SMALL TARGET DETECTION USING AN OPTIMIZATION-BASED FILTER

Kai Xie[1,2], *Keren Fu*[1,2], *Tao Zhou*[1,2], *Jie Yang*[1,2]*, *Qiang Wu*[3], *Xiangjian He*[3]

[1]Institute of Image Processing & Pattern Recognition, Shanghai Jiao Tong University, China
[2]Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai
[3]School of Computing and Communications, University of Technology, Sydney, Australia

ABSTRACT

Small target detection is a critical problem in the Infrared Search And Track (IRST) system. Although it has been studied for years, there are some challenges remained, e.g. cloud edges and horizontal lines are likely to cause false alarms. This paper proposes a novel method using an optimization-based filter to detect infrared small target in heavy clutter. First, we design a certain pixel area as active area. Second, a weighted quadratic cost function is performed in the active area. Finally, a filter based on statistics of active area is derived from the cost function. Our method could preserve heterogeneous area, meanwhile, remove target region. Experimental results show our method achieves satisfied performance in heavy clutter.

***Index Terms*—** small target detection; heavy clutter; active area; optimization-based filter; area statistics

1. INTRODUCTION

Small target detection is a critical problem in the Infrared Search And Track (IRST) system. Although it has been studied for years [1, 2, 3, 4, 5], there are some challenges remained. The reasons are as follows: first, features such as texture and color are unavailable for small targets when they are far away from the infrared sensor. Second, heterogeneous areas such as cloud edges, sky-sea lines etc. may be falsely detected as small targets.

Background estimation based small target detection method is widely studied in recent years [6, 7]. These methods detect small targets from the residual image, i.e. subtracting the estimated background image from the original input. The detection performance relies highly on the quality of the estimated background. A good background estimation method should preserve the heterogeneous area as much as possible and exclude the target region. In other word, it requires the method to effectively distinguish the target region and the heterogeneous area. The 2-D least mean square (TDLMS) method [6] minimizes the difference between an input image and a background image that is estimated by the weighted average of neighboring pixels. The TopHat method [7] estimates background by a morphological opening operator with structure element. These methods are less effective because they could not distinguish the target region and the heterogeneous region, and the background estimation image contains much clutter.

Existing background suppression methods are mainly based on the filtering methods [8, 9]. The LS-SVM [9] method uses filter templates which can suppress most part of the correlative background but may be easily interfered because of the strong fluctuation of background clutters. Yang et al. [8] provide an adaptive Butterworth high pass filter by suppressing low frequency components to achieve the purpose for enhancing targets. This method makes use of the diversity between small targets and backgrounds in frequency domain. Insufficiently, it is inoperative to noises or clutter fluctuation which also possesses high frequency in the image.

These methods [6, 7, 8, 9] may achieve good results on simple background, but not on heavy clutter background. The main reason is they cannot distinguish small targets from heavy clutter effectively. In this paper, we focus on removing target while preserving the high-frequency component in the background. We proposes a novel method using an optimization-based filter to solve this problem. A weighted quadratic cost function is performed in active areas, then a filter based on statistics of active areas is derived. The proposed method achieves better performance, when compared with TDLMS [6], TopHat [7], LS-SVM [9], especially in the heavy clutter.

2. THE PROPOSED METHOD

Generally, the IR image model can be formulated as:

$$f = f_T + f_B + n, \tag{1}$$

where f, f_T, f_B and n are IR image, target image, background image and random noise, respectively. n is assumed to follow Gaussian distribution with mean 0 and variance σ^2 [10].

*Jie Yang is the corresponding author(email: jieyang@sjtu.edu.cn). This research is partly supported by NSFC, China (No: 61273258) and SAIST Funding.

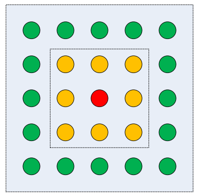

Fig. 1. Green color represents pixels in active area, yellow color represents pixels inside inner window, red color is center pixel.

