Journal of Information Hiding and Multimedia Signal Processing ©2016 ISSN 2073-4212
Ubiquitous International Volume **7**, Number **2**, March **2016**

A Small Target Detection Method based on Human Visual System and Confidence Measurement

Jingli Yang, Fudan Li, Zhen Sun and Shouda Jiang

School of Electrical Engineering and Automation
Harbin Institute of Technology
No.2 Yi-Kuang Street NanGang District , Harbin, China
icehit0615@163.com

Received January, 2016; revised February, 2016

Abstract. *Robust small target detection of low signal-to-noise ratio (SNR) is very important in infrared search and track applications.Due to the complex background, current algorithms have some unsolved issues about small target detection capabilities, such as detection rate, false alarm rate and speed. They generally improve one or two of the aforementioned detection capabilities while sacrificing the others. In order to pursue good performance in detection rate, false alarm rate and speed simultaneously, an infrared small target detection algorithm based on mechanism of visual attention and two layers structure was proposed.Firstly, we detected salient regions that may contain targets by morphological Top-hat and threshold operation. Then target recognition was performed by confidence measurement based on histogram in the salient regions. Experimental results show the proposed algorithm has ideal robustness and efficiency for real infrared small target detection applications.*
Keywords: Small target detection, Human visual system, Top-hat morphological filter, Confidence measurement

1. **Introduction.** Infrared (IR) small-target detection plays a critical role in large amounts of practical projects such as infrared warning and defense alertness, in which not only accuracy is needed but also robustness is required[1]. Various algorithms have been developed in the past few decades. Including filter-based algorithm[2, 3, 4, 5], wavelet based algorithms [6], mathematical morphology based algorithms [7] and a series of algorithm based on Fourier Transform [8]. Filter-based algorithms such as kernel filters, max-mean filter, max-median filter and adaptive lattice algorithm enhance small target and suppress the background clutter by using filters to filter the possible target regions. Two-dimensional least mean square (TDLMS) filter acts as a background prediction operator when applied to the infrared small target detection. Wavelet based algorithms enhance small target through extracting the different properties between the target region and clutter background. Mathematical morphology based algorithms mainly choose suitable structuring element and operations. Algorithms based on Fourier Transform convert the airspace information into the frequency domain.

Although numerous methods have been proposed, many of them may fail in certain circumstances, e.g., ground sky background [9], which is common in the helicopter view. In this situation, targets are easily mixed up with the background clutters in size and easily overlapped by vegetations, roads, rivers, bridges[10], resulting in a huge false rate in traditional algorithm.

In order to design an appropriate method, a small target detection method inspired by human visual system (HVS) and confidence measurement has been designed in this paper. HVS is a kind of layered image processing system consisting of optical system, retina and visual pathways. The rest of this paper is organized as follows. In section 2, the framework of the proposed algorithm for small target detection will be described. The experimental results for the proposed method are presented in Section 3 and conclusions are drawn in Section 4.

2. Target Detection Algorithm Based on HVS and Confidence Measurement.

2.1. Framework of the Proposed Two Stage Algorithm - A Brief Description.

HVS divides the scene into small patches and select important information through visual attention selection mechanism to make it easy to understand and analyze. On the other hand, as a component of low-level artificial vision processing, it facilitates subsequent procedures by reducing computational cost, which is a key consideration in real-time applications. Based on the above knowledge, we propose a framework consisting of two stages inspired by HVS[11] as follows (See Figure.1). In our algorithm, the detection process is divided into two steps, pre-detection based on mechanism of visual attention for orientation and detection stage. In the pre-detection stage, a saliency map (SM) is obtained and the most salient regions through threshold operation are picked up to improve detection speed. In the second stage, a confidence measurement for every candidate region is taken to get the real target.

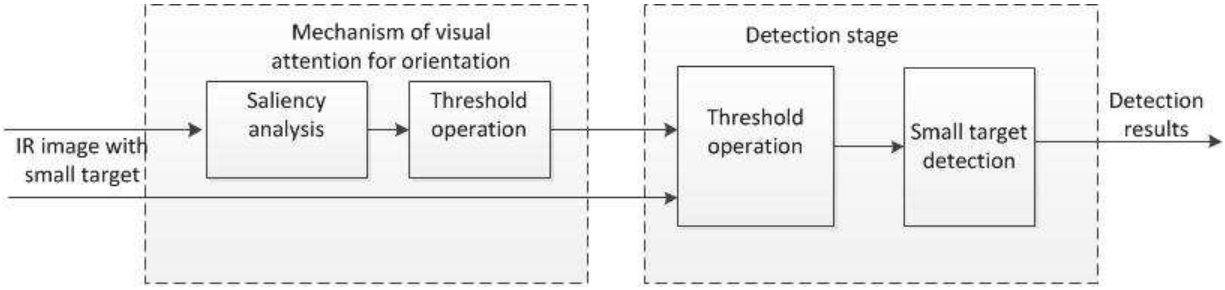


FIGURE 1. The proposed framework inspired by HVS

Due to the two-layer structure, the algorithm computational complexity becomes the prime concern. Saliency detection method about morphological operation is chosen to simplify computational complexity and improve processing speed. Details will be described in section 2.2.Candidate targets can be got through morphological operation followed by threshold operation.

