A Small Infrared Target Detection Method Based on Coarse-to-Fine Segmentation and Confidence Analysis

Hang Li[1(✉)], Ze Wang[1], Tian Tian[2], and Hui Huang[1]

[1] Research and Development Center, China Academy of Launch Vehicle Technology, Beijing 100076, People's Republic of China
Hanglee@hust.edu.cn
[2] China University of Geosciences, Wuhan 430074, People's Republic of China

Abstract. In this paper, a small target detection algorithm in infrared image is proposed. First, an infrared image is coarse-to-fine segmented automatically by self-adaptive histogram segmentation. After detecting small abnormal region in segmented image and defining them to candidate targets, the abnormality based confidence of each candidate target is calculated and sorted. Finally, the candidate target with the maximum confidence is weighed to be real one. The experiments demonstrate that the proposed method is efficient and robust.

Keywords: Infrared small target detection · Coarse-to-fine segment · Confidence

1 Introduction

Small infrared target detection is a key technology in infrared image process and pattern recognition. It is a hard problem for the variability of the target's appearance due to different atmospheric conditions, background, camera ego-motion and et al. [1]. By assuming hot targets appears as bright spots in the scene, variable methods are proposed to solve this problem. Gu et al. [3] proposed a kernel-based nonparametric regression method for background prediction and clutter removal, furthermore applied in target detection. Wang et al. [4] presented a highpass filter based on LS-SVM to detect small target. Wang et al. [5] provided a real time small target detection method based on cubic facet model. In general, traditional small target detection can be divided into two classes. Gao et al. [6] applied self-similarity descriptor to model and detect targets. One uses grayscales based threshold to segment the infrared image and detected small target [2-4], the other fuses multiple features to segment the target region in fusion image [5-8].

Both of these methods rely on the 'abnormality' of the target region. The 'abnormality' means that, compared with the neighboring background regions, the target regions has lower or higher intensity and its intensity distribution is more uniform. In the segmentation stage, the present methods always apply a single segmenting threshold or adaptive threshold in fixed window. However, considering the interrupts derived from some factors, such as noise and sea clusters, it is hard to identify the segment boundary

© Springer-Verlag Berlin Heidelberg 2015
T. Tan et al. (Eds.): IGTA 2015, CCIS 525, pp. 423–432, 2015.
DOI: 10.1007/978-3-662-47791-5_47

by using one-time segmentation. Thus the segmenting methods mentioned above cannot effectively detect the small target under complex background.

To solve this problem, a small target detection approach is proposed here by using the principle of maximum entropy (POME) based self-adaptive segmentation and confidence analysis. As shown in Fig.1, the approach has three stages. First, a self-adaptive segmentation is applied to coarse-to-fine segments the infrared image. Then, the candidate targets are searched and detected in the segmented image. Finally, by calculating and sorting the confidence of each candidate target, the real target is weighed to be the one with the highest confidence. The experimental results validate the good performance of the proposed method in infrared small target detection.

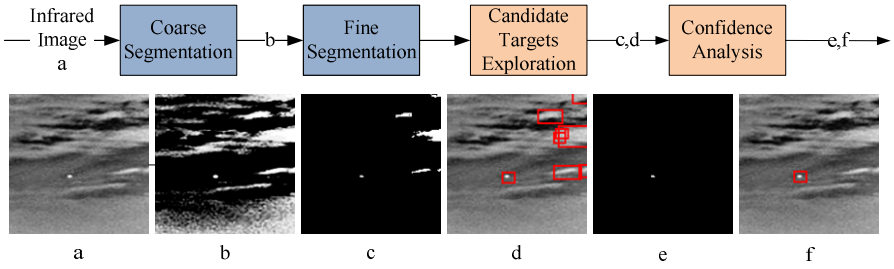

Fig. 1. The procedure of proposed method. a is an infrared image. It is coarsely segmented to b and then b is finely segmented to c. Candidate targets are searched in c and performed in infrared image d. By calculating the confidences of candidate targets, the one with maximum confidence is weighed to be the real target and shown in e and f.