Background estimation based small target detection method aims to estimate f_B. We propose an optimization-based filtering method to solve this problem. A weighted quadratic cost function is performed in active areas, then a filter based on statistics of active areas is derived. We apply our filter on IR image using sliding window manner, then we can obtain a background estimation image of the same size of original input.

active area: Active area is an area between inner and outer window, shown in Fig.1. M, N are sizes of inner window and outer window, respectively. We denote active area as $\Omega = \{1, 2, ..., n\}$. Where n is the total number of pixels in the area. Intensity of IR image's pixel $i \in \Omega$ is denoted by x_i, and intensity of background image's pixel $i \in \Omega$ is denoted by y_i.

weighted quadratic model: The key assumption of our method is that a local linear model existing in pixels of background image [11, 12], but IR image contains not only background but also target. The filter should be designed to eliminate effects of target region, so we propose a weighted quadratic model to solve this problem.

The local linear model is:

$$y_i = ax_i + b, \quad i \in \Omega \tag{2}$$

the weighted quadratic model is:

$$\min_{[a,b]^T} \sum_{i\in\Omega} (y_i - \begin{bmatrix} x_i \\ 1 \end{bmatrix}^T \begin{bmatrix} a \\ b \end{bmatrix})^2 w_i + \epsilon a^2 \tag{3}$$

where a, b are parameters in local linear model, and ϵ a regularization parameter penalizing large a. w_i gives different importance to different pixel in active area. It can be calculated by intensity value of center pixel and pixel i: $w_i = e^{-c*||x_0 - x_i||_2^2}$, where c a constant and x_0 the intensity of center pixel in the original input. We notice that the more similar intensity values of x_0, x_i are, the larger w_i is.

By introducing some new denotes: $Y = [y_1, y_2, ..., y_n, 0]^T$,

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_n & \sqrt{\epsilon} \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \tag{4}$$

and

$$W = \begin{bmatrix} w_1 & 0 & \cdots & 0 & 0 \\ 0 & w_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & w_n & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \tag{5}$$

the cost function can be rewrote as:

$$\min_{[a,b]^T} (X^T \begin{bmatrix} a \\ b \end{bmatrix} - Y)^T W (X^T \begin{bmatrix} a \\ b \end{bmatrix} - Y) \tag{6}$$

We could easily obtain:

$$\begin{bmatrix} a \\ b \end{bmatrix} = (XWX^T)^{-1} XWY \tag{7}$$

With some matrix operations, a, b could be calculated:

$$a = \frac{\frac{\sum_{i\in\Omega} x_i y_i w_i}{\sum_{i\in\Omega} w_i} - \frac{\sum_{i\in\Omega} x_i w_i}{\sum_{i\in\Omega} w_i} * \frac{\sum_{i\in\Omega} y_i w_i}{\sum_{i\in\Omega} w_i}}{\frac{\sum_{i\in\Omega} x_i^2 w_i}{\sum_{i\in\Omega} w_i} - (\frac{\sum_{i\in\Omega} x_i w_i}{\sum_{i\in\Omega} w_i})^2 + \frac{\epsilon}{\sum_{i\in\Omega} w_i}} \tag{8}$$

$$b = \frac{\sum_{i\in\Omega} y_i w_i}{\sum_{i\in\Omega} w_i} - a * \frac{\sum_{i\in\Omega} x_i w_i}{\sum_{i\in\Omega} w_i} \tag{9}$$

We notice that $\frac{w_i}{\sum_{i\in\Omega} w_i}$ is always in the range $[0, 1]$, and $\sum_{i\in\Omega} \frac{w_i}{\sum_{i\in\Omega} w_i} = 1$. It could be considered as a normalized probability. From the view point of statistics, a, b could be considered as:

$$a = \frac{E(xy) - E(x)E(y)}{E(x^2) - (E(x))^2 + \frac{\epsilon}{\sum_{i\in\Omega} w_i}} \tag{10}$$

$$b = E(y) - a * E(x) \tag{11}$$

where x, y are random variables that associate with x_i, y_i, respectively.