In addition, HVS will conduct some complex processing to separate targets from background clutters. Since background clutters are similar to the infrared targets in the size and shape, it is very difficult to distinguish them with the method of traditional algorithms. The confidence measurement which incorporates statistical features of candidates will be very effective to distinguish the real target between clutters. Details will be described in section 2.3.

2.2. Mechanism of Visual Attention for Orientation.

The thermal images are obtained by sensing the radiation in the infrared (IR) spectrum, which is either emitted or reflected by the object in the scene. Due to this property, IR images can provide information which is not available in visual image. However, in contrast to visual images, the infrared image have extremely low signal to noise ratio(SNR) which results in limited

information for performing detection tasks. Actually, although the target has extremely low SNR, there are still some difference between the target and surroundings. According to the delicate difference especial about intensity and morphology, HVS can rapidly find the saliency region or the region of interest (ROI) and transfer perspective according to the degree of salience.

Mathematical morphological operation measures and extracts the corresponding image forms using a certain structuring element, which obviously simplifies image data, keeps basic image feature and dislodges useless structure. Based on the theory above, Top-hat transform of gray-scale morphology is chosen due to its good performance in enhancing dim targets and the ability of parallel computation.

All of the mathematical morphological operations work with two sets. One set is the analysis of original image and the other set is called structuring element. For gray-scale image, the dilation and erosion operation of pixel $f(x, y)$ are defined by

$$\begin{aligned}(f \oplus g)(x, y) &= \underset{i,j}{Max} \left(f(x - i, y - j) + g(i, j)\right) \\ (f \ominus g)(x, y) &= \underset{i,j}{Min} \left(f(x - i, y - j) - g(i, j)\right)\end{aligned} \tag{1}$$

Here, $g(x, y)$ represents structuring element which serves as a kind of template. The operation of dilation using $g(x, y)$ as template which searches the maximum summation of the image within the scope of structuring element, while the operation of erosion searches the minimum minus of gray-scale.

The morphological opening and closing operations of f and g are defined respectively by combination dilation and erosion operations as follows:

$$\begin{aligned}f \circ g &= (f \ominus g) \oplus g \\ f \bullet g &= (f \oplus g) \ominus g\end{aligned} \tag{2}$$

Based on opening and closing operations, opening Top-hat (OTH) operation is defined as:

$$OTH_{f,g}(x, y) = f(x, y) - f \circ g(x, y) \tag{3}$$

Opening operation is an iterated operation which does erosion operation first and then makes dilation operation using the same structuring element. Due to this property, we can receive a significant map by opening Top-hat operation. After the final significant map is obtained, a threshold operation is used to pick out the most salient points. Then, we define the thresholds as followed:

$$T = \frac{1}{N_I} \sum_{j=1}^{N_I} S_j + k * \left(Max\left(S_j\right) - \frac{1}{N_I} \sum_{j=1}^{N_I} S_j \right) \tag{4}$$

Where S_j represents the saliency value of the j_{th} pixel ,N_I is the number of the pixels of the image. k is a parameter from 0 to 1 and be used to control the threshold changing from $\frac{1}{N_I} \sum_{j=1}^{N_I} S_j$ to $Max\left(S_j\right)$. We only deal with those areas above the threshold. Setting threshold can effectively reduce the number of salient areas, thus reduce the amount of calculation.

2.3. **The Theory of Confidence Measurement in the Second Stage.** Due to the target gray-scale information is not rich, a series of features of target and background are

overlaps which makes it barely possible to part extended targets. To overcome this shortcoming, we present a confidence measurement approach that incorporates four statistical characteristics of gray histogram, which can not only distinguish target between clutters but also effectively compensate features distribution contingency.

For gray-scale image, gray-scale histogram is defined as followed:

$$H_i = \frac{n_i}{N} \quad i = 0, 1 \cdots L-1 \tag{5}$$

Where i and L represent gray levels and their type numbers respectively, n_i represents the number of pixels which belongs to gray level i , and N is the total number of pixels.

Base on gray-scale histogram, there are several statistics indicators which can reflect histogram features. They are respectively defined as followed:

(1) Variance

$$\sigma^2 = \sum_{i=0}^{L-1} (i-\mu)^2 H(i) \tag{6}$$

Where μ is the mean gray.