2 POME Based Self-adaptive Segmentation

2.1 The Principle of Maximum Entropy (POME)

Entropy defines the uncertainty of the random variables. When the entropy is larger, the random variables are more unstable. In this sense, the principle of maximum entropy presents a fact that, if part of prior knowledge is known, the most reasonable inference about the unknown distribution is the most uncertain one according to the known knowledge. It's the only choice because the other choices would be added other constraints and assumptions, and at the same time these constraints and assumptions according to the information which cannot be launched. The objective function based on POME is

$$J = \max_{p \in P} H(Y/X) = \sum_{(x,y)} p(x,y) \log \frac{1}{p(y/x)} \tag{1}$$

Where $H(Y|X)$ is denoted as conditional entropy. $p(x,y)$ is joint probability density and $p(y|x)$ is conditional probability density.

2.2 Objective Function Based on POME

Assuming that the infrared image I has n grayscales $X=\{x_i\}_{i=0}^{n-1}$, where $x_i \in \mathbb{Z}$ and $x_i \in [0,255]$. The probability of grayscale x_i is defined as $p(x_i)$. Thus the entropy of infrared image I is

$$H(X)=-\sum_{i=0}^{n-1} p(x_i)\log p(x_i), \sum p(x_i)=1 \tag{2}$$

Suppose that there is t which defines ROI regions as $\mathbb{Z}_{ROI}=\{x_0,x_1,...,x_i=t-1\}$ and defines the background as $\mathbb{Z}_{background}=\{x_i=t,x_{i+1},...,x_{n-1}\}$. The entropies of ROI regions and background are denoted respectively as

$$\begin{cases} H_{ROI}=-\sum_{x_i=t}^{255} p(x_i)\log p(x_i) \\ H_{background}=-\sum_{x_i=0}^{t-1} p(x_i)\log p(x_i) \end{cases} \tag{3}$$

where $H(X)=H_{ROI}+H_{background}$. By using threshold $\tilde{t}$, I is segmented to be $\tilde{I}$, whose entropy is $\tilde{H}(\tilde{X})$ where $\tilde{X}=\{\tilde{x}_0=0,\tilde{x}_1=255\}$. Thus there is

$$\tilde{H}(\tilde{X})=\tilde{H}_{ROI}+\tilde{H}_{background} \tag{4}$$

$$\begin{cases} \tilde{H}_{ROI}=-\left(\sum_{x_i=\tilde{t}}^{255} p(x_i)\right)\log\sum_{x_i=\tilde{t}}^{255} p(x_i) \\ \tilde{H}_{background}=-\left(\sum_{x_i=0}^{\tilde{t}-1} p(x_i)\right)\log\sum_{x_i=0}^{\tilde{t}-1} p(x_i) \end{cases} \tag{5}$$

Base on Eq. (1), the objective functions are

$$J_1=\max\left\{H_{ROI}=H\left(\{x_i\}_{i=0}^{n-1}/\{x_i\geq\tilde{t}\}\subseteq\mathbb{Z}_{ROI}\right)\right\} \tag{6}$$

$$J_2=\max\left\{H_{background}=H\left(\{x_i\}_{i=0}^{n-1}/\{x_i<\tilde{t}\}\subseteq\mathbb{Z}_{background}\right)\right\} \tag{7}$$

$$p\left(x_i/\{x_i<\tilde{t}\}\subseteq\mathbb{Z}_{background}\right)=\begin{cases}\dfrac{p(x_i)}{\sum_{x_i}p(x_i)} & x_i<\tilde{t}\\ 0 & x_i\geq\tilde{t}\end{cases} \tag{8}$$

$$p\left(x_i/\left\{x_i \geq \tilde{t}\right\} \subseteq \mathbb{Z}_{ROI}\right) = \begin{cases} \dfrac{p\left(x_i\right)}{\sum_{x_i} p\left(x_i\right)} & x_i \geq \tilde{t} \\ 0 & x_i < \tilde{t} \end{cases} \tag{9}$$

