Multiple Feature Analysis for Infrared Small Target Detection

Yanguang Bi, Xiangzhi Bai, Ting Jin, and Sheng Guo

***Abstract*— Detection of small target has been an important and challenging task in infrared systems. Most detection algorithms which only use single metric are difficult to separate target from clutter completely. The false alarm may be high when there exists complex backgrounds. In this letter, multiple novel features are proposed from four aspects to establish elaborate description. Each feature reflects specific characteristic of small target. The best feature vector is selected to apply these features for detection. Then, learning-based classifier is trained to screen candidate targets which are obtained by initial segmentation. Experimental results demonstrate that the proposed features could discriminate small targets from various clutters effectively. The better detection performance is achieved compared with other methods in different infrared backgrounds.**

***Index Terms*— Feature extraction, infrared background, learn-based classifier, small target detection.**

I. Introduction

INFRARED systems have applied to various detection fields such as surveillance and remote sensing [1], [2]. The early detection of small targets like antiship skimming missiles, surface ships, and low-flying aircraft is of great importance. Due to long imaging distance, the target tends to be dim and negligibly small [3]. It is difficult to use structure and texture information in such low signal-to-noise ratio and complex environments. In addition, many interferences include cloud edge, sky–sea line, different weathers, and sensor noises could cause high false alarms [4].

Over the past few decades, many algorithms have been designed to detect small target. Filter-based algorithms view target as a raised spot and construct specific filters to eliminate or highlight them. Conventional methods include max–mean, max–median [5], top-hat [6], [7], facet model [8], and kernel-based regression [9]. However, most of them need priori information about target or would be affected by heavy noise clutters.

Human visual system based algorithms highlight the saliency of target by constructing specific metrics. Difference of Gaussian (DoG) is one frequently used tool [10]. To alleviate the interference of edges, an improved difference of Gabor was proposed by incorporating multiple direction information [11]. Other similar methods include local contrast measure (LCM) [12], weighted local difference measure [13], and directional saliency-based measure [14].

Considering that single metric has limited description ability, learning-based algorithms detect targets via multiple aspects. Some common features in different methods involve the phase spectrum of Fourier transform [15], the histogram of oriented gradients (HOGs) [16], and relevant manual crafted features [17], [18].

To improve the performance of classification and detection, we analyze multiple characteristics of target and background from four aspects and propose novel and more effective features. Our contributions can be concluded as follows.

1) Seven novel features are proposed to distinguish small target from backgrounds. Direction, scale, gray-level intensity, spatial distribution, and variation are considered to improve the classification ability under complex clutters.
2) To apply proposed features for detection, the best feature vector is selected. The detection problem is converted to classification problem.

This letter is organized as follows. In Section II, we analyze the seven features and use them for detection. In Section III, the experimental results of feature processing and detection are presented. Finally, we conclude this letter in Section IV.

II. Analysis of Features and Target Detection

As shown in Fig. 1, the local area in the following analysis is divided into three areas according to the relative location of pixels. The three areas are center area, boundary area, and surround area. The surround area is divided into eight sections.

A. Gray-Level-Based Features

1) Center-Surround Contrast: The maximum value in center area is chose as $I_{\max}$ and the contrast C_i of the ith section in surround area is calculated by

$$C_i = \frac{1}{n}\sum_{j=1}^{n}(I_{\max} - (S_i(j)) \tag{1}$$

where $S_i(j)$ denotes the gray level of the jth pixel in the ith section in surround area which has n pixels totally. The minimum contrast among eight sections is selected as the Center-Surround Contrast (CSC). In this way, most edge regions are suppressed.

$$\mathrm{CSC} = \min_{i=1,2,\ldots,8} C_i. \tag{2}$$

For small target, $I_{\max}$ is usually higher than surround area so the contrast is larger than background. Although small targets may have different sizes, $I_{\max}$ could maintain constant and thus achieve multiscale detection.

Manuscript received November 20, 2016; revised February 14, 2017 and March 24, 2017; accepted May 24, 2017. Date of publication July 3, 2017; date of current version July 20, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61271023, in part by the Program for New Century Excellent Talents in Universities under Grant NCET-13-0020, and in part by the Fundamental Research Funds for the Central Universities under Grant YWF-17-BJ-Y-69. *(Corresponding author: Xiangzhi Bai.)*

Y. Bi and S. Guo are with the Image Processing Center, Beijing University of Aeronautics and Astronautics, Beijing 100191, China.