(2)Skewness

$$\mu_s = \sigma^3 \sum_{i=0}^{L-1} (i-\mu)^3 H(i) \tag{7}$$

(3) Kurtosis

$$\mu_k = \sigma^4 \sum_{i=0}^{L-1} (i-\mu)^4 H(i) - 3 \tag{8}$$

(4) Entropy

$$\mu_e = \sum_{i=0}^{L-1} H(i) log_2[H(i)] \tag{9}$$

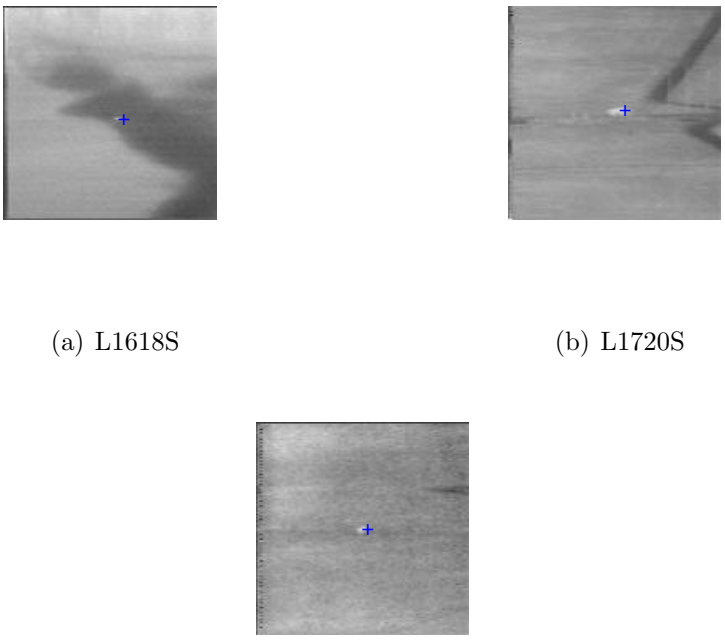
What the variance reflects is the numeral discrete distribution of the gray scale of one image. Where there the gray-scale is relatively uniform, the local variance is smaller; where there the fluctuation is obvious, the local variance is larger. Therefore, the difference between the target and background clutter can be calculated by the size of the variance. Skewness, which represents asymmetry of histogram, has a positive correlation with asymmetry, that is to say, the larger the skewness level is, and the more asymmetric the histogram is. We can easily know from the equation 7 that the skewness level variance of the target is positive, while background clutters one is negative. Kurtosis, which reflects the approximate state when the gray-scale distribution of the image approaches the mean value, is used to identify whether the gray-scale distribution of the image is centralized around the mean value, if the kurtosis is small, it indicates that the distribution is centralized. On the contrary, the distribution is dispersive. Entropy denotes uniformity coefficient of gray-scale distribution. In the whole image, the change of entropy might be submerged by noise, but in small region, with the appearance of the target, the coherence of the image was broken, which leads to change of entropy. Especially in small region, the change of gray-scale can cause great change of entropy. Once significant map has been achieved based on HVS, a confidence measurement for each candidate region is computed as the product of sigmoid function based on the above theory:

$$C = \prod_{i=1}^{4} C_i = \prod_{i=1}^{4} \frac{1}{1 + e^{-\lambda_i(\mu_i - \mu_{i0})}} \tag{10}$$

Where λ_1, λ_1, λ_1,λ_1 control the slope of the sigmoid respectively and μ_{10},μ_{20}, μ_{30}, μ_{40} are the offsets of the sigmoid μ_1, μ_2, μ_3, μ_4 are the variance, skewness, kurtosis and entropy of a region.

If a target region is bright with a high contrast with its neighborhood, then equation10 assigns C_1,C_2,C_3 ,C_4 close to 1, otherwise the confidence will be closed to 0. On the base of the above preparation, we define the centers of salient areas obtained in the section 2.2 as the suspected target centers and use the confidence measurement to classify these pixels with its neighbors. The candidate target with the highest confidence is decided to be the real one. Attributing to confidence measurement successfully separating the target from clutters, the accuracy can be improved to a satisfactory degree.

3. **Experimental Section.** Experiments on the real IR image sequence have been done. The sequence contains three groups of images with a resolution of $128 * 128$, obtained from United States Army Aviation and Missile Command (AMCOM). Each image involves one target, and the real targets are marked by asterisk(see Figure 2). All experiments were implemented by MATLAB software on a PC with 4-GB memory and 3.2-GHz Intel i5 dual processor.



(a) L1618S (b) L1720S

(c) L2208S

FIGURE 2. The three groups of the real IR image picked in different groups for experiments.

