Received February 28, 2019, accepted March 10, 2019, date of current version April 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2906076

Infrared Small Target Detection Through Multiple Feature Analysis Based on Visual Saliency

YUWEN CHEN[1], BIN SONG[1], (Senior Member, IEEE), XIAOJIANG DU[2], (Senior Member, IEEE), AND MOHSEN GUIZANI[3], (Fellow, IEEE)

[1]State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710071, China
[2]Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122, USA
[3]Department of Computer Science and Engineering, Qatar University, Doha, Qatar

Corresponding author: Bin Song (bsong@mail.xidian.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61772387 and Grant 61802296, in part by the Fundamental Research Funds of Ministry of Education and China Mobile under Grant MCM20170202, and in part by the ISN State Key Laboratory.

ABSTRACT Infrared small target detection in extreme environments such as low illumination or complex background with low signal clutter ratio is of crucial significance and counted as a difficult task in infrared search and tracking systems. In this paper, an effective infrared small target detection method is proposed based on the human visual system characteristics and multiple feature analysis. By using the contrast mechanism and visual attention mechanism, spatial gray level-based feature map and saliency extraction based feature map from frequency domain are obtained. Based on the characteristics of infrared dim target images and human visual attention mechanism, the target saliency features are revealed through the feature analysis in the spatial domain and frequency domain respectively. The saliency features from each feature map are applied to the final saliency map. By this means, the background clutter and noise are inhibited and the targets are distinct for the various scenes in the infrared images. The experimental results show that the proposed method has a robust and effective performance in terms of detection and false alarm rates. Comparing with the other methods in the experiments, the proposed method is feasible and adaptable in the various scenes of infrared images.

INDEX TERMS Infrared image, small target detection, visual contrast mechanism, visual attention mechanism.

I. INTRODUCTION

It is well known that object detection is widely applied in various fields, such as aerospace, remote sensing, surveillance, security, healthcare and so on [1]–[7]. In the image or video, object detection such as face detection, color or shape detection, action detection, object of interest detection and so on, play crucial roles in numerous applications. In some fields, such as smart healthcare system, target detection and action detection are important research issues of the smart healthcare technologies for the old people, which is used to track the people and detect the falling action. The object of interest detection in the medical images also can assist to develop the healthcare system technologies. The face detection is widely used in the security system. The emotion awareness and authentication based on the face detection are popular research issues for the multimedia system. Dim and small target detection in infrared (IR) images is one of the crucial techniques in the practical projects, such as precise guidance, pre-warning, video surveillance, search and tracking system, and so on [8]–[10]. Generally, the IR target contains a little pixels and presents as spot-like feature [1], which is submerged in the background easily. In the infrared search and tracking system, unconstrained and extreme environments such as low illumination, unknown background clutter and noise introduced by the sensor system and natural factors make the discovery of the IR small target hard. Therefore, the IR small target detection is difficult and challenging in the practical applications.

Over the past few decades, many typical algorithms are proposed and applied in the IR small target detection, such as wavelet transform algorithms [9], facet model [11], morphology method [12], sparse matrix and low-rank matrix decomposition method [13], and so on. In recent years,

The associate editor coordinating the review of this manuscript and approving it for publication was Huimin Lu.

human visual system characteristics are introduced in the IR small target detection and have superior performance. The properties of human visual system (HVS) such as contrast mechanism, multi-resolution, size adaptation, pop-out phenomena, and so on, can benefit the object detection in image processing. Human vision has selective attention property, and the salient regions attract attention of human quickly in the complex background. As a typical representation of the visual attention, saliency map is now widely used in many computer vision applications such as image segmentation, object detection and recognition [14]–[17]. Qi *et al.* [18] proposed a fast saliency method which can detect small salient regions well. Based on gradient enhancement operation combined with Gaussian smoothing, a facet kernel operator can highlight the target and suppress the background. Huang *et al.* [19] introduced the directional difference of Gaussian filter to cope with complicated background because the directional difference of Gaussian can restrain the irregular edges with fan like shapes. The salient regions such as the small targets in the image can be remained. Han *et al.* [17] utilized the multiscale relative local contrast measure which can make the salient regions have significant contrast and deal with different sizes of small targets. In this way, the target is enhanced and the background is suppressed. However, some of these methods cannot effectively enhance the small target in the complex background. The bright noise in the background would be remained by some methods. The parameters in some methods would not be chosen appropriately for various image scenes.

In this paper, an IR small target detection method is proposed through constructing the multiple feature maps, which combines the spatial saliency characteristics of spatial domain with extracted saliency information of frequency domain. As the infrared target image presents different saliency features in different image domains. The saliency characteristics of the IR images are analyzed from the spatial and frequency domain respectively. The spatial gray level based feature map and the saliency extraction based feature map are designed respectively. A spatial gray level based feature map is based on the HVS contrast mechanism. The map is obtained by the modified morphology filters and weighted gray map, which contains the multiple orientations and scales for removing the background and enhancing the target respectively. By analyzing frequency characteristics of the IR target images, the other saliency extraction based feature map is computed by the difference of bilateral filters to extract various frequency components. The multiple features from different saliency maps present the saliency information of IR image from different views which can help improve the signal-to-clutter ratio of IR images. The multiple features from the different saliency maps are integrated in the final stage to benefit the target enhancement and the background inhibition. The final enhanced saliency map can have the robust ability to separate the IR small target from the jamming objects.

The rest of paper is organized as follows. In Section 2, spatial gray level based feature map and multi-scale saliency based feature map are described, including the morphological filter in multi-direction, multi-scale weighted gray map and saliency extraction. Experimental results are presented in Section 3, and the conclusion is in Section 4.