In order to learn $[a, b]^T$, we set $y = x$ in learning phase, and we obtain:

$$a = \frac{\sigma^2(x)}{\sigma^2(x) + \frac{\epsilon}{\sum_{i\in\Omega} w_i}} \tag{12}$$

$$b = (1 - a) * E(x) \tag{13}$$

where $E(x), \sigma^2(x)$ are the expectation and variance of intensities of pixels in active area, respectively. They are two statistical quantities of pixels.

Thus, the background pixel's intensity:

$$y^* = ax_0 + b = ax_0 + (1 - a) * E(x) \tag{14}$$

a trade-off between the intensity of center pixel in original input and the expectation of intensities of pixels in active area.

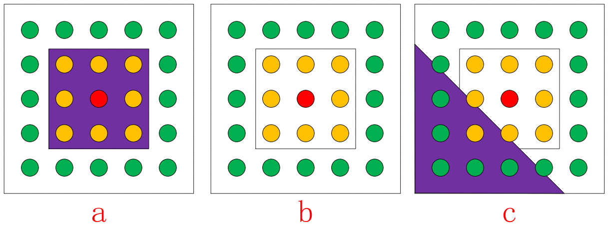

Fig. 2. Green color represents pixels in active area, yellow color represents pixels inside inner window, red color is center pixel. Purple color represents high intensity region, and white color represents low intensity region.

3. MOTIVATION AND ANALYSIS

In order to eliminate effects of target region, we design certain active area and weighted quadratic model mentioned above. Fig.2 represents local patch centering at three different regions. From left to right, they are target region, homogenous region and inhomogeneous region, respectively. When the local patch is located in the target region, all weights are small, as is shown in Fig.2(a). When the local patch is located in the homogenous region, all weights are big, as is shown in Fig.2(b). When the local patch is located in the inhomogeneous region, some weights are big, the others are small, as is shown in Fig.2(c).

From Fig.2, we notice that weights in target region and homogenous region equal approximately; weights in inhomogeneous region fall into different categories, pixels that have similar intensities with center pixel acquire big weights, and vice versa. On the other hand, intensities of active pixels in target region and homogenous region equal approximately; intensities of active pixels in inhomogeneous region fall into different categories in accordance with weights, and in the same category intensities varies to a small extent. As a result of above analysis, $\sigma^2(x)$ in inhomogeneous region prefers variance of intensities of pixels in the category with big weights, it should be a small value, and in target region and homogenous region, $\sigma^2(x)$ should also be small. Thus, we could obtain:

$$\sigma_a^2 \approx \sigma_c^2 \approx \sigma_b^2 \tag{15}$$

On the other hand, we could obtain inequations:

$$\{\sum_{i\in\Omega} w_i\}_a < \{\sum_{i\in\Omega} w_i\}_c < \{\sum_{i\in\Omega} w_i\}_b \tag{16}$$

where footnotes a, b, c are in accordance with Fig.2, and we could immediately obtain inequations:

$$a_a < a_c < a_b \tag{17}$$

according to (12).

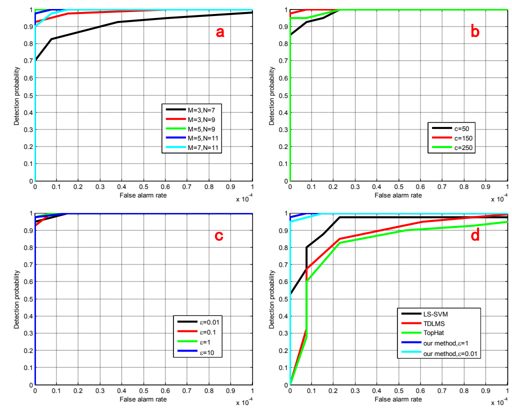

Fig. 3. ROC curves of different parameters and methods in test data.

We notice that $E(x)$ is the expectation of intensities of pixels in active area: in target and homogeneous region it should be the mean of intensities(because of all weights are approximate equivalent), and in inhomogeneous region it is approximately the partial mean of intensities of pixels which belong to the category with big weights.