We randomly choose the background regions sized $16 * 16$ in the image beside the target area. The results of confidence measurement which are respectively applied to

three groups of images are shown in Figure3. It can be obviously seen that the confidence measurement can efficient distinguish the target between background.

$$\begin{aligned} TPR &= \frac{Quantity\ \ of\ \ true\ \ targets\ \ detected\ \ in\ \ images}{Quantity\ \ of\ \ true\ \ targets\ \ existing\ \ in\ \ images} * 100\% \\ FPR &= \frac{Quantity\ \ of\ \ false\ \ detected\ \ in\ \ images}{Quantity\ \ of\ \ true\ \ targets\ \ existing\ \ in\ \ images} * 100\% \end{aligned} \tag{11}$$

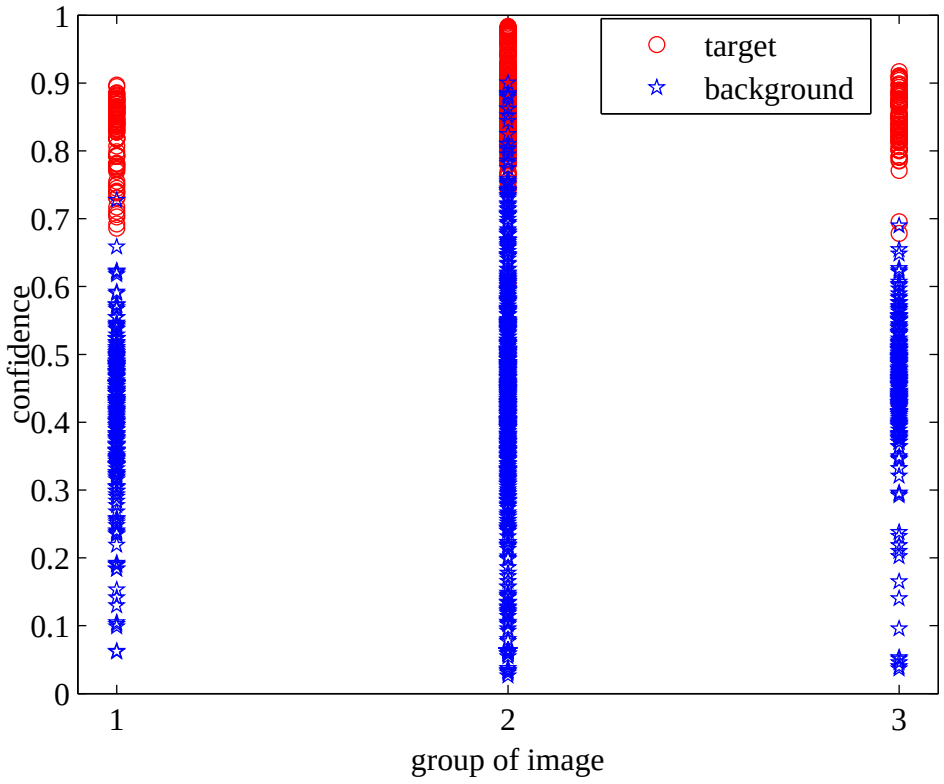


FIGURE 3. The confidence distribution of target and background.

To demonstrate the detection effect of the proposed algorithm, we compare it with other five IR small target detection methods, including max-medium method, the traditional Top-hat method, two-dimensional least mean square (TDLMS) method, LCM method and PFT method. To reveal the advantage of the proposed algorithm to the other five algorithms,we asset its threshold to different values and draw the relationship between detection rate (TPR) and false alarm rate (FPR) which are shown in Figure4- Figure.12.

Experimental results show that the parameter in equation 10 is suitable to be set from 0.2 to 0.6 in our work. We can get high TRP with all kinds of algorithms by selecting this parameters section so as to meet the requirements. At the same time, we can get the conclusion (see Figure.4, Figure.7 and Figure.10) that Top-hat and our method can achieve the good performance in three groups of image. LCM algorithm and PFT method have lower TPR. The result reveals that confidence measurement may reduce the detection rate of the algorithm, which is determined by the classifiers own accuracy.