Thus according to Eq. (8) and Eq. (9), the objective function can be simplified by

$$J_1 = \max\left\{H_{ROI} - \tilde{H}_{ROI}\right\} \tag{10}$$

$$J_2 = \max\left\{H_{background} - \tilde{H}_{background}\right\} \tag{11}$$

In traditional segment method based on POME, Eq. (10) and Eq. (11) always become

$$J = \max\left\{H_{ROI} - \tilde{H}_{ROI} + H_{background} - \tilde{H}_{background}\right\} = \max\left\{H\left(x_i\right) - \tilde{H}\left(x_i\right)\right\} \tag{12}$$

And this kind of method iterates each grayscales to find $\tilde{t}$ to satisfy Eq. (12). However the traversal algorithm wastes time and usually mistakenly segments part of target regions to background. To ensure segmenting the target regions completely, coarse-to-fine segmentation is used to get rid of most background regions first and then to obtain the ROI regions which are composed by target regions and false alarms.

$$J_1 = \max\left\{H\left(X\right) = H_{ROI} + H_{background}\right\} \tag{13}$$

$$J_2 = \left\{ find\ \tilde{t} \,\middle|\, \min\left\{H\left(\tilde{X}\right) = \tilde{H}_{ROI} + \tilde{H}_{background}\right\} and\ \tilde{H}_{ROI} > 0 \right\} \tag{14}$$

Because the uniform distribution has maximum entropy, Eq. (13) is transformed as

$$J_1 : f\left(x_i, x_i \in X\right) \rightarrow \left\{ y_j : p\left(y_j, Y = \left\{y_j\right\}_{j=0}^{n'-1}\right) = \frac{1}{n'} \right\} \tag{15}$$

Where $Y = \left\{y_j\right\}_{j=0}^{n'-1}$ follows uniform distribution and $Y \in \mathbb{Z}^m$, $y_j \in [0, 255]$. Due to the entropy function is convex, its minimum value is equal to 0 under the situation that $\tilde{t} = 0$ or $\tilde{t} = 255$. Thus the Eq. (14) is transformed as

$$J_2 : \tilde{t} = y_{n'-2} \tag{16}$$

2.3 The Coarse-to-fine Segmentation

Considering that segmentation in one time is hard to completely segment targets, we use segmentation twice to coarse-to-fine segment infrared image to ROI regions and background. As shown in Figs.2 (a2-b2), histogram equalization is firstly used to

homogenize the intensity space and transform the distribution of histogram to be uniform distribution. Then by using a self-adaptive threshold $thres_0$, the infrared image, shown in Fig.2 b1, is coarsely segmented in order to get rid of most background regions, shown in Fig.2 c1. Later histogram equalization is used to segment image, shown in Fig.2 (c2-d2), to enhance the contrast of the segmented image. Finally another self-adaptive threshold $thres_1$ is applied to segment the enhanced image. The candidate targets shown in Fig.1d are searched and detected in the segmented image shown in Fig.2 e1.

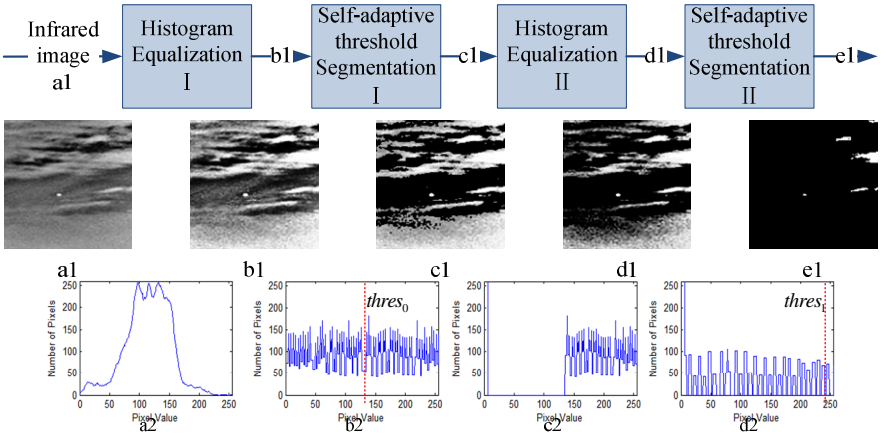

Fig. 2. The flowchart of segmentation based on non-linear histogram equalization, coordinate images in each steps and their histograms. Images (a1 to e1) respectively show respectively the input infrared image, intermediate results and segmented image while Images (a2-d2) show the histograms corresponding to a1-d1.