X. Bai is with the Image Processing Center, Beijing University of Aeronautics and Astronautics, Beijing 100191, China and also with the State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China (e-mail: jackybxz@buaa.edu.cn).

T. Jin is with the Beijing Institute of Spacecraft System Engineering, Beijing 100086, China.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2017.2711047

1545-598X © 2017 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: McMaster University. Downloaded on March 04,2020 at 16:51:49 UTC from IEEE Xplore. Restrictions apply.

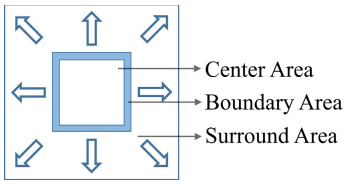

Fig. 1. Division of small neighboring area.

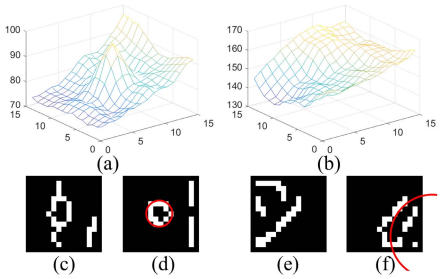

Fig. 2. (a) Small target. (b) Background. (c) Edge extracted from (a) directly and (d) after top hat. (e) Edge extracted from (b) directly and (f) after top hat.

2) Center-Fitting Residual: The bivariate quadratic function in

$$z(x, y) = ax^2 + by^2 + cxy + dx + ey + f \tag{3}$$

is used to fit the underlying gray-level surface in center area using the pixels of surround area. $z(x, y)$ denotes the fitting pixel gray level in coordination (x, y). a–f are coefficients deduced only by surround pixels using the least-square algorithm.

To embody the protruding attribute of small target in small neighboring area, the residual information in the sth scale is given using r_s as follows:

$$r_s = \frac{1}{n}\sum_{j=1}^{n}(I(x_j, y_j) - z(x_j, y_j)) \tag{4}$$

where $I(x_j, y_j)$ denotes the original gray level of the jth pixel in center area. This area has n pixels totally. Multiple scales are considered by changing the size of the center area to adapt the small targets with different sizes:

$$\text{CFR} = \max_{s=1,2,\ldots}(r_s). \tag{5}$$

The maximum r_s of all scales is chose as the center-fitting residual (CFR) expressed in (5). Due to the discontinuity of small target, the fitting residual should be larger than backgrounds.

In terms of the actual size of small target, three scales are considered in the entire features analysis. The corresponding sizes of the center area are 3×3, 5×5, and 7×7, respectively.

B. Edge-Based Features

As shown in Fig. 2, the distribution of small target is similar to the Gaussian model in shape so the edge usually resembles to circle. By contrast, the edge of background tends to be irregular.

Top-hat transformation is performed to avoid the influence of background before canny edge detection [19]. The edge which is the nearest to the region center is selected to fit a circle, i.e., the red curves in Fig. 2, using the least-square algorithm.

1) Radius: The fitting circle of target tends to have stable and small radius whereas the fitting circle of background has arbitrary size and location. Therefore, the radius is taken as the third feature.

2) Bias of Center: The fitting circle of target is usually located in region center. The distance between the circle center and region center is thus much smaller than background.

3) Variance of Distance With Center: The radius and center cannot describe a circle adequately when the circle is fit using the least-square algorithm. Thus, the variance of distances between each edge pixel with the fitting circle center is calculated as complement.

For target, the distance between each edge pixel with fitting circle center tends to be similar which leads to minor variance. However, the situation is opposite for background. Therefore, this variance is chose as the fifth feature.

C. Information Entropy-Based Feature

1) Benchmark Information Entropy Contrast: The mean gray level of boundary area is calculated as the benchmark coefficient I_b which is added into the information entropy E formula as follows:

$$E = -\frac{1}{n}\sum_{i=1}^{l} p_i \log_2 p_i (I_i - I_b) \tag{6}$$

where p_i denotes the probability of the ith gray level I_i when there are l gray levels and n pixels totally.