Although the targets are successfully detected, there still exist many false alarms. We hope that the false alarm rate will be reduced. In Figure.5, Figure.8 and Figure.11, FPR

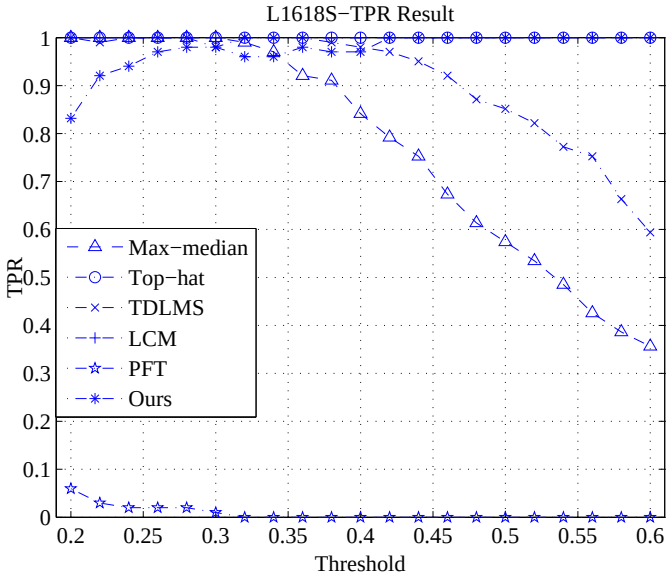


FIGURE 4. TPR of all method in images of L1618S.

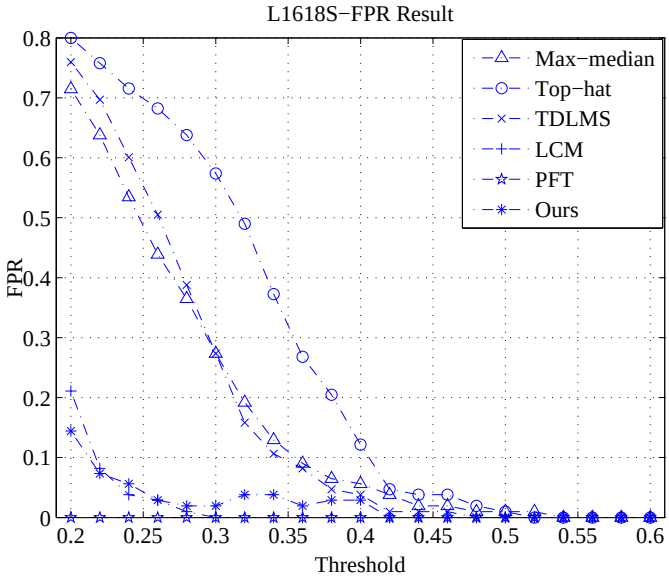


FIGURE 5. FPR of all method in images of L1618S.

of traditional methods without confidence are always higher than FPR of traditional methods without confidence are always higher than 50% in the same cases, nevertheless, FPR of our method is encouragingly lower than 20%.

Finally, we compare the computational cost of these detection methods for three groups of images (see Figure.6, Figure.9and Figure.12). The experimental results illustrate that Top-hat and PFT have the fastest computing speed among all these methods. Due to the newly added confidence measurement operation, the proposed algorithm is just a few seconds slower. While other algorithm can achieve better performance such as LCM in TPR and FPR in same scenes(see Figure.4 and Figure.8), but its detection speed is too slow. The proposed algorithm, can not only speed up the detection process, but also show a good performance in TPR and FPR in all the six methods.

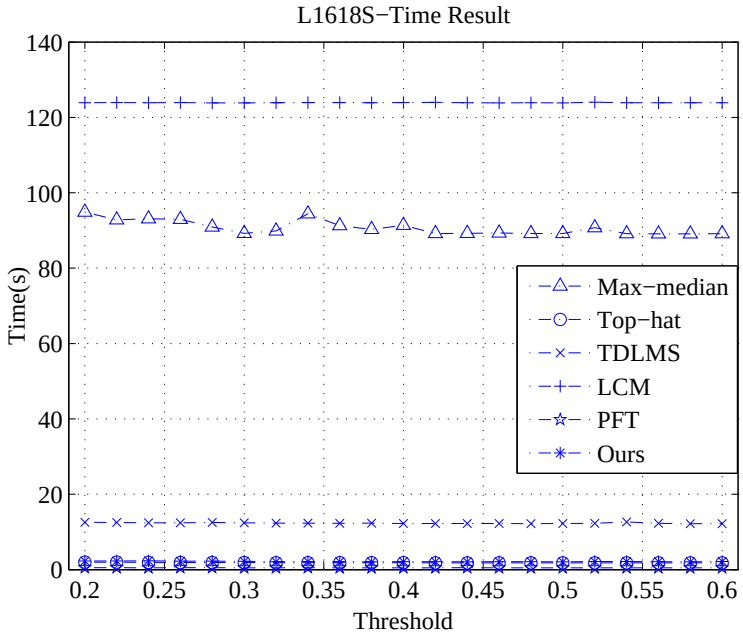


FIGURE 6. Time of all method in images of L1618S.