Assuming that the infrared image has m grayscales defined by $X=\{x_i\}_{i=0}^{m-1}$, its cumulative distribution function in grayscale x_i is defined as

$$c(x_i)=\sum_{i=0}^{m-1}p(x_i) \tag{17}$$

Where $p(x_i)$ denotes as the probability of grayscale x_i in histogram. It is clear that, after histogram equalization, the distribution of grayscales follows uniform distribution. Thus assuming that there are m_1 grayscales $Y=\{y_j\}_{j=0}^{m_1-1}$ reserved after histogram equalization, based on Eq.(15) and Eq.(16), the histogram equalization function is defined as:

$$y_j=\left\lfloor\frac{(m-1)c(x_i)}{\sum_i p(x_i)}+0.5\right\rfloor\times\left\lfloor\frac{256}{m_1}\right\rfloor \tag{18}$$

Then the self-adaptive threshold $thres_0$ is denoted as

$$thres_0 = \left\{ y_t \left| \sum_{j=0}^{t} p\left(y_j\right) \le 0.5 \sum_{j=0}^{m_1-1} p\left(y_j\right) \le \sum_{j=0}^{t+1} p\left(y_j\right) \right. \right\} \tag{19}$$

The threshold is only related to m_1, where m_1 can be determined by

$$m_1 = \left\{ m \left| 1 \le m_1 \le m, \min\left\{ \left| \sum_{j=1}^{t} y_j - 0.5 \right| \right\} \right. \right\} \tag{20}$$

Then histogram equalization is applied again. Assuming that there are m_2 grayscales $Z = \{z_k\}_{k=0}^{m_2-1}$ reserved and the histogram equalization function is defined as:

$$z_k = \left\lfloor \frac{(m_1 - 1)c\left(y_j\right)}{\sum_j p\left(y_j\right)} + 0.5 \right\rfloor \times \left\lfloor \frac{256}{m_2} \right\rfloor \tag{21}$$

Where $m_2 = m_1/2$ and the self-adaptive threshold $thres_1$ is denoted as

$$thres_1 = \frac{z_{m_2-2} + z_{m_2-1}}{2} \tag{22}$$

Histogram equalization is a contrast adjustment without affecting the global contrast by extending intensity range of image and merging neighboring, low probability gray levels. This property ensures the feasibility of our segmenting method that the high intensity regions in original image can always map to the high one in equalized image. But it also leaves some background pixels whose intensities are similar to that of target region, in the segmenting result. To capture the real target from those backgrounds, a selection method is then presented by using confidence measure and empirical size constraint.

3 Confidence Measure of Candidate Targets

By defining the regions except $\mathbb{A}$ in the original infrared image as $\mathbb{B}$, the confidence of A_i takes the form

$$C\left(A_i\right) = C_1\left(A_i\right) \times C_2\left(A_i\right) \times C_3\left(A_i\right) \tag{23}$$

Where C_1 shows the size confidence

$$C_1(A_i)=\begin{cases}0 & r_i > t_r \\ 1 & r_i \le t_r\end{cases} \tag{24}$$

And t_r indicates the maximum value of the radius r_i. C_2 shows the intensity confidence [9]

$$C_2(A_i)=\frac{1}{1+e^{-\lambda_1(\mu_{\mathbb{A}}-\mu_{A_i})}} \tag{25}$$

$\mu_{\mathbb{A}}$ is the mean grey value of $\mathbb{A}$. μ_{A_i} is the mean grey value of A_i. The higher the intensity of A_i, the more the value of C_2 is close to 1; otherwise, the value of C_2 is close to 0. C_3 shows the contrast confidence

$$C_3(A_i)=\frac{1}{1+e^{-\lambda_2(\mu_{\mathbb{A}}-\mu_{\mathbb{B}}-(\mu_{A_i}-\mu_{B_i}))}} \tag{26}$$

$\mu_{\mathbb{B}}$ is the mean grey value of $\mathbb{B}$ and μ_{B_i} is the mean grey value of B_i. If the candidate target has a higher contrast than its neighboring background, the value of C_3 is close to 1; otherwise, the value is close to 0. C_2 and C_3 are sigmoid functions whose slopes are controlled, respectively, by λ_1 and λ_2, where $\lambda_1, \lambda_2 \sim [1,+\infty)$. To sum up, the candidate target A_i with the largest confidence is decided to be the real one.