The target regions usually have higher information entropy because gray-level changes dramatically in a small area. The modified entropy in (6) of target could be enhanced with the difference between each gray level with I_b. For background, the gray levels which are higher and lower than I_b could offset each other so the metric is suppressed.

In the sth scale, the contrast EC_s is given by subtracting the modified entropy of surround area from center area:

$$\text{EC}_s = E_{\text{center}} - E_{\text{surround}}. \tag{7}$$

EC is calculated in different scales by changing the size of center area. The maximum is selected as the final contrast:

$$\text{EC} = \max_{s=1,2,\ldots} \text{EC}_s. \tag{8}$$

D. Texture-Based Feature

The angular second moment (ASM) [19] in gray-level co-occurrence matrix (GLCM) [19] is frequently used in texture analysis and could be understood as the texture energy. Large ASM indicates the gray-level changes smoothly. Thus. the ASM of targets tend to be small compared with background.

1) Texture Energy Contrast: For entire region, GLCM is calculated evenly from four directions. The minimum ASM extracted from four matrixes is taken as $\text{ASM}_{\text{entire}}$ which denotes the most drastic direction. For center area, GLCM is calculated similarly but the maximum ASM which means the smoothest direction is selected as $\text{ASM}_{\text{center}}$. Texture energy contrast (TEC) in the sth scale is

$$\text{TEC}_s = \frac{1 - \text{ASM}_{\text{center}}}{1 - \text{ASM}_{\text{entire}}} \tag{9}$$

Authorized licensed use limited to: McMaster University. Downloaded on March 04,2020 at 16:51:49 UTC from IEEE Xplore. Restrictions apply.

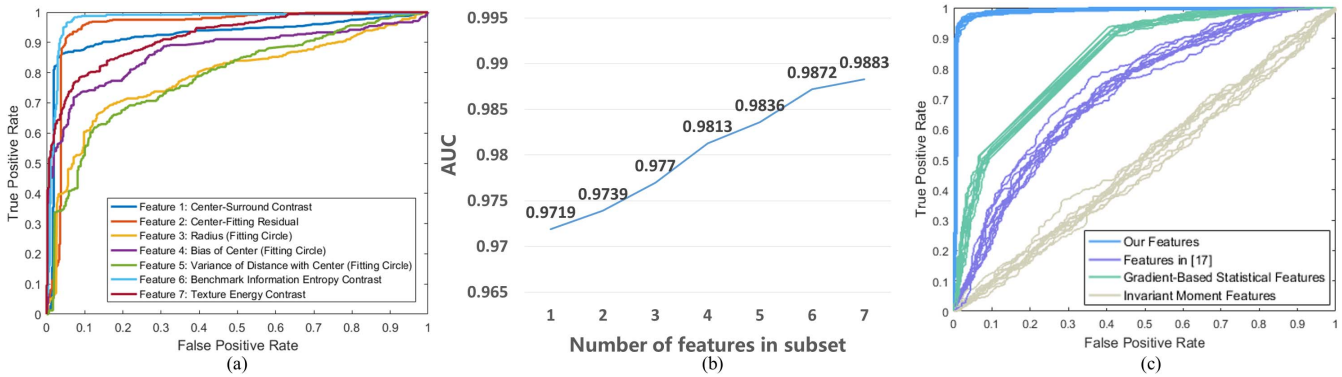

Fig. 3. Feature processing. (a) ROC curves for each proposed feature. (b) AUC of feature subset during selection. (c) Comparison with other features.

and the maximum TEC_S is chose as the last feature. Edge clutters are suppressed in this way but target would not be affected.

$$\mathrm{TEC} = \max_{s=1,2,\ldots} \mathrm{TEC}_s. \tag{10}$$

This contrast could be considered as the proportion of center area texture energy in entire region texture energy. For target, the proportion is usually larger than most backgrounds.

E. Application for Detection

Several basic tools like canny edge and GLCM are used as preprocessing operations during the analysis. Based on these primitive operations and small target properties, new and more advanced features are redesigned by incorporating direction and scale information. With top-hat transformation and support vector machine [19], the multiple features could thus be used to classify the candidate targets.