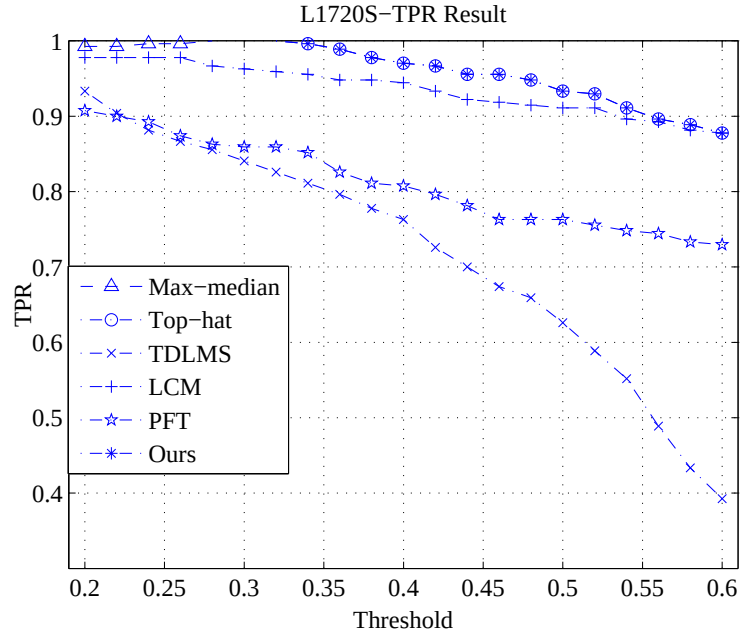


FIGURE 7. TPR of all method in images of L1720S.

We choose 0.44 as the threshold to compare the performance of all methods and list the index in table1, table2 and3. The result of the experiment shows that traditional algorithms have inevitable difficulties in detecting small target in complex background on the ground. By contrast, the proposed algorithm can greatly reduce the FPR.

4. **Conclusions.** In this paper, a robust IR small target detection algorithm based on HVS has been proposed. First, Top-hat and a threshold operation based on mechanism of visual attention are used to choose the most significant areas. Afterwards, confidence

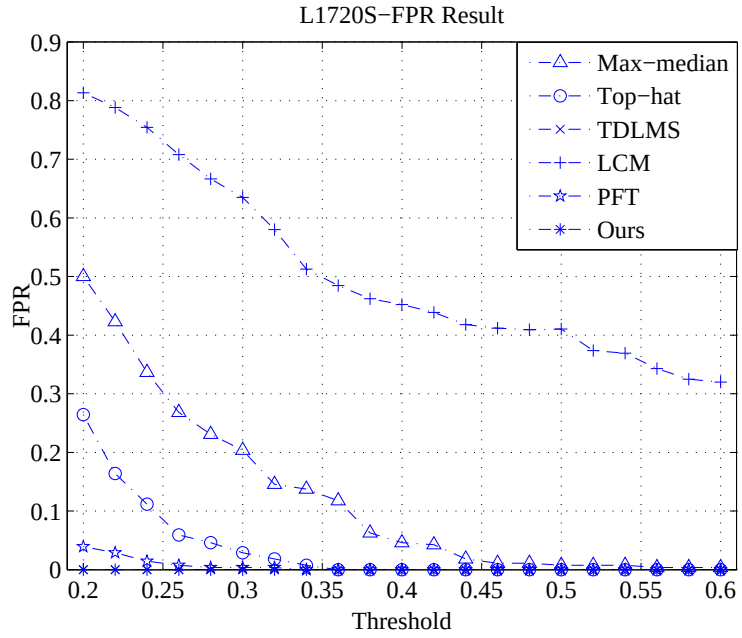


FIGURE 8. FPR of all method in images of L1720S.

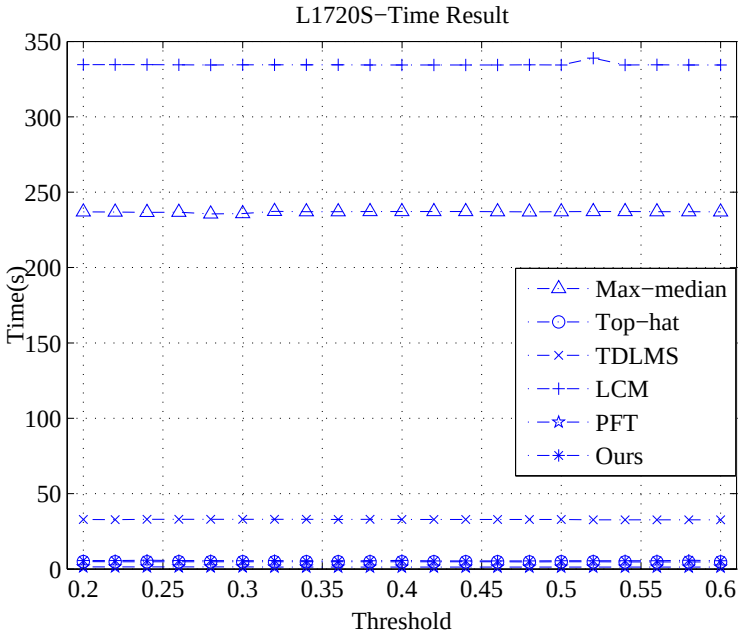


FIGURE 9. Time of all method in images of L1720S.

measurement is applied to separate targets from background clutters. Experimental results show that the proposed algorithm is robust to resist pseudo targets and can achieve a high detection rate in less than 0.02s with a fast calculation speed. It is worth noting that the FPR of proposed algorithm is far below other algorithms. This algorithm can be either directly used in single-frame target detection or used as a foundation module in sequential target tracking for real-time applications.