4 Experimental Work

Five typical infrared images sequences are selected to test the proposed method. Each image sequence contains 100 images with size 150×150 and the experiments were performed on MATLAB R2011a software in a 3.2GHZ personal computer. The slopes $\lambda_1=\lambda_2=5$ and the radius' maximum value $t_r=7$.

First, five images chosen from the five infrared image sequences, respectively, were tested and the results were intuitively performed in Fig. 3. Figs. 3a-e indicate the infrared images are under complex background such as roadside, sky, and sky-sea, sea background and heavy fog. It is clear that the real targets in these images are detected correctly.

Then, the Receiver Operation Characteristic (ROC) curves of the target regions detected by the LSSVM method [1], the Facet method [2], the GST method [3] and our method were compared for the robustness of the detected target region. The true positive rate is defined as the ratio of the detected pixels' number to the one of the real targets' pixels, and the false-positive rate is denoted as the ratio of the false alarms' number to the total pixels' in the infrared image [5]. The ROC curves in Figs. 4a–e

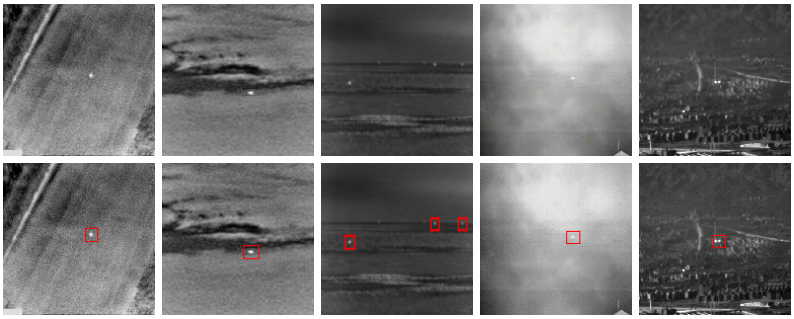

Fig. 3. Infrared images and result images under different backgrounds, where a-e label a couple of infrared image and its result

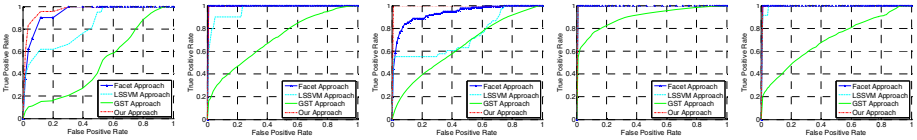

Fig. 4. The ROC curves of four methods in five infrared images shown in Fig.3 a~e

successively referred to the targets shown in Figs. 3a–e. As shown in Fig. 4, the proposed method represented by the green curves performs much better than all the other three methods.

In order to objectively compare our method with the others, a common metric [5] is applied to evaluate the accuracy of small target detection method by assessing their background suppression.

$$\text{Signal to clutter ratio gain:} \qquad SCR\ Gain = (S/C)_{out} / (S/C)_{in} \tag{27}$$

where S is the signal amplitude and C is the clutter standard deviation in a single frame. Meanwhile the time consuming performance is also considered for comparison.

Experimental results, including SCR gain and run time, are listed in Table 1, which indicates that our method and LSSVM method have the best two time consuming performance. Moreover our method performs better in accuracy than LSSVM method. In addition, the methods with best SCR gain are our method and GST methods. But GST method is slower than ours. It is evident that our method maintains better performances under different backgrounds. Thus it is a correct and efficient method for small target detection in infrared image.