The real small target may occupy only a few pixels in an image thus the background is the majority which may affect detection performance. In fact, these features could effectively enhance small targets and suppress backgrounds. Specifically, gray-level-based features enhance target according to its high gray intensity and suppress edge or slope clutters which usually appear in heavy cloud scenes. Edge-based features remove common nonuniformity and noises which have similar high gray intensity but irregular shapes. Through the variation trend of gray level, information entropy and texture-based features enhance target and exclude incidental glints which could occur in artificial structures and sea clutters. Moreover, top-hat transformation is performed first to remove most background regions and obtain candidate targets. Therefore, the imbalance in classification could be relieved.

There may exist redundant and irrelevant features, selection is performed before detection to check whether each feature is indispensable. To balance practical application and efficiency, wrapper-based methods [20] and sequential forward search [21] are taken. First, the AUC (area under receiver operating characteristic (ROC) curve) metric of all features are calculated and sort. Then, put the best feature which has the highest AUC value into subset and evaluate the subset. If the AUC of subset increases, it indicates that this feature is valuable and should be retained, otherwise abandoned. Check all features in this way and the final feature subset is obtained.

Ten-fold cross validation is applied to evaluate features. The total data set is divided into ten exclusive subsets. Every subset is chose to evaluate the model trained using other nine subsets and the average AUC is calculated as final evaluation result.

TABLE I
MEAN AND STD OF TEN AUC METRICS FOR EACH FEATURE

Feature	I	II	III	IV	V	VI	VII
Mean	0.922	0.950	0.781	0.852	0.789	0.972	0.909
STD	0.006	0.006	0.014	0.013	0.010	0.006	0.005

III. EXPERIMENTAL RESULTS

A. Samples Collection

The experimental images, include cloud-sky, ground-sky, and cumulus scenes, are provided by the project partners. The wavelength ranges from 5 to 12 μm. The positive samples have target existing in center area and the negative samples are collected from background directly. The size of each sample is 15×15.

B. Feature Evaluation and Selection

The ROC curves of seven features in one validation are plotted in Fig. 3(a). According to the ten-fold cross validation, the mean and standard deviation (STD) of ten AUC metrics for each feature are listed in Table I. The small STD values indicate stable validation process and credible evaluation results. The mean values are final evaluation results and the higher mean AUC value illustrates better classification ability. Edge-based features, which are sensitive to complex background, have lower mean AUC values ($<$0.9). Other features are relatively satisfactory. The mean AUC value of feature VI is the highest (0.972), so this feature is the best.

The AUC metric of feature subset during selection is shown in Fig. 3(b). Beginning with feature VI, the evaluation of subset is persistently increasing when it contains more features. The highest AUC is achieved when all the features are retained. This selection result also validates the necessity of each feature.

The subset is evaluated only seven times using sequential forward search. Considering that every feature contributes to classification, the time consumption of evaluation is decreasing as more features are put into subset. Therefore, the entire time consumption is acceptable. The selection aims to justify the necessity of each feature so it is only performed once.

C. Feature Comparison

Invariant moment features, gradient-based statistical features and manual crafted features [17] are evaluated as comparisons. Different from classical features like HOG [16]

Authorized licensed use limited to: McMaster University. Downloaded on March 04,2020 at 16:51:49 UTC from IEEE Xplore. Restrictions apply.