II. TARGET DETECTION BASED ON THE MULTIPLE FEATURES

The IR small target contains a few pixels in the IR image and is easily submerged in the background clutter because of the long distance between the targets and sensors. Usually, the IR small target has high intensity than that of the surrounding background and like a spot in the image. In the raw IR images, the small targets are defined to have a total spatial extent of less than 80 pixels [1]. Usually, a raw IR image $F(x,y)$ can be formulated as:

$$F(x, y) = F_T(x, y) + F_B(x, y) + F_N(x, y) \tag{1}$$

where, (x, y) is the coordinate of the image pixel, $F_T(x, y)$ is the gray value of target pixel, $F_B(x, y)$ and $F_N(x, y)$ are gray values of background and noise respectively. According to the HVS characteristics, the response of ganglion cell to contrast pattern is linear to contrast [19]. It means that the HVS can perceive the brightness of target according to the contrast between the target and background. The contrast between the intensity of target and background, instead of the amplitude of a visual target signal, can make a target or pixel stand out against its neighbors. Meanwhile, human visual attention has selective property for the salient region in the complex scene. Visual attention model can search the salient target efficiently without the background interference in the field of target detection and recognition. Based on the contrast mechanism and visual attention mechanism, spatial gray level based feature map and multi-scale saliency based feature map for improving the target signal-to-clutter rate and extracting the multi-scale saliency information are described in the follows.

A. MATHEMATICAL MORPHOLOGY THEORY

Mathematical morphology theory is deduced from the set theory and integral geometry, which is a vital theory in image processing and pattern recognition [21]. Reconstruction is a part of the mathematical morphology theory, and related to the two basic morphological operators: dilation ($\oplus$) and erosion ($\ominus$), which are defined as follows:

$$\begin{aligned}(F \oplus B)(x, y) = max\Big\{F\left(x - x', y - y'\right) \\ + B\left(x', y'\right) \mid \left(x', y'\right) \in D_B\Big\}\end{aligned} \tag{2}$$

$$\begin{aligned}(F \ominus B)(x, y) = min\Big\{F\left(x + x', y + y'\right) \\ - B\left(x', y'\right) \mid \left(x', y'\right) \in D_B\Big\}\end{aligned} \tag{3}$$

where, D_B is the domain of structuring element B. If the mask image is F, the marker image is F_m, the dilation result computed between F and F_m by structuring element B is denoted by $\mathcal{D}_F^{(1)}$.

$$\mathcal{D}_F^{(1)}(F_m) = \min((F_m \oplus B), F) \tag{4}$$

Iteratively, $\mathcal{D}_F^{(n)}$ is defined as follows:

$$\mathcal{D}_F^{(n)}(F_m) = \mathcal{D}_F^{(1)}\left(\mathcal{D}_F^{(n-1)}(F_m)\right) \tag{5}$$

where, $\mathcal{D}_F^{(0)}(F_m) = F_m$. Reconstruction by dilation is denoted as $\mathcal{R}_F^{\mathcal{D}}(F_m) = \mathcal{D}_F^{(k)}(F_m)$. It is a stable value of $\mathcal{R}_F^{\mathcal{D}}(F_m)$ when $\mathcal{D}_F^{(k)}(F_m) = \mathcal{D}_F^{(k+1)}(F_m)$. Opening by reconstruction $\mathcal{O}_R$ is the reconstruction by dilation of F_m which is the erosion result of F by using the structuring element C.

$$\mathcal{O}_R(F) = \mathcal{R}_F^{\mathcal{D}}(F_m) = \mathcal{R}_F^{\mathcal{D}}(F \ominus C) \tag{6}$$

So the white tophat-by-reconstruction operator $RWTH(F)$ is defined as:

$$RWTH(F) = F - \mathcal{O}_R(F) \tag{7}$$

Similarly, reconstruction by erosion is denoted by $\mathrm{R}_F^{\mathrm{E}}$, which is defined as follows:

$$\varepsilon_F^{(1)}(F_m) = \max((F_m \ominus B), F) \tag{8}$$

$$\varepsilon_F^{(n)}(F_m) = \varepsilon_F^{(1)}\left(\varepsilon_F^{(n-1)}(F_m)\right), \quad \text{where, } \varepsilon_F^{(0)}(F_m) = F_m. \tag{9}$$

$$\mathcal{R}_F^{\varepsilon}(F_m) = \varepsilon_F^{(k)}(F_m), \quad \text{where, } \varepsilon_I^{(k)}(F_m) = \varepsilon_F^{(k+1)}(F_m) \tag{10}$$

Then closing by reconstruction is denoted by $\mathcal{C}_R$,

$$\mathcal{C}_R(F) = \mathcal{R}_F^{\mathcal{E}}(F_m) = \mathcal{R}_F^{\mathcal{E}}(F \oplus C) \tag{11}$$

The black tophat-by-reconstruction operator $RBTH(F)$ is defined as:

$$RBTH(F) = \mathcal{C}_R(F) - F \tag{12}$$

The local bright features can be smoothed through opening by reconstruction in the image, to the contrary, some dim features can be removed using closing by reconstruction. The RWTH and RBTH can obtain the local bright features and the dim features respectively. Considering the gray value of IR small target pixel is usually higher than that of its neighbor background, we discuss the work in the condition of bright IR target in detail.

B. SPATIAL GRAY LEVEL BASED FEATURE MAP

According to the definition above, the marker image F_m determines the region of a mask image F which is reconstructed or not. Therefore, the choice of the marker image plays an important role. To obtain the target region and remove the background by RWTH, the target pixels in F_m need to be eroded and the background regions need to be preserved as much as possible. Generally, F_m is the erosion result of F. The background clutter in the IR image has complicated edges and texture features. If we select a single rectangle or disk structuring element as usual, the result of suppressing the background would not be satisfied. So, we proposed a modified RWTH (MRWTH).