Based on the above analysis: according to (14), when the local patch is located in the target region, the intensity of background estimation pixel prefers the mean of pixels in active area; when the local patch is located in the homogenous region, the intensity of background estimation pixel favours the intensity of center pixel in original input; when the local patch is located in the inhomogeneous region, the intensity of background estimation pixel is a trade-off value between the intensity of center pixel in original input and the partial mean of intensities of pixels which belong to the category with big weights. Thus, our method could remove target, meanwhile, preserve inhomogeneous region.

4. EXPERIMENTS AND RESULTS

To compare the performance of methods quantitatively, receiver operating characteristic (ROC) curves are employed. We choose TopHat [7] and TDLMS [6] filtering method as two baseline methods. Moveover, LS-SVM [9] filtering method is also chosen as the comparison method in this paper since the method is well studied and has a good performance. The images chosen in the experiments contain more than one hundred targets and eight different categories of clutter environments. We take these images as test data.

4.1. Analysis On Effects Of Parameters

The proposed method have four key parameters: M, N, c and ϵ:

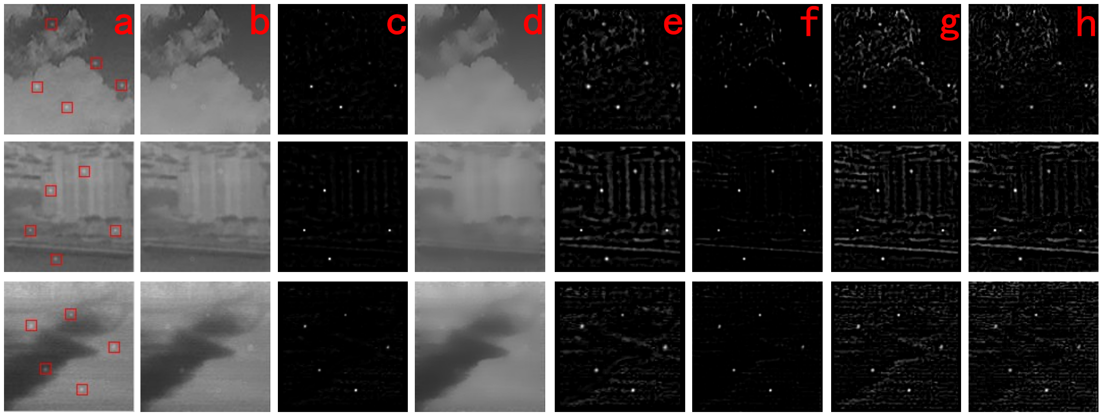

Fig. 4. Comparison between four methods. Red rectangles contain ground truths of small targets. (a)original IR image, (b)background estimation image with $\epsilon = 0.01$, (c)the residual image of (b), (d)background estimation image with $\epsilon = 1$, (e)the residual image of (d), (f)LS-SVM, (g)TDLMS, (h)TopHat.

N determines the size of local patch, controls the influence sphere of our filtering method. A large N implies that more pixels are considered for our method and vice versa. N should be as small as possible, for correlation between pixels decreases when distance of them increases. M is the size of inner window, and it should be big enough to cover the reference object. The difference between M and N determines the total number of pixels in the active area, and it should be well set up. If it is too small, the number of samples for area statistics is insufficient. On the other hand, a large value causes a large N or a small M. We choose five groups of M and N to test our method with parameters configuration of $c = 150, \epsilon = 1$. The evaluation results are shown in Fig.3(a).

c is the parameter using calculating w_i. When $c = 0$, all the weights equal to one. It makes our weighted quadratic model degenerate to non-weighted model. When $c = +\infty$, all the weights equal to zero. It lets our model make no sense. The effect of changing c with parameters configuration of $M = 5, N = 11, \epsilon = 1$ is shown in Fig.3(b).

ϵ is a regularization parameter penalizing large a. When $\epsilon = 0$, $a \equiv 1$, and the filtering image is just the original IR image, according to (12) and (14). On the other hand, when $\epsilon = +\infty$, $a \equiv 0$. when the local patch is located in the homogeneous and target region, our method performs as a mean filter in active area. Meanwhile, the local patch is located in the inhomogeneous region, our method performs as an edge-preserving filter. The effect of changing ϵ with parameters configuration of $M = 5, N = 11, c = 150$ is shown in Fig.3(c).