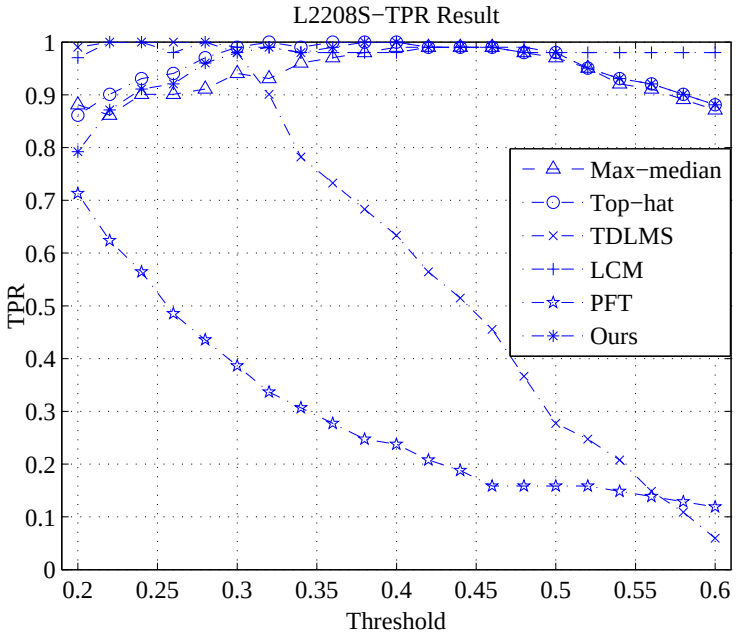


FIGURE 10. TPR of all method in images of L2208S.

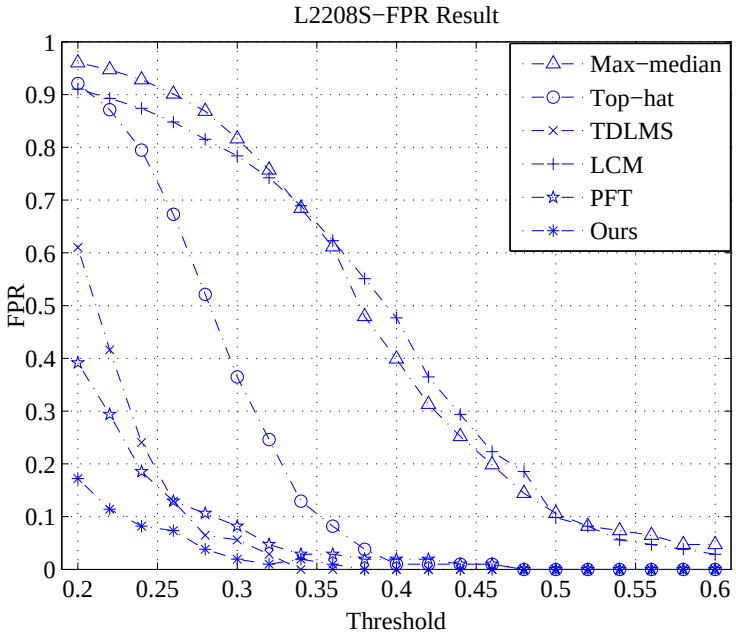


FIGURE 11. FPR of all method in images of L2208S.

Acknowledgment. The work was supported by the National Natural Science Foundation of China (Grant No. 61301207). Prof. Zhen Sun is the corresponding author.

REFERENCES

[1] G. Wenpu, S. Jiyin, L. Hao, Dynamic Template Preparing and Application on FLIR Target Recognition. *Electrical and Control Engineering (ICECE), 2010 International Conference on*, pp.864-867, 2010.

[2] Leonov, Sergei, Nonparametric method for clutter removal. *Aerospace and Electronic Systems, IEEE Transactions on*, vol.37, no.3, pp.832-848, 2001.