We introduce the directional operation to get a proper marked image. The multi-directional linear structuring elements are exploited, which can be sensitive to orientations for inhibiting irregular edges of background clutter. It is described as follows:

$$F_n = F \ominus C_n, \quad n = 1, 2, \ldots, N \tag{13}$$

where, $C_n(n = 1, 2, \ldots, N)$ are the linear structuring elements in each direction, which can erode target in multiple orientations. The result of erosion is selected as follows:

$$F_m(x, y) = \max(F_n(x, y)), \quad n = 1, 2, \ldots, N \tag{14}$$

In the experiments, we use eight directions and the interval of direction is $\pi/8$. Considering a real IR small target is usually less than 80 pixels, the length of linear structuring element is set as 9. The mask image F_m get from (14) is used in RWTH forming the MRWTH which can cope with the orientation information to suppress the irregular background edge. The contrast between the target and background can be improved by suppressing the background.

The local region which contains the small target pixels is usually brighter than the surrounded regions and not spatially correlated with its surroundings. The MRWTH can improve the contrast between the target and its surroundings by inhibiting the irregular background. In order to enlarge the contrast further through enhancing the target signal, a weighted gray map based on the MRWTH is proposed. A multi-scale binomial kernel induced from the Laplacian operator, which is as follows:

$$W_{t\times t} = \begin{bmatrix} -1 & -1 & \ldots & -1 & -1 \\ -1 & 0 & \ldots & 0 & -1 \\ \vdots & \vdots & 4n-4 & \vdots & \vdots \\ -1 & 0 & \ldots & 0 & -1 \\ -1 & -1 & \ldots & -1 & -1 \end{bmatrix}_{t\times t}, \quad t \geq 3 \tag{15}$$

The binomial kernel is considered as a weighted map to enhance the result of MRWTH, which is used to obtain the convolution with the MRWTH. It can be seen that $W_{3\times 3}$ is the original Laplacian operator. If the pixel is in the homogeneous background, the result of the convolution would be zero approximately. If the pixel in the salient region, the result of the convolution would be larger than zero because of the local salient region is discontinuous with its neighbor. So, the larger contrast between a pixel and its surroundings, the larger result of the kernel convolution. Then, the MRWTH and weighted gray template can suppress the background and strengthen the target simultaneously. Since the target size is unknown in practical applications and the real IR targets do not exceed the 80 pixels, the parameter t is odd and belongs to [3], [9]. The spatial gray level based feature map is computed as follows:

$$w_t = W_{t\times t} * MRWTH \tag{16}$$

$$S1 = \prod_t w_t \tag{17}$$

where $*$ is the convolution operation, *S1* is spatial gray level based feature map. The calculation of spatial gray level based feature map is in Algorithm 1.

Algorithm 1 The Spatial Gray Level Based Feature Map

Input: the original image F
Output: the spatial gray level based feature map S1
1.The number of orientations is l, the size of the image is $X \times Y$.
2.For n = 1 : l do

$$F_n = F \ominus C_n$$

3.End for
4.For x = 1:X do
5.For y = 1:Y do

$$F_m(x, y) = \max(F_n(x, y)), \quad n = 1, 2, \ldots, N$$

6.End for
7.End for

$$MRWTH(F) = F - \mathcal{R}_F^{\mathcal{D}}(F_m)$$
$$w_t = W_{t \times t} * MRWTH(F)$$
$$S1 = \prod_t w_t$$

C. THE BILATERAL FILTER THEORY

Gaussian kernel is commonly applied to reduce the noise and smooth the image. The multi-scale Gaussian kernels can present different scales saliency details in the linear scale space. Based on the spatial and range Gaussian kernels, the bilateral filter can preserve the saliency details in the spatial and range domains, which can be suitable for various image scenes. The pixel u of image f filtered by the bilateral filter is calculated as follows [22]:

$$f_b(u) = \frac{1}{H} \sum\nolimits_{v \in \mathrm{P}} h_{\sigma_s}(\|\text{u-v}\|)\, h_{\sigma_r}(|f(u) - f(v)|) f(v) \tag{18}$$

$$H = \sum\nolimits_{v \in \mathrm{P}} h_{\sigma_s}(\|\text{u-v}\|)\, h_{\sigma_r}(|f(u) - f(v)|) \tag{19}$$

where h is denoted as the Gaussian function, σ_s and σ_r are the standard deviations of the Gaussian functions, P is the local region domain.

D. SALIENCY EXTRACTION BASED FEATURE MAP

In a real IR image, the gray values of the background regions change slowly and have homogeneous distribution approximately. In the frequency domain, the background region is considered as low frequency part. On the contrary, the target region is not consistent with its neighbor regions and presents difference among its surroundings. The target signal has high frequency. In addition, some clutter from the sensors always has higher frequency. So the salient region which contains the small target can be extracted by suppressing the background and clutter in the frequency domain through band pass filters [15]. The difference of Gaussian is usually used to obtain the saliency map. As the different value of the standard deviations can smooth the fine scale details of the image saliency features in different degrees, we design the several band pass filters based on difference of bilateral (DoB) filters as follows:

$$\sum\nolimits_{n=1}^{N} DoB\left(u, \sigma_{s1}^n, \sigma_{r1}^n, \sigma_{s2}^n, \sigma_{r2}^n\right) = \sum\nolimits_{n=1}^{N} f_b(u, \sigma_{s1}^n, \sigma_{r1}^n) - f_b(u, \sigma_{s2}^n, \sigma_{r2}^n) \tag{20}$$

where N is the number of the DoB filters. The results of the DoB filters reveal the band frequency at several scales. If the σ_{s2}^n is equal to σ_{s1}^{n-1}, and σ_{r2}^n is equal to σ_{r1}^{n-1}, (20) is simplified as follows:

$$\sum\nolimits_{n=1}^{N} DoB\left(u, \sigma_{s1}^n, \sigma_{r1}^n, \sigma_{s2}^n, \sigma_{r2}^n\right) = f_b\left(u, \sigma_{s1}^N, \sigma_{r1}^N\right) - f_b(u, \sigma_{s2}^1, \sigma_{r2}^1) \tag{21}$$