Table 1. Performances comparison of four small target detective methods while a, b, c, d and e present the infrared images shown in Fig.3

Method	Metrics	a	b	c	d	e
Facet	SCR gain	5.437	6.906	11.43	8.975	1.781
	Time (s)	0.446	0.420	0.381	0.659	0.530
LS-SVM	SCR gain	6.331	8.490	9.127	5.255	2.349
	Time (s)	0.527	**0.233**	**0.250**	0.603	**0.222**
GST	SCR gain	20.25	**27.03**	**19.84**	12.42	2.454
	Time (s)	0.416	0.431	0.612	0.757	0.457
Ours	SCR gain	**28.64**	21.76	14.66	**43.90**	**6.881**
	Time (s)	**0.277**	0.263	0.302	**0.376**	0.384

Table 2. The detection ratio (%) of four target detection methods in five infrared image set, where a~e correspond to infrared image a~e shown in Fig.3 respectively

Method	a	b	c	d	e
LSSVM	82.1	80.5	80.8	60.3	52.3
GST	74.5	20.5	75.6	63.5	47.5
Facet	59.8	79.5	75.1	87.3	53.5
Ours	**96.7**	**87.0**	**93.0**	**96.8**	**70.9**

For further comparison, five infrared image sets are used to validate these methods. Table 2 lists the detective ratios, which show the proportions of detected real targets for four methods in the five infrared image sequences. In Table 2, it is clear that LS-SVM method shows stable performance contrast with GST method. The facet method in all database performs robust as well as our method, but its detection ratio is lower. It is evident that our method shows better detection ratios under different situations, thus it is a robust method for infrared small target detection.

5 Conclusion

Based on coarse-to-fine segmentation and confidence analysis, our approach can be used in infrared small target detection and verified good performances via experiments. The objective function origins from POME ensure the targets segmented from the infrared image completely. And the abnormality based confidence can be able to distinguish real targets and false alarms. However, due to the assumption that targets are brighter than neighboring background, our method cannot be used in infrared image whose targets in dark. Experimental results suggest that the presented method is efficient and robust.

References

1. Shaik, J., Iftekharuddin, K.M.: Detection and tracking of targets in infrared images using Bayesian techniques. Optics and Laser Tech. **41**, 832–842 (2009)
2. Yang, L., Yang, J., Yang, K.: Adaptive detection for infrared small target under sea-sky complex background. Electron. Lett. 40(17), 1083–1085 (2004)
3. Gu, Y.F., Wang, C., Liu, B.X., Zhang, Y.: Kernel-based nonparametric regression method for clutter removal in infrared small-target detection applications. IEEE Geosci. Remote Sens. Lett. **7**, 469–473 (2010)
4. Wang, P., Tian, J.W., Gao, C.Q.: Infrared small target detection using dinfraredectional highpass filters based on LS-SVM. Electron. Lett. **45**(3), 156–158 (2009)
5. Wang, G.D., Chen, C., Shen, X.B.: Facet-based infrared small target detection method. Electron. Lett. **41**(22), 1244–1246 (2005)
6. Gao, C.D., Zhang, T.Q., Li, Q.: Small infrared target detection using sparse ring representation. IEEE Aerosp. Electron. Syst. Mag. **27**(3), 21–30 (2012)
7. Qi, S.X., Ma, J., Tao, C.: A Robust Directional Saliency-Based Method for Infrared Small-Target Detection Under Various Complex Backgrounds. IEEE Trans. Geosci. and Remote Lett. **99**, 1–5 (2012)
8. Wang, X., Lv, G.F., Xu, L.Z.: Infrared dim target detection based on visual attention. Infrared Physics and Technology **55**(6), 513–521 (2012)
9. Yilmaz, K., Shafique, M.: Target Tracking in Airborne Forward Looking Infrared Imagery. Image and Vision Comput. J. **21**(7), 623–635 (2003)
10. Parag, T.: Coupled label and intensity MRF models for IR target detection. In: CVPR Workshops, pp. 7–13 (2011)
11. Zhang, L., Wu, B., Nevatia, R.: Pedestrian detection in infrared images based on local shape features. In: CVPR Workshop OTCBVS, pp. 1–8 (2007)