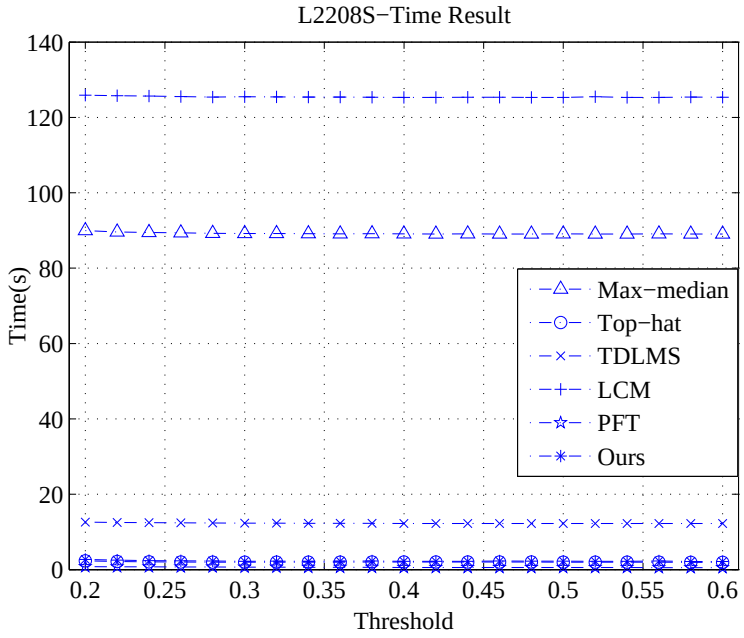


FIGURE 12. Time of all method in images of L2208S.

TABLE 1. TPR, FPR and Time results among different methods in L1618S (k =0.44)

Method	TPR	FPR	Time(s)
Max-median	1	0.2035	235.7630
Top-hat	1	0.0288	4.7890
TDLMS	0.8407	0	32.9630
LCM	0.9630	0.6351	334.4490
PFT	0.8593	0.0037	1.3730
Our method	1	0	5.5380

TABLE 2. TPR, FPR and Time results among different methods in L1720S (k =0.44)

Method	TPR	FPR	Time(s)
Max-median	1	0.2734	89.2470
Top-hat	1	0.5738	1.8720
TDLMS	0.9802	0.2734	12.4330
LCM	1	0	123.8640
PFT	0.0099	0	0.4990
Our method	0.9802	0.194	2.2150

[3] S. D. Deshpande, H. E. Meng, R. Venkateswarlu, P. Chan, Max-mean and max-median filters for detection of small-targets. *SPIE's International Symposium on Optical Science, Engineering, and Instrumentation*, International Society for Optics and Photonics, pp.74-83, 1999.

[4] P. Ffrench, J. R. Zeidler,W. H. Ku, Enhanced detectability of small objects in correlated clutter using an improved 2-D adaptive lattice algorithm. *Image Processing, IEEE Transactions on*, vol.6, no.3, pp.383-397, 1997.

TABLE 3. TPR, FPR and Time results among different methods in L2208S (k =0.44)

Method	TPR	FPR	Time(s)
Max-median	0.9406	0.8160	89.1850
Top-hat	0.9901	0.3648	1.9970
TDLMS	0.9802	0.0561	12.3550
LCM	0.9901	0.7837	125.4710
PFT	0.3861	0.0818	0.6550
Our method	0.9802	0.0194	2.2000

[5] T. W. Bae, Y. C. Kim, S. H. Ahn, K. I. Sohng, An efficient two-dimensional least mean square (TDLMS) based on block statistics for small target detection. *Journal of Infrared, Millimeter, and Terahertz Waves*, vol.30, no.10, pp.1092-1101, 2009.

[6] T. Arodz, M. Kurdziel, T. J. Popiela, E. OD. Sevre , D. A. Yuen, Detection of clustered micro calcifications in small field digital mammography. *computer methods and programs in biomedicine*, vol.81, no.1, pp.56-65, 2006.

[7] M. Zeng, J.Li, Z. Peng, The design of top-hat morphological filter and application to infrared target detection. *Infrared Physics & Technology*, vol.48, no.1, pp.67-76, 2006.

[8] C. Guo, Q. Ma, L.Zhang, Spatio-temporal saliency detection using phase spectrum of quaternion fourier transform. *Computer vision and pattern recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1-8, 2008.

[9] C. Q. Gao, D. Y. Meng, Y. Yang, Y. T. Wang, Infrared patch-image model for small target detection in a single image. *Image Processing, IEEE Transactions on*, vol.22, no.12, pp.4996-5009, 2013.

[10] J. Li, M. D. Levine, X. An, X. Xu, H. He, Visual saliency based on scale-space analysis in the frequency domain. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol.35, no.1, pp.996-1010, 2013.

[11] H. Li, Y. H. Tan, Y. S. Li,J. W. Tian, Image layering based small infrared target detection method. *Electronics Letters*, vol.50, no.1, pp.42-44, 2014.