It can be seen that the band pass filters based on the DoB can be constructed by the difference of the two bilateral filters, which benefit the preservation of the spatial frequency from the IR image for computing the saliency map. Here, the band pass filter based on the DoB is used to extract the saliency features for simplicity as follows:

$$D(u) = |f_b(u, \sigma_{s1}, \sigma_{r1}) - f_m(u)| \tag{22}$$

where f_m is based on the neighborhood average method to get the low cut-off frequency and the f_b gets the high cut-off frequency. D is the saliency extraction result. We set the values of σ_{s1} and σ_{r1} are 4 and 0.1 to extract the fine-scale saliency features in the simulation. The local region domain P is set as square regions with the side of nine pixels because the IR small targets do not exceed 80 pixels. The calculation of saliency extraction based feature map is the Algorithm 2.

Algorithm 2 Saliency Extraction Based Feature Map

Input: the original image F
Output: the saliency extraction based feature map S2
1.The number of pixels in image is U
2.For u = 1:U do

$$S2 = D(u) = |f_b(u, \sigma_{s1}, \sigma_{r1}) - f_m(u)|$$

3.End for

E. SALIENCY MAP ENHANCEMENT

The spatial gray level based feature map and saliency extraction based feature map are based on the contrast mechanism and visual attention mechanism respectively. The multiple features from the spatial and frequency domains would be complementary the final saliency map enhancement. The final saliency map is as follows:

$$S(u) = S1(u) \cdot S2(u) \tag{23}$$

According to the multiple feature maps, the saliency characteristics are presented in the different domains. The final saliency map can be enhanced based on the multiple feature maps. The target position can be easily obtained by the simple threshold segmentation. The Fig. 1 shows the flow chart of the proposed method.

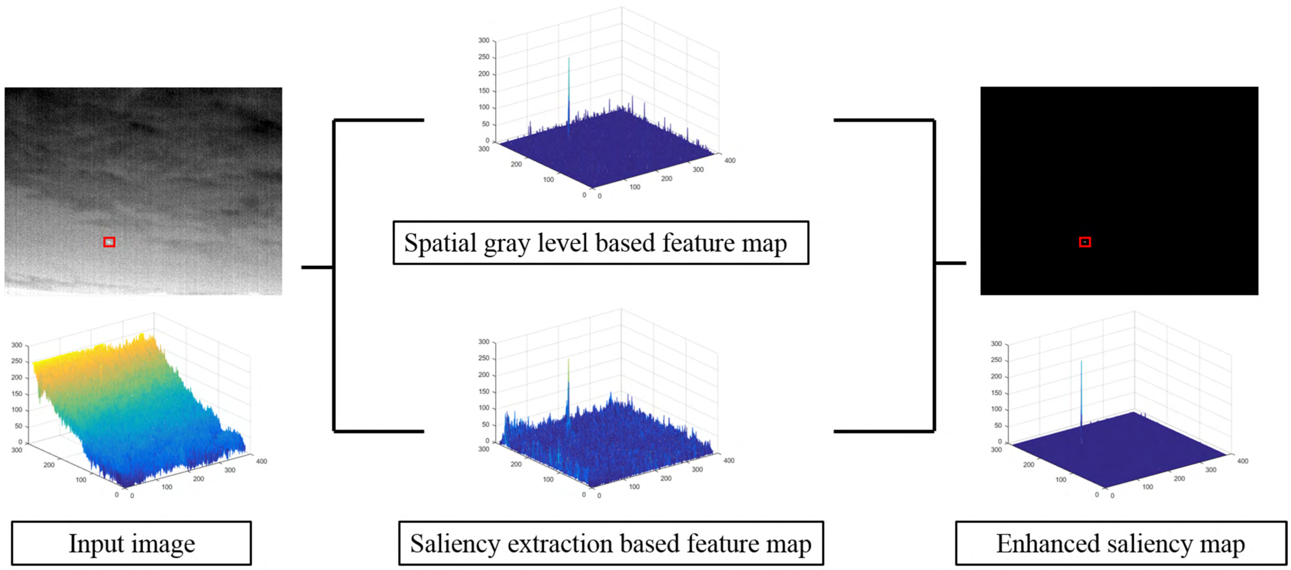

FIGURE 1. The flow chart of the proposed method.

