the location of target. Some recent novel methods and massive infrared image data are used for comparison in the experiment. The extensive results demonstrate that proposed algorithm not only obtain the higher SCRG, BSF, but also can achieve higher detection probability and lower false alarm. However, there are two limitations. Firstly, our method could not handle multiple target size in a large range. This will prospect in future study. Secondly, the proposed method uses multiple iterations in local operation. Accordingly, a faster version of current algorithm should be investigated in future work.

Acknowledgements

This work is supported by the National Natural Science Foundation of China under Grant 61701259.

References

References

[1] S. Kim, J. Lee, Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track,

Pattern Recognition 45 (1) (2012) 393 – 406. doi:10.1016/j.patcog.2011.06.009.

[2] D. Wang, H. Lu, Z. Xiao, M. Yang, Inverse sparse tracker with a locally weighted distance metric, IEEE Transactions on Image Processing 24 (9) (2015) 2646–2657. doi:10.1109/TIP.2015.2427518.

[3] H. Zhu, S. Liu, L. Deng, Y. Li, F. Xiao, Infrared small target detection via low-rank tensor completion with top-hat regularization, IEEE Transactions on Geoscience and Remote Sensing (2019) 1–13doi:10.1109/TGRS.2019.2942384.

[4] X. Wang, S. Shen, C. Ning, M. Xu, X. Yan, A sparse representation-based method for infrared dim target detection under sea–sky background, Infrared Physics & Technology 71 (2015) 347 – 355. doi:10.1016/j.infrared.2015.05.014.

[5] Y. Deng, H. Duan, Biological eagle-eye-based visual platform for target detection, IEEE Transactions on Aerospace and Electronic Systems 54 (6) (2018) 3125–3136. doi:10.1109/TAES.2018.2845178.

[6] Y. Li, Y. Zhang, Robust infrared small target detection using local steering kernel reconstruction, Pattern Recognition 77 (2018) 113 – 125. doi:10.1016/j.patcog.2017.12.012.

[7] L. Zhang, J. Zhang, A new saliency-driven fusion method based on complex wavelet transform for remote sensing images, IEEE Geoscience and Remote Sensing Letters 14 (12) (2017) 2433–2437. doi:10.1109/LGRS.2017.2768070.

[8] C. Gao, L. Wang, Y. Xiao, Q. Zhao, D. Meng, Infrared small-dim target detection based on markov random field guided noise modeling, Pattern Recognition 76 (2018) 463 – 475. doi:10.1016/j.patcog.2017.11.016.

[9] X. Bai, F. Zhou, Analysis of different modified top-hat transformations based on structuring element construction, Signal Processing 90 (11) (2010) 2999 – 3003. doi:10.1016/j.sigpro.2010.04.021.

[10] T. Chen, Q. Wu, R. Rahmani-Torkaman, J. Hughes, A pseudo top-hat mathematical morphological approach to edge detection in dark regions, Pattern Recognition 35 (1) (2002) 199 – 210, shape representation and similarity for image databases. doi:10.1016/S0031-3203(01)00024-3.

[11] F. Zhang, C. Li, L. Shi, Detecting and tracking dim moving point target in ir image sequence, Infrared Physics & Technology 46 (4) (2005) 323 – 328. doi:10.1016/j.infrared.2004.06.001.

[12] R. Lerallut, Étienne Decencière, F. Meyer, Image filtering using morphological amoebas, Image and Vision Computing 25 (4) (2007) 395 – 404, international Symposium on Mathematical Morphology 2005. doi:10.1016/j.imavis.2006.04.018.

[13] X. Bai, F. Zhou, B. Xue, Infrared dim small target enhancement using toggle contrast operator, Infrared Physics & Technology 55 (2) (2012) 177 – 182. doi:10.1016/j.infrared.2011.12.002.

[14] X. Bai, F. Zhou, Hit-or-miss transform based infrared dim small target enhancement, Optics & Laser Technology 43 (7) (2011) 1084 – 1090. doi:10.1016/j.optlastec.2011.02.003.

[15] X. Bai, F. Zhou, Analysis of new top-hat transformation and the application for infrared dim small target detection, Pattern Recognition 43 (6) (2010) 2145 – 2156. doi:10.1016/j.patcog.2009.12.023.

[16] J. Guo, G. Chen, Analysis of selection of structural element in mathematical morphology with application to infrared point target detection, in: Infrared Materials, Devices, and Applications, Vol. 6835, International Society for Optics and Photonics, 2008, p. 68350P. doi:10.1117/12.754627.

[17] F. Mandanas, C. Kotropoulos, M-estimators for robust multidimensional scaling employing l2,1 norm regularization, Pattern Recognition 73 (2018) 235 – 246. doi:10.1016/j.patcog.2017.08.023.

[18] Y. Yang, Y. Feng, J. A. K. Suykens, Robust low-rank tensor recovery with regularized redescending m-estimator, IEEE Transactions on Neural Networks and Learning Systems 27 (9) (2016) 1933–1946. doi:10.1109/TNNLS.2015.2465178.

[19] L. Fang, W. Zhao, X. Li, X. Wang, A convex active contour model driven by local entropy energy with applications to infrared ship target segmentation, Optics & Laser Technology 96 (2017) 166 – 175. doi:https://doi.org/10.1016/j.optlastec.2017.05.008.

[20] H. Deng, X. Sun, M. Liu, C. Ye, X. Zhou, Entropy-based window selection for detecting dim and small infrared targets, Pattern Recognition 61 (2017) 66 – 77. doi:10.1016/j.patcog.2016.07.036.

[21] H. Deng, X. Sun, M. Liu, C. Ye, X. Zhou, Small infrared target detection based on weighted local difference measure, IEEE Transactions on Geoscience and Remote Sensing 54 (7) (2016) 4204–4214. doi:10.1109/TGRS.2016.2538295.

[22] L. Jiang, Y. Han, B. Xie, Edge-preserving image smoothing via a total variation regularizer and a nonconvex regularizer, Procedia Computer Science 154 (2019) 603 – 609, proceedings of the 9th International Conference of Information and Communication Technology [ICICT-2019] Nanning, Guangxi, China January 11-13, 2019. doi:10.1016/j.procs.2019.06.095.

[23] C. L. P. Chen, H. Li, Y. Wei, T. Xia, Y. Y. Tang, A local contrast method for small infrared target detection, IEEE Transactions on Geoscience and Remote Sensing 52 (1) (2014) 574–581. doi:10.1109/TGRS.2013.2242477.

[24] Y. Wei, X. You, H. Li, Multiscale patch-based contrast measure for small infrared target detection, Pattern Recognition 58 (2016) 216 – 226. doi:10.1016/j.patcog.2016.04.002.

[25] S. Qi, J. Ma, C. Tao, C. Yang, J. Tian, A robust directional saliency-based method for infrared small-target detection under various complex backgrounds, IEEE Geoscience and Remote Sensing Letters 10 (3) (2013) 495–499. doi:10.1109/LGRS.2012.2211094.