4.2. Comparison To Baseline Methods

Fig.4 gives comparisons between the proposed method and the baseline methods TopHat [7], TDLMS [6], LS-SVM [9], whose parameters are well adjusted to achieve their best performances in test data. To compare the performance of these methods quantitatively, ROC curves are employed, shown in Fig.3(d).

As is shown in Fig.4(g)(h), the residual image of TDLMS and TopHat contains much clutter, for they can not sperate inhomogeneous region from target region. Small target may be detected in the inhomogeneous region falsely. Performance of quantitative evaluation is shown in Fig.3(d).

LS-SVM could suppress most of the clutter in the residual image, but unfortunately in the same time target may be suppressed to an extent, as is shown in Fig.4(f). The visual effects are quite good, but performance of quantitative evaluation is not so good, as is shown in Fig.3(d).

When ϵ is set to 0.01 and 1, the background estimation image and the residual image are shown in Fig.4(b)(c) and (d)(e), respectively. A large ϵ causes losing of small fluctuation in background estimation image, and small target is removed much cleaner. It makes the residual image contains some small fluctuation of clutter, but the target is more clear. The visual effects with small ϵ are quite good, but performance of quantitative evaluation with large ϵ are better, as is shown in Fig.3(d), where parameters configuration is $M = 5, N = 11, c = 150$.

5. CONCLUSION

This paper proposes a novel method using an optimization-based filter to detect infrared small target in heavy clutter. A weighted quadratic cost function is performed in active areas, then a filter based on statistics of active areas is derived. Experimental results show our method achieves satisfied performance in heavy clutter.

6. REFERENCES

[1] Ruiming Liu, Yanhong Lu, Chenglong Gong, and Yang Liu, "Infrared point target detection with improved template matching," *Infrared Physics & Technology*, vol. 55, no. 4, pp. 380–387, 2012.

[2] Sungho Kim, Yukyung Yang, and Joohyoung Lee, "Horizontal small target detection with cooperative background estimation and removal filters," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1761–1764.

[3] Baojun Qi, Tao Wu, Bin Dai, and Hangen He, "Fast detection of small infrared objects in maritime scenes using local minimum patterns," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 3553–3556.

[4] Keren Fu, K Xie, C Gong, Irene YH Gu, and Jie Yang, "Effective small dim target detection by local connectedness constraint," in *IEEE int'l conf ASSP (ICASSP 2014)*, 2014.

[5] Suyog D Deshpande, H Er Meng, Ronda Venkateswarlu, and Philip Chan, "Max-mean and max-median filters for detection of small targets," in *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*. International Society for Optics and Photonics, 1999, pp. 74–83.

[6] Mohiy M Hadhoud and David W Thomas, "The two-dimensional adaptive lms (tdlms) algorithm," *Circuits and Systems, IEEE Transactions on*, vol. 35, no. 5, pp. 485–494, 1988.

[7] Xiangzhi Bai and Fugen Zhou, "Analysis of new top-hat transformation and the application for infrared dim small target detection," *Pattern Recognition*, vol. 43, no. 6, pp. 2145–2156, 2010.

[8] L Yang, J Yang, and K Yang, "Adaptive detection for infrared small target under sea-sky complex background," *Electronics Letters*, vol. 40, no. 17, pp. 1083–1085, 2004.

[9] P Wang, JW Tian, and Ch Q Gao, "Infrared small target detection using directional highpass filters based on ls-svm," *Electronics letters*, vol. 45, no. 3, pp. 156–158, 2009.

[10] Kai Xie, Keren Fu, Tao Zhou, Junhao Zhang, Jie Yang, and Qiang Wu, "Small target detection based on accumulated center-surround difference measure," *Infrared Physics & Technology*, 2014.

[11] Kaiming He, Jian Sun, and Xiaoou Tang, "Guided image filtering," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 6, pp. 1397–1409, 2013.

[12] Anat Levin, Dani Lischinski, and Yair Weiss, "A closed-form solution to natural image matting," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 2, pp. 228–242, 2008.