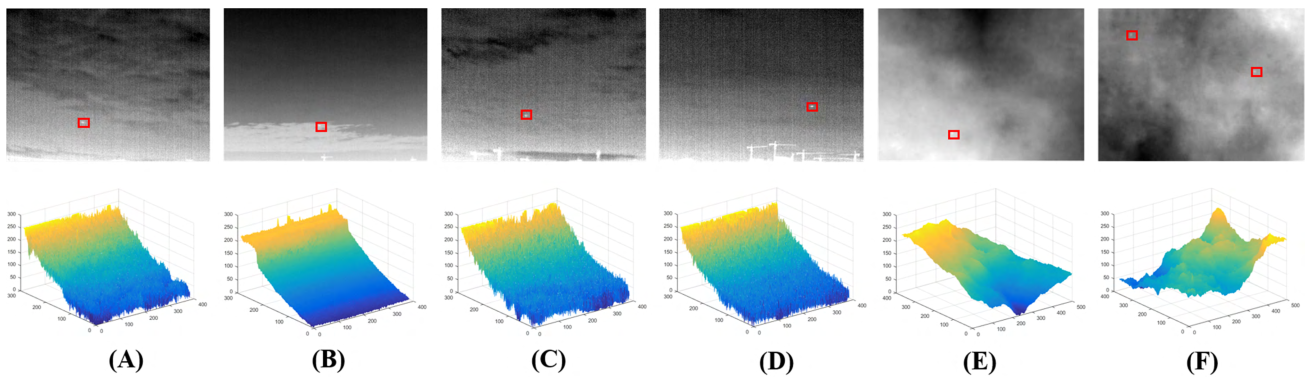

FIGURE 2. The real IR images and simulated images.

III. EXPERIMENTAL RESULTS

To evaluate the target detection performance of the proposed method, the simulation experiments based on the real IR images and the simulated images are introduced. In Fig. 2, the original gray images are shown in the first row and their three-dimensional gray level distributions are presented in the second row. The images (A), (B), (C), and (D) are the real IR images are collected by the IR detector. The small target is the airplane and labeled by the rectangle in the image. The images (A), (C) and (D) have numerous clutter in the background, and some regions of the images are high light. But the targets are dim and submerged in the clutter. The target is in the cloud with low signal to clutter ratio and cannot be perceived easily in image (B). The images (E) and (F) are the simulated images which simulate the cloud through the method in literature [23]. The simulated IR small targets are distributed randomly in the images. The targets are weak and the simulated cloud background are irregular. Some methods based on the HVS or saliency maps are chosen for the comparison in the experiments, including modified top-hat transformation (MT) [12], fast saliency method (FS) [18], multi-patch contrast method (MPCM, as mentioned in literature RLCM [17]), and relative local contrast method (RLCM) [17]. These methods and the proposed method are tested based on the images in Fig. 2. Through the three-dimensional gray level distributions of the results, the performance of background suppression and target enhancement of these methods can be perceived directly by the visual sense. In order to further assess these methods about the difficulty degree of small target detection, some widely used metrics are introduced. The metrics to evaluate the performances of the methods are signal-to-clutter ratio gain (SCRG) and background suppression factor (BSF) [24], which are as follows:

$$SCRG = \frac{(S/C)_{out}}{(S/C)_{in}}, BSF = \frac{C_{in}}{C_{out}} \tag{24}$$

where S and C are the signal amplitude and the clutter standard deviation respectively. The *in* and *out* are the original image and the result of the method. The SCRG and BSF

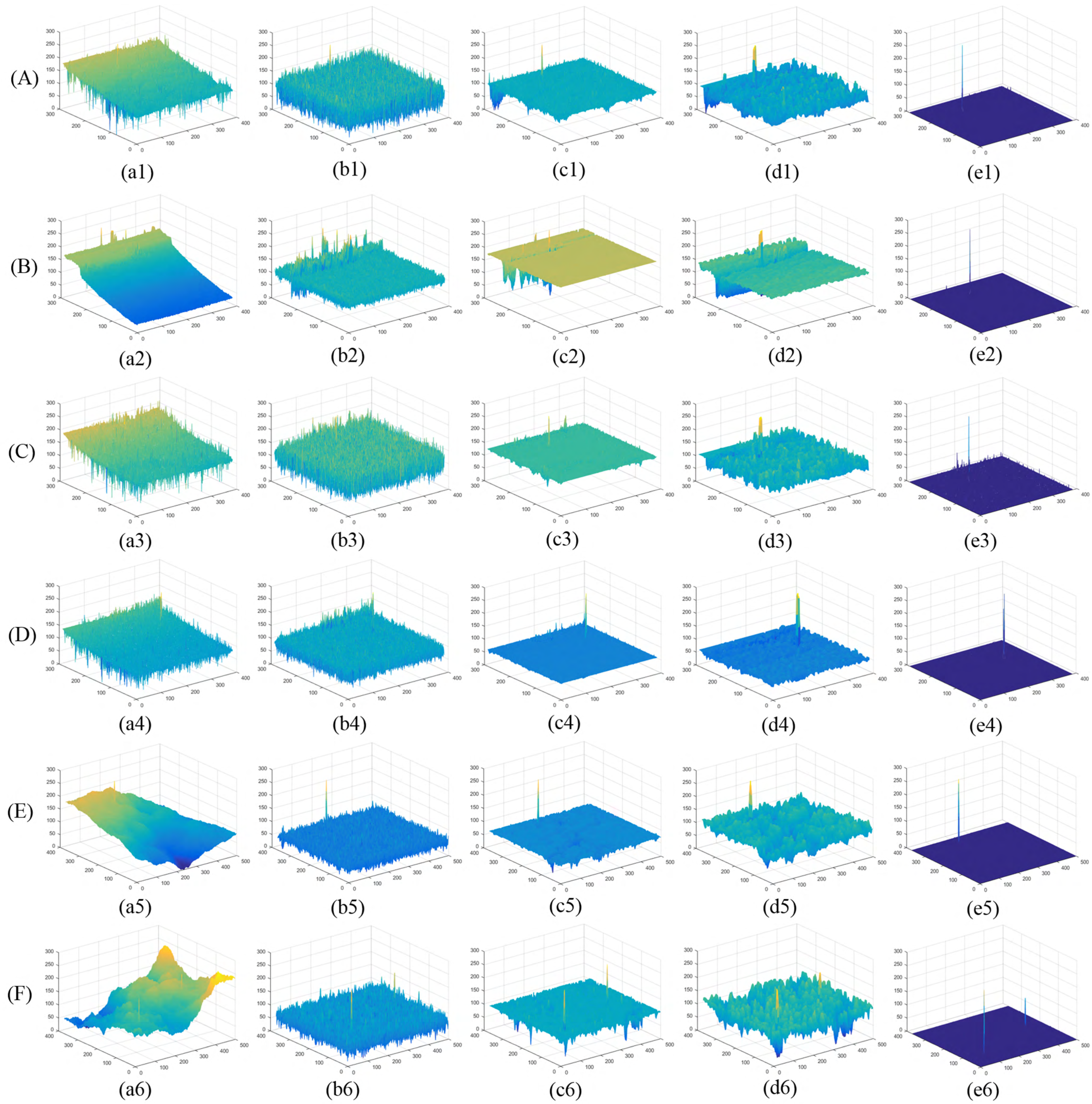

FIGURE 3. The results of the methods in the experiments based on three-dimensional gray level distributions. The each row is the results using the images in Fig.2 respectively. (a1)-(a6) and (e1)-(e6) are the results of the methods MT, FS, MPCM, RLCM and the proposed methods.

present the degree of the target enhancement and the background suppression. The metric reflecting the relationship between the detection rate Pd and false alarm rate Pf is receiver operator characteristic (ROC) curve [12], which is defined as follows:

$$Pd = \frac{number\ of\ detected\ pixels}{number\ of\ real\ target\ pixels} \tag{25}$$

$$Pf = \frac{number\ of\ false\ alarm\ pixels}{number\ of\ pixels\ in\ the\ whole\ image} \tag{26}$$

In the Fig. 3, the three-dimensional gray level distributions of the results based on the methods in the experiments are shown. (a1)-(a6) and (e1)-(e6) are the results of methods MT, FS, MPCM, RLCM and the proposed methods. The gray value ranges of these results in the experiments are normalized to [0,255]. The row (A)-(F) are the results based on the corresponding images in Fig. (2). It can be seen that method MT can enhance the small target in some degree in the results (a1)-(a4), but much clutter in the background can

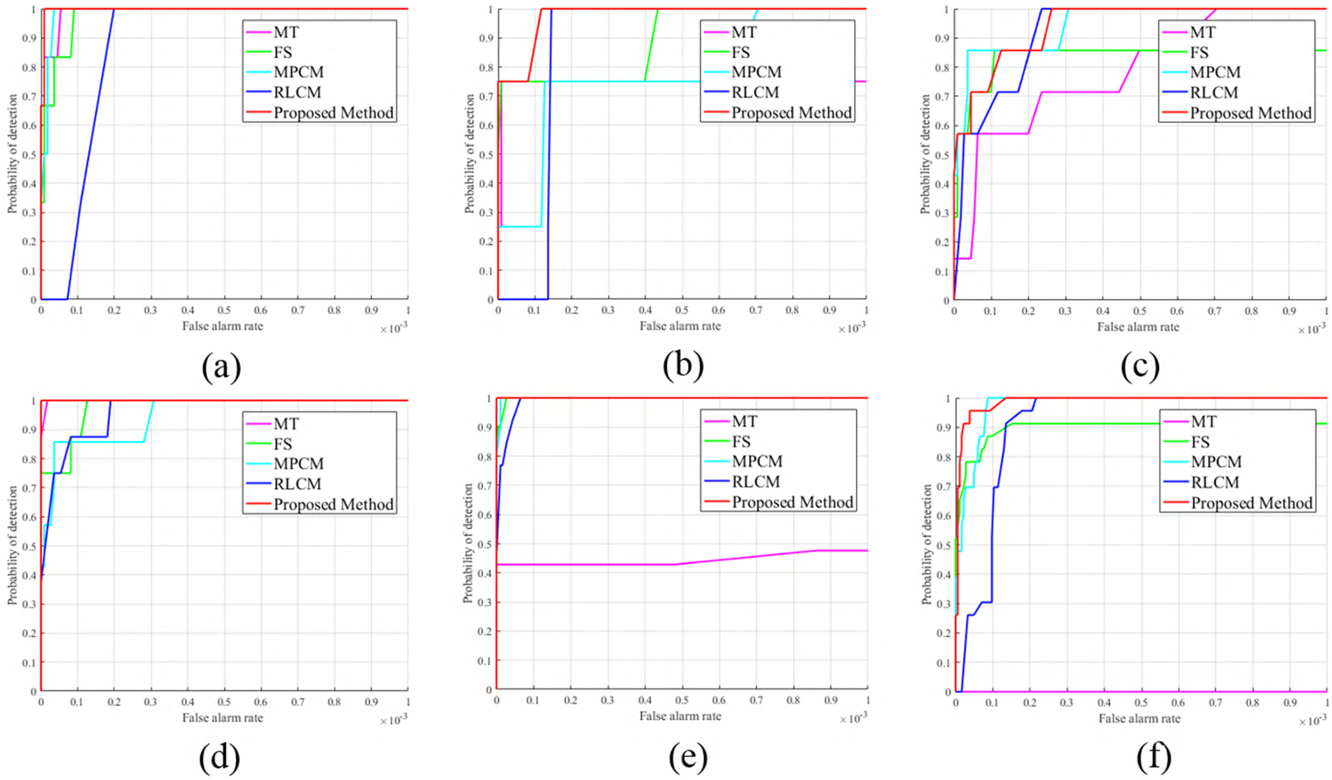

FIGURE 4. The ROC curves of the methods in the experiments.

