Smoke detection in infrared images based on superpixel segmentation

Dai, Min, Gao, Peng, Sha, Mozhou, Tian, Jinwen

Min Dai, Peng Gao, Mozhou Sha, Jinwen Tian, "Smoke detection in infrared images based on superpixel segmentation," Proc. SPIE 11432, MIPPR 2019: Remote Sensing Image Processing, Geographic Information Systems, and Other Applications, 1143213 (14 February 2020); doi: 10.1117/12.2538195

Event: Eleventh International Symposium on Multispectral Image Processing and Pattern Recognition (MIPPR2019), 2019, Wuhan, China

Smoke detection in infrared images based on superpixel segmentation

Min Dai*[a], Peng Gao[a], Mozhou Sha[b], Jinwen Tian[a]
[a] School of Artificial Intelligence and Automation, HUST, 1037 Luoyu Road, Wuhan, Hubei, China;
[b]Institute of Mechanical and Electrical Engineering, Beijing.

ABSTRACT

Infrared smoke interference technology seriously infected the combat effectiveness of photoelectric guided weapons in modern warfare. As a result of the occlusion caused by smoke screen, the robustness of image matching guidance algorithm will decrease. Thus, to judge whether there is smoke interference in images and smoke screen area extraction are of great importance for the accuracy of image matching guidance algorithm. However, most of the smoke detection methods aimed at fire early warning, so that they focused on whether smoke exists or not. While both of the discrimination of smoke interference and smoke screen area extraction are what we concern. In this paper, a smoke detection method based on superpixel segmentation and region merging is proposed. Firstly, over-segmentation regions of input infrared image with superpixel segmentation are obtained. Then, fusion texture feature of the image is computed. Finally, superpixel regions are merged based on the fusion features of each superpixel block obtained in the previous step and smoke screen area extraction is completed.

Keywords: smoke detection, infrared image, superpixel segmentation, fusion texture features.

1. INTRODUCTION

Photoelectric imaging terminal guidance weapon has high efficiency-cost ratio and shows great power in battlefield applications [1]. At the same time, IR smoke interference technology is widely used in modern warfare as an important means of photoelectric countermeasures [2]. By releasing smoke bombs, a large smoke screen could form in a short time to shield military targets, such as command posts, power plants, etc. Which makes it difficult for infrared imaging devices to obtain target images. Smoke interference greatly reduces the operating range and target recognition ability of the infrared imaging guidance system, seriously affecting the performance of the image matching guidance algorithm, and even completely invalidating it [3], which greatly weakens the combat effectiveness of the weapon. An American expert pointed out that more than 90% of the target capture systems in the United States will be affected by existing smoking equipment [4].

The specific mechanism of smoke interference is mainly reflected in the following two aspects. Firstly, the suspended particles in the smoke screen can absorb and scatter radiation of a specific wavelength, and then emit electromagnetic waves outward in other wavelengths or propagation directions. This means that the infrared radiation is attenuated as it passes through the smoke screen, causing the energy received by the photodetector on the infrared seeker to be greatly attenuated, even below the sensitivity of the detector, making the target difficult to detect. Secondly, the smoke screen generate a large amount of high-temperature aerosol particles to radiate infrared radiation [2] stronger than the target, thus covering the infrared characteristics of the target, and finally the target is covered by a highlighted area in the imaging result, so that the target cannot be resolved. To sum up, smoke interference technology will seriously affect the impacts of infrared imaging guided weapons, and extraction of smoke screen area is an effective way to improve the performance of infrared imaging automatic target recognition.

*meiwen@hust.edu.cn; phone 1 867 286-2141;

MIPPR 2019: Remote Sensing Image Processing, Geographic Information Systems, and Other Applications,
edited by Zhiguo Cao, Jie Ma, Zhong Chen, Yu Shi, Proc. of SPIE Vol. 11432, 1143213 · © 2020 SPIE
CCC code: 0277-786X/20/$21 · doi: 10.1117/12.2538195

Recently, most of the application scenarios of smoke detection are early warnings of fires[5]-[8]. In fires, smoke screens are often produced in the smoldering stage before the fire is generated. C. Yu [5] used the optical flow method to extract the motion area in smoke image, and then completed classification with other motion features. D. Yang [6] selected the diffusion characteristics of smoke in the time domain and the irregular contour feature in the space domain to analyze the smoke, and extracted the texture features of the suspicious area of smoke. J. Yuan [7] used three frame difference method to test the suspected area of smoke, whose texture feature combined with LBP and GLCM was extracted to identify of smoke and non-smoke. F. Yuan et al. [8] used the semantic segmentation method based on fully convolutional networks for smoke screen extraction, and achieved about 70% of mIOU on their synthesized smoke images datasets. Smoke detection algorithms above which have achieved good results on fire early warning are unable to meet the application requirements of this paper. Because most of them aimed at identifying whether smoke exists or not, but didn't extract smoke screen area at pixel level. This paper proposed a smoke detection method based on SLIC superpixel segmentation [9] and region merging with a fusion texture feature of Gabor wavelet feature and fractal dimension feature. Meanwhile, it is found that the smoke screen area changes rapidly in infrared smoke interference image sequences, and the fractal characteristics of images with smoke change drastically. So the fractal dimension variation can be computed to determine whether there is smoke interference or not.

Therefore, the proposed smoke detection method consists of two major components, one is the discrimination of smoke interference, the other is the extraction of smoke screen area：

1) When judging whether there is smoke in the infrared image or not, the image is firstly processed into blocks, and then the fractal dimensions of each block image are separately counted, finally the discrimination of smoke interference is judged comprehensively. Compared with the single fractal feature extraction method of the whole image, this method can reduce the misjudgment caused by the internal variation of smoke and improve the accuracy of the discrimination of smoke interference.

2) Because of the similarity of the gray scale of the smoke screen and the background in infrared smoke interference images, it is difficult to distinguish them with grayscale feature. Therefore, when extracting the smoke screen area, classical and effective texture features are adopted. At the same time, the SLIC superpixel segmentation algorithm can generate compact, approximately uniform superpixels [9]-[11] on the one hand, facilitating subsequent region merging based on feature similarity. On the other hand, the target contour can be well maintained, which can greatly improve the accuracy of the smoke screen area extraction.

The rest of this paper is organized as follows. Sec. 2 describes the proposed smoke detection method based on SLIC superpixel segmentation and texture features extraction, then experimental results and analysis are presented in Sec. 3. Finally the conclusions are summarized in Sec. 4.

2. INFRARED IMAGE SMOKE DETECTION ALGORITHM

The two major parts of the proposed smoke detection algorithm are respectively shown in figure 1 and figure 2. One is the discrimination of infrared image smoke interference, the other is the overall flow of the extraction of smoke screen area.

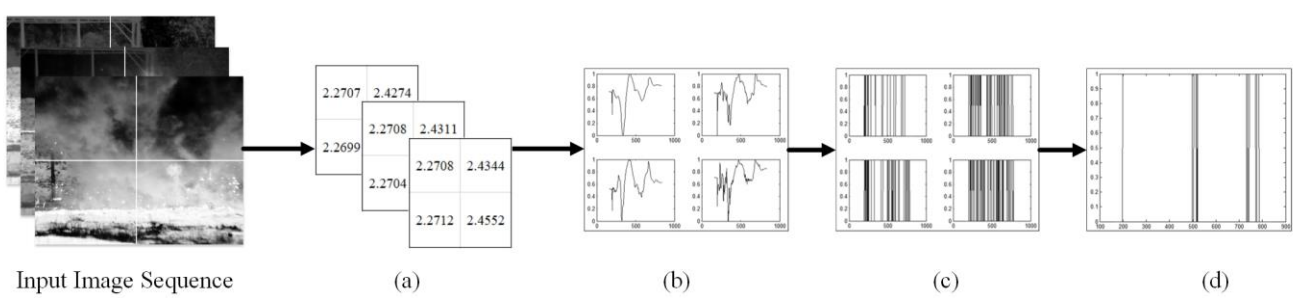


Figure 1. The architecture of the discrimination of smoke interference. (a) The fractal dimension of each block. (b) The normalized fractal dimension variation of per 10 frames. (c) The discrimination result of each block. (d) The final discrimination result of whole infrared smoke interference image sequence.

As shown in figure 1, the input images are evenly divided into 4 blocks, then the fractal dimension of each block is calculated separately, and the result is normalized, and then the fractal dimension variation of per 10 frames is calculated. The block discrimination is determined according to the magnitude of the variation. Finally, the results of each block discrimination are combined to determine whether the smoke discrimination of the whole image. The calculation process will be described in detail in Sec. 2.1.

As shown in figure 2, the input image is subjected to SLIC superpixel segmentation, Gabor wavelet transform, and fractal dimension. According to the extracted texture features, the statistical characteristics of each superpixel block are obtained. And then regions are merged based on feature similarity of different block. Finally, neighboring regions are merged to correct the smoke screen area. The principle and calculation process will be introduced in detail in Sec. 2.2~2.4.

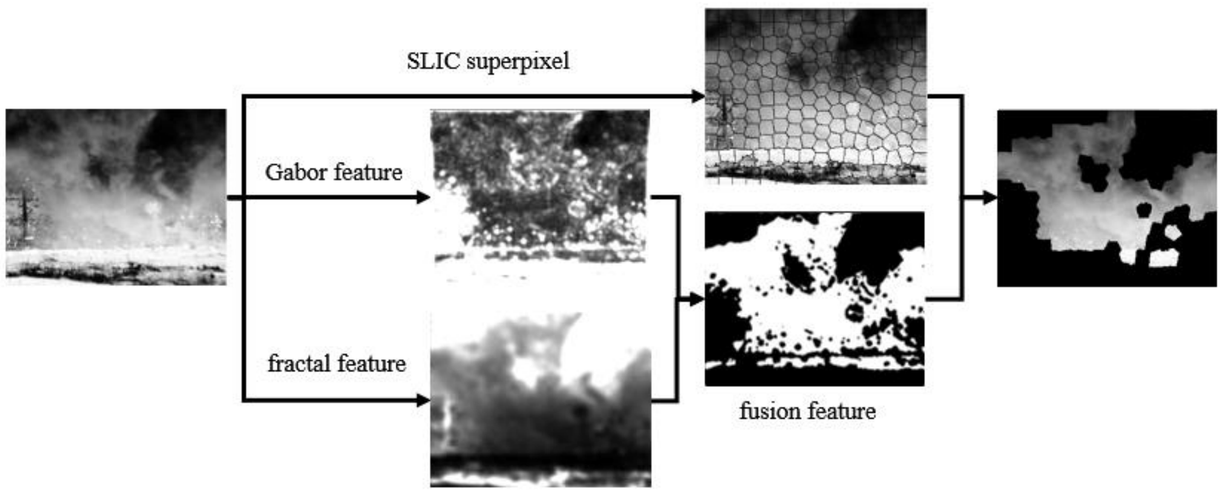


Figure 2.The overall flow of the smoke screen extraction algorithm.

2.1 The discrimination of smoke interference

For an input image sequence, every image is divided into 4 blocks. The fractal dimension of each block is shown in Fig. 1 (a), defined as $F=\{F(i,j)\,|\,i=100,101,\ldots,800;\ j=1,2,3,4\}$, where i denotes the i-th frame of sequence, and j denotes the j-th block from top to bottom, left to right. The calculation process of the fractal dimension will be described in detail in Sec. 2.3. Then their normalized variation and result is shown in Fig. 1 (b), defined as $NF=\{NF(i,j)\}$. $NF(i,j)$ is calculated as:

$$NF(i,j)=\frac{F(i,j)\text{-}F(i-1,j)}{\max F(:,j)-\min F(:,j)} \tag{1}$$

A threshold is used to split the normalized variation curve of each block. As shown in Fig. 1 (c), discrimination result of the four blocks $B=\{B(i,1),B(i,2),B(i,3),B(i,4)\}$ can be formulated as follows:

$$B(i,j)=\begin{cases}1,\left|NF(i,j)\right|>T\\0,\left|NF(i,j)\right|\le T\end{cases} \tag{2}$$

The final discrimination result of whole infrared smoke interference image sequence is denoted as $R=\{R(i)\}$, which is calculated as:

$$R(i)=B(i,1)\,||\,B(i,2)\,||\,B(i,3)\,||\,B(i,4) \tag{3}$$

2.2 SLIC superpixel segmentation

Superpixel refers to the irregular block of pixels with certain visual perception meaning composed of adjacent pixels with similar features. Superpixel segmentation method groups the pixels with the similarity of features between pixels, and uses a small number of superpixels to replace a large number of pixels to describe image features. We use a simple linear iterative clustering (SLIC) method [9] to generate superpixels for preliminary image segmentation.

SLIC superpixel segmentation requires very few parameters to be set. By default, only the number of superpixels K is required