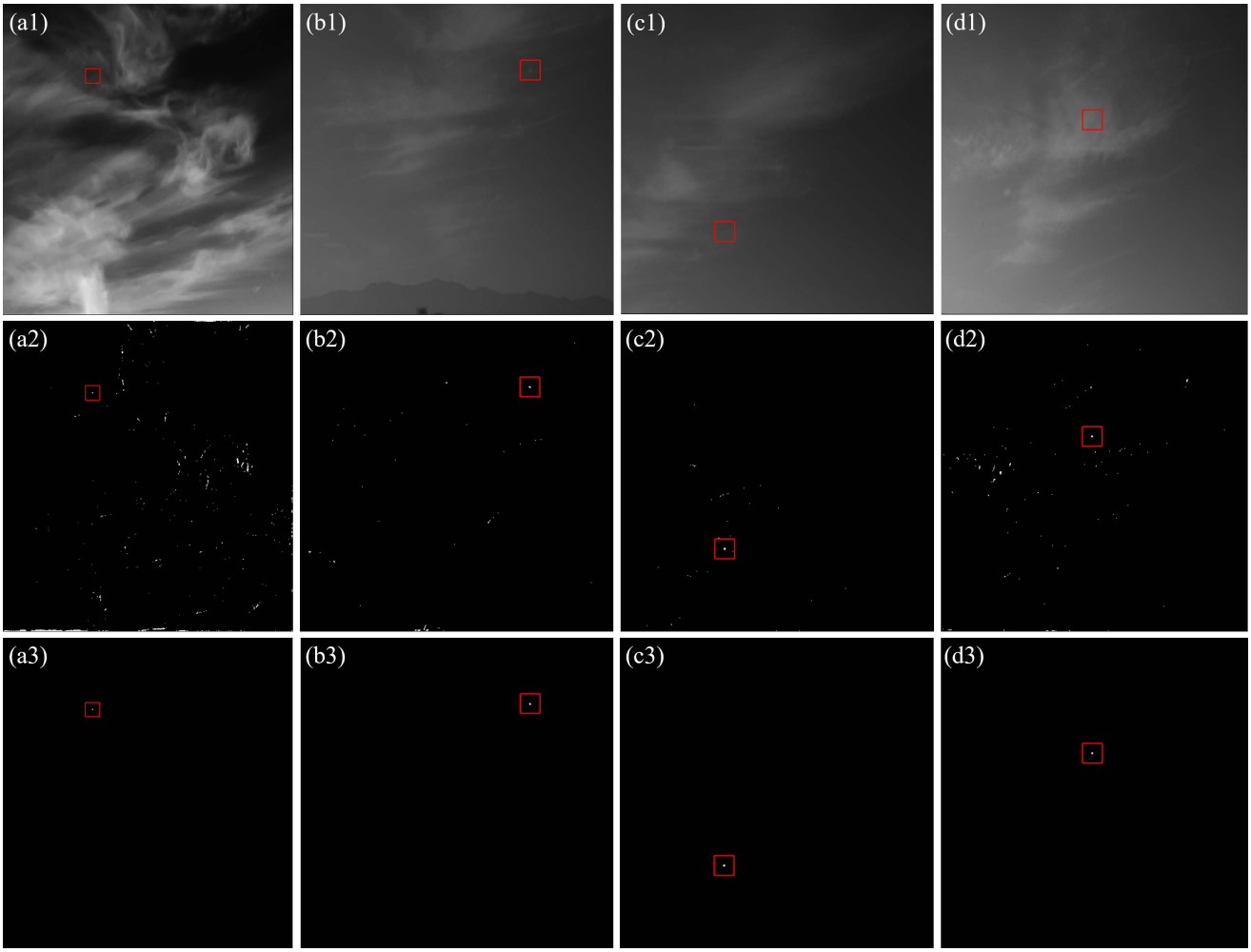

Fig. 4. (a1)–(d1) Original images, (a2)–(d2) segmentation results from top-hat transformation, and (a3)–(d3) final detection results after classification. The sizes of original infrared images are 623×667, 491×487, 491×487, and 481×487.

which has high dimensions compared with samples, gradient-based feature just consists of weighted gradients from eight directions evenly.

As shown in Fig. 3(c), invariant moment features locate in the diagonal which means the worst classification performance. It is difficult to use simple geometrical features for establishing effective description due to the less pixels and irregular shape of small targets. The gradient-based features have certain classification ability because the statistical weighted gradient of small target shows robustness. Features in [17] do not show expected performance. It indicates that simple statistic features like standard derivation cannot been used directly. The ROC curve of our features is located in the upper-left corner which means the best classification performance. The consideration of multiple aspects improves discrimination ability effectively.

D. Detection Result

The detection experiments are performed on different infrared images. Original infrared images are shown in Fig. 4(a1)–(d1), respectively; Fig. 4(a2)–(d2) denote the segmentation results from top-hat transformation and the threshold T is calculated in

$$T = \bar{TH} + k \times \sigma_{\mathrm{TH}}. \tag{11}$$

Fig. 4(a3)–(d3) denote the final detection results after classification. The real small targets are marked with red rectangles in all images.

In (11), TH is the result of top-hat transformation; $\bar{TH}$, σ_{TH}, and k denote the mean value, the STD, and an adjustable parameter, respectively. This equation is usually used to choose threshold automatically in different scenes [12]. The optimal selection of k is from statistical experiment results and can be set to 5 in most situations.

It could be observed that the results segmented from top-hat transformation tend to have much false alarms. Most of them are complex edges and corner regions. When classification is applied, the real targets are retained whereas the false alarms decrease obviously and are even eliminated. It proves that our features could separate small target from background robustly and the classification works well in detection.

The proposed features are deduced based on the approximate Gaussian distribution property of infrared small target. This detection method could also work in other wavelength images if the approximate Gaussian distribution condition of the target region is satisfied.

E. Comparison With Other Detection Methods

Conventional methods such as accumulated center-surround difference measure (ACSDM) [4], max-mean,

Authorized licensed use limited to: McMaster University. Downloaded on March 04,2020 at 16:51:49 UTC from IEEE Xplore. Restrictions apply.

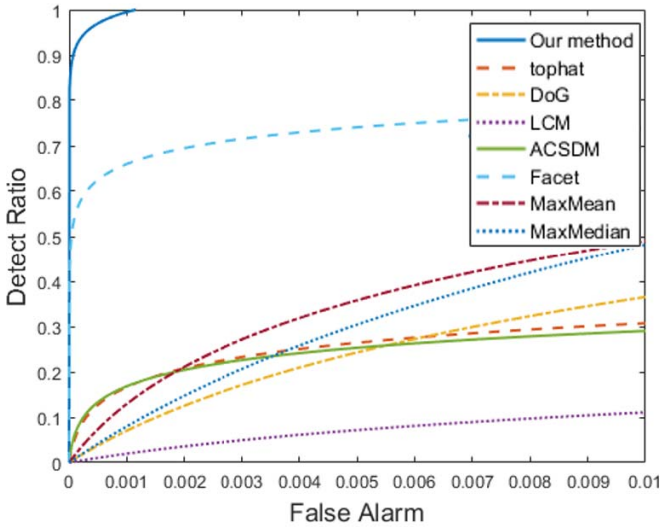

Fig. 5. ROC curves of eight methods on test images.

max-median [5], facet model [8], top-hat [6], DoG [10], and LCM [12] are also performed as comparisons. When the segmentation threshold T in (11) changes, the ROC curves are plotted in Fig. 5. The false alarm is defined in

$$\text{false alarm} = \frac{\text{number of false pixels detected}}{\text{number of total pixels in image}}. \quad (12)$$

The detection ratio is defined in

$$\text{detection ratio} = \frac{\text{number of true pixles detected}}{\text{number of total true pixels detected}} \quad (13)$$

and normalized to [0, 1]. Furthermore, these smooth curves are fit synthetically using the entire detection results of images in Fig. 4.

During the detection, DoG and LCM do not obtain expected results considering the magnitude of target is negligibly small which affects the target enhancement. The detection results of max–mean, max–median, and ACSDM are well only when the size of filters is similar to small target. Facet model shows good detection ability but false alarms still exist. The result of top hat is unsatisfactory because the fixed structuring elements cannot adapt to various backgrounds. After classification, the false alarm is reduced significantly and the real targets are maintained. Our method thus achieves the best detection results. This comparison indicates that multiple features could establish more comprehensive description which contributes to better detection performance than ordinary methods.

IV. Conclusion

In this letter, seven features are proposed from four aspects for small target detection. Features I and II describe target from the spatial distribution of gray level to highlight the center area. Features III–V analyze the edge characteristics which reflect the location and gradient information of target. Features VI and VII embody the variation degree of target via entropy and texture. The consideration of direction and scale makes features more effective and robust. The ROC curves show that the proposed feature vector has better classification ability than other features. With the contribution of multiple features, the detection results achieve high detection ratio (>0.9) but low false alarm (<0.001) compared with other methods. The proposed method could also be implemented effectively under various scenes like noisy cloud sky, cumulus, or mountain edges.

References

[1] M. Malanowski and K. Kulpa, "Detection of moving targets with continuous-wave noise radar: Theory and measurements," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 9, pp. 3502–3509, Sep. 2012.

[2] C.-I. Chang and D. Heinz, "Constrained subpixel target detection for remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 3, pp. 1144–1159, May 2000.

[3] A. Dawoud, M. S. Alam, A. Bal, and C. H. C. Loo, "Decision fusion algorithm for target tracking in infrared imagery," *Opt. Eng.*, vol. 44, no. 2, p. 026401, Feb. 2005.

[4] K. Xie, K. Fu, T. Zhou, J. Zhang, J. Yang, and Q. Wu, "Small target detection based on accumulated center-surround difference measure," *Infr. Phys. Technol.*, vol. 67, pp. 229–236, Nov. 2014.

[5] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," in *Proc. SPIE*, vol. 3809. Oct. 1999, pp. 74–83.

[6] X. Bai and F. Zhou, "Infrared small target enhancement and detection based on modified top-hat transformations," *Comput. Elect. Eng.*, vol. 36, no. 6, pp. 1193–1201, Nov. 2010.

[7] V. T. Tom, T. Peli, M. Leung, and J. E. Bondaryk, "Morphology-based algorithm for point target detection in infrared backgrounds," in *Proc. SPIE*, vol. 1954. Oct. 1993, pp. 2–11.

[8] G.-D. Wang, C.-Y. Chen, and X.-B. Shen, "Facet-based infrared small target detection method," *Electron. Lett.*, vol. 41, no. 22, pp. 1244–1246, Oct. 2005.

[9] Y. Gu, C. Wang, B. Liu, and Y. Zhang, "A kernel-based nonparametric regression method for clutter removal in infrared small-target detection applications," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 3, pp. 469–473, Jul. 2010.

[10] X. Dong, X. Huang, Y. Zheng, L. Shen, and S. Bai, "Infrared dim and small target detecting and tracking method inspired by human visual system," *Infr. Phys. Technol.*, vol. 62, pp. 100–109, Jan. 2014.

[11] J. Han, Y. Ma, J. Huang, X. Mei, and J. Ma, "An infrared small target detecting algorithm based on human visual system," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 452–456, Mar. 2016.

[12] C. L. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2014.

[13] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Small infrared target detection based on weighted local difference measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 4204–4214, Jul. 2016.

[14] S. Qi, J. Ma, C. Tao, C. Yang, and J. Tian, "A robust directional saliency-based method for infrared small-target detection under various complex backgrounds," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 3, pp. 495–499, May 2013.

[15] Z. Cui, J. Yang, S. Jiang, and C. Wei, "Target detection algorithm based on two layers human visual system," *Algorithms*, vol. 8, no. 3, pp. 541–551, Jul. 2015.

[16] C. W. Miller, J. A. Edelberg, M. L. Wilson, and K. Novak, "Application of rich feature descriptors to small target detection in wide-area persistent ISR systems," in *Proc. SPIE*, vol. 9092. Jun. 2014, p. 909202-1–909202-7.

[17] S. Kim, "Analysis of small infrared target features and learning-based false detection removal for infrared search and track," *Pattern Anal. Appl.*, vol. 17, no. 4, pp. 883–900, Nov. 2014.

[18] M. S. Alam, E.-H. Horache, and S. F. Goh, "Performance evaluation for cluttered infrared image based on fringe-adjusted joint transform correlator," in *Proc. SPIE*, vol. 5437. Apr. 2004, pp. 63–74.

[19] E. R. Davies, *Machine Vision: Theory, Algorithms, Practicalities*, 3nd ed. New York, NY, USA: Academic, 2004.

[20] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, nos. 1–2, pp. 273–324, 1997.

[21] M. Dash and H. Liu, "Feature selection for classification," *Intell. Data Anal.*, vol. 1, nos. 1–4, pp. 131–156, 1997.

Authorized licensed use limited to: McMaster University. Downloaded on March 04,2020 at 16:51:49 UTC from IEEE Xplore. Restrictions apply.