TABLE 1. The values of SCRG for the methods in the experiments.

	MT	FS	MPCM	RLCM	Proposed method
(A)	2.9397	2.9109	9.2716	4.0979	**26.8974**
(B)	1.3400	5.4784	7.5579	4.2053	**17.9963**
(C)	3.0304	2.7331	9.4919	4.4405	**13.5129**
(D)	3.9295	3.9280	14.5099	6.7602	**25.6163**
(E)	1.2825	6.6458	9.5459	3.4511	**38.5919**
(FL)	1.0786	3.8177	6.9198	3.2335	**14.2704**
(FR)	1.0354	2.9533	5.1333	2.3521	**10.6827**

TABLE 2. The values of BSF for the methods in the experiments.

	MT	FS	MPCM	RLCM	Proposed method
(A)	3.0379	2.8269	9.5108	3.8919	**63.7477**
(B)	1.5499	6.3127	7.7595	3.7961	**75.8414**
(C)	2.9030	2.6068	9.2459	3.5481	**40.8591**
(D)	3.9676	4.1983	18.6092	6.0094	**49.9213**
(E)	1.2538	6.4967	9.3318	3.3737	**37.7266**
(F)	0.9999	3.3510	5.1559	1.9773	**35.5118**

disturb the target signal. In the results (a5)-(a6), the most of the background is not suppressed effectively and the targets are covered by the background. The method FS can enhance the target well in the simple background. However, if the background scenes are complex, the saliency maps would be influenced by the bright noise, such as the results (b1)-(b4). The method MPCM can magnify the target signal-to-clutter rate in some results such as (c1), (c4), (c5) and (c6). But in the other images, the background regions with high intensity are not suppressed sufficiently which can disturb the target detection. RLCM can strengthen the signals of targets without the effect of background clutter in some results such as (d2) and (d4). But in the other images, the great majority of background is not eliminated. The proposed method can remove the background effectively based on the images in Fig. 2. In some images which have low signal-to-clutter ratio, the proposed method still have robust performances to improve the target signal. It can be seen that the proposed method has feasible and adaptive ability for the various and complex scenes in IR images. Table 1 and Table 2 are about SCRG and BSF indexes. The higher value of each index reflects a better capability for the target detection. The line (FL) and (FR) SCRG are the left and right targets in image (F) respectively. In the tables, it can be seen that the proposed method have a better performance compared with the other methods in the experiments for improving the target signal-to-clutter rate and inhibiting the background. In the terms of the ROC curves in the Fig. 4, the different methods have own advantages under different conditions. At the same false alarm rate, the proposed method has higher detection rates than the other methods in the experiments. The proposed method can have robust ability with high detection rates and low false alarm rates simultaneously.

According to the experiments, it can be seen that the proposed method can have effective and robust performance in terms of background inhibition and target enhancement. However, there are some limitations in the proposed method which can be improved in the future. In the actual environment, the inhomogeneous and deficient images would affect the performance of the proposed method. For the heavy interference of the background in the images, the adaptability of the proposed method could be improved further.

IV. CONCLUSION

In this paper, we proposed an effective IR target detection method based on the multiple feature analysis. Based on the contrast mechanism and visual attention mechanism, spatial gray level based feature map and saliency extraction based feature map are designed for the final saliency map enhancement. It can highlight the target signal and suppress the background, which can separate the target from the complex background. The experimental results show that the proposed method is robust in improving the values of SCRG and BSF of the images, and has significant performance in terms of the detection rate and false alarm rate.

REFERENCES

[1] 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.

[2] 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.

[3] X. Du, M. Guizani, Y. Xiao, and H.-H. Chen, "A routing-driven elliptic curve cryptography based key management scheme for heterogeneous sensor networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1223–1229, Mar. 2009.

[4] Y. Xiao, X. Du, J. Zhang, and S. Guizani, "Internet protocol television (IPTV): The killer application for the next generation Internet," *IEEE Commun. Mag.*, vol.45, no. 11, pp. 126–134, Nov. 2007.

[5] S. Bourennane, C. Fossati, and A. Cailly, "Improvement of target-detection algorithms based on adaptive three-dimensional filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1383–1395, Apr. 2011.

[6] J. Guo, B. Song, F. R. Yu, Y. He, and M. Sookhak, "A survey on compressed sensing in vehicular infotainment systems," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2662–2680, 4th Quart., 2017.

[7] 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.

[8] M. Nasiri, M. R. Mosavi, and S. Mirzakuchaki, "Infrared dim small target detection with high reliability using saliency map fusion," *IET Image Process.*, vol. 10, no. 7, pp. 524–533, Jul. 2016.

[9] F. A. Sadjadi, "Infrared target detection with probability density functions of wavelet transform subbands," *Appl. Opt.*, vol. 43, no. 2, pp. 315–323, Feb. 2004.

[10] S. Kim and J. Lee, "Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track," *Pattern Recognit.*, vol. 45, no. 1, pp. 393–406, Jan. 2012.

[11] 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.

[12] X. Bai and F. Zhou, "Analysis of new top-hat transformation and the application for infrared dim small target detection," *Pattern Recognit.*, vol. 43, no. 6, pp. 2145–2156, Jun. 2010.

[13] C. Q. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4996–5009, Dec. 2013.

[14] S. Kim and J.-H. Lee, "Robust scale invariant target detection using the scale-space theory and optimization for IRST," *Pattern Anal. Appl.*, vol. 14, no. 1, pp. 57–66, Feb. 2011.

[15] X. Wang, G. Lv, and L. Xu, "Infrared dim target detection based on visual attention," *Infr. Phys. Technol.*, vol. 55, no. 6, pp. 513–521, Nov. 2012.

[16] X. Kang, B. Song, J. Guo, X. Du, and M. Guizani, "Attention-mechanism-based tracking method for intelligent Internet of vehicles," *Int. J. Distrib. Sensor Netw.*, vol. 14, no. 10, pp. 1–16, 2018.

[17] J. Han, K. Liang, B. Zhou, X. Zhu, J. Zhao, and L. Zhao, "Infrared small target detection utilizing the multiscale relative local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 612–616, Apr. 2018.

[18] S. Qi, G. Xu, Z. Mou, D. Huang, and X. Zheng, "A fast-saliency method for real-time infrared small target detection," *Infr. Phys. Technol.*, vol. 77, pp. 440–450, Jul. 2016.

[19] S. Huang, M. Li, X. Wang, X. Zhao, L. Yang, and Z. Peng, "Infrared small target detection with directional difference of Gaussian filter," in *Proc. 3rd IEEE Int. Conf. Comput. Commun.*, Dec. 2017, pp. 1698–1701.

[20] S. Kim, Y. Yang, J. Lee, and Y. Park, "Small target detection utilizing robust methods of the human visual system for IRST," *J. Infr., Millim., Terahertz Waves*, vol. 30, no. 9, pp. 994–1011, Sep. 2009.

[21] X. Bai, F. Zhou, and B. Xue, "Image fusion through local feature extraction by using multi-scale top-hat by reconstruction operators," *Optik-Int. J. Light Electron Opt.*, vol. 124, pp. 3198–3203, Sep. 2013.

[22] Z. Zhou, B. Wang, S. Li, and M. Dong, "Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters," *Inf. Fusion*, vol. 30, pp. 15–26, Jul. 2016.

[23] D. Yang and K. Yang, "Cloud scenes generation based on the improved random midpoint displacement method," in *Proc. Int. Conf. Inf. Sci. Cloud Comput.*, 2013, pp. 166–171.

[24] C. I. Hilliard, "Selection of a clutter rejection algorithm for real-time target detection from an airborne platform," *Proc. SPIE*, vol. 4048, pp. 74–84, Jul. 2000.

YUWEN CHEN received the M.S. degree in signal and information processing from Shaanxi Normal University, Xi'an, China, in 2017. She is currently pursuing the Ph.D. degree in communication and information systems with Xidian University, Xi'an, China. Her research interests include pattern recognition, machine learning, and multimodal data fusion.

BIN SONG (M'10–SM'19) received the B.S., M.S., and Ph.D. degrees in communication and information systems from Xidian University, Xi'an, China, in 1996, 1999, and 2002, respectively. In 2002, he joined the School of Telecommunications Engineering, Xidian University, where he is currently a Professor of communications and information systems. He is also an Associate Director of the State Key Laboratory of Integrated Services Networks. He has authored over 60 journal papers or conference papers and 30 patents. His research interests and areas of publications include video compression and transmission technologies, video transcoding, video signal processing based on compressed sensing, computer vision, big data, the IoT networks, and multimedia communications.

XIAOJIANG (JAMES) DU received the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland at College Park, in 2002 and 2003, respectively. He is currently a Tenured Professor with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA, USA. His research interests include wireless communications, wireless networks, security, and systems. He has authored over 300 journal and conference papers in these areas, as well as a book (published by Springer). He is a Life Member of the ACM. He received the Best Paper Award from the IEEE GLOBECOM 2014 and the Best Poster Runner-Up Award from the ACM MobiHoc 2014. He has served as the Lead Chair for the Communication and Information Security Symposium, the IEEE International Communication Conference 2015, and the Co-Chair of the Mobile and Wireless Networks Track of the IEEE Wireless Communications and Networking Conference 2015. He is (was) a Technical Program Committee (TPC) member of several premier ACM/IEEE conferences.

MOHSEN GUIZANI (S'85–M'89–SM'99–F'09) received the B.S. (Hons.) and M.S. degrees in electrical engineering, and the M.S. and Ph.D. degrees in computer engineering from Syracuse University, Syracuse, NY, USA, in 1984, 1986, 1987, and 1990, respectively. He is currently a Professor with the CSE Department, Qatar University, Qatar. Previously, he has served in different academic and administrative positions with the University of Idaho, Western Michigan University, University of West Florida, University of Missouri-Kansas City, University of Colorado Boulder, and Syracuse University. He has authored nine books and over 500 publications in refereed journals and conferences. He has guest edited a number of special issues in the IEEE journals and magazines. His research interests include wireless communications and mobile computing, computer networks, mobile cloud computing, security, and smart grid. He has also served as a member, the Chair, and the General Chair for a number of international conferences. He is a Senior Member of the ACM. Throughout his career, he received three teaching awards and four research awards. He has received the 2017 IEEE Communications Society WTC Recognition Award and the 2018 AdHoc Technical Committee Recognition Award for his contribution to outstanding research in wireless communications and Ad-Hoc Sensor networks. He was the Chair of the IEEE Communications Society Wireless Technical Committee and the Chair of the TAOS Technical Committee. He is currently the Editor-in-Chief of the *IEEE Network Magazine*, serves on the editorial boards of several international technical journals, and the Founder and the Editor-in-Chief of the *Wireless Communications and Mobile Computing Journal* (Wiley). He has served as the IEEE Computer Society Distinguished Speaker, and is currently the IEEE ComSoc Distinguished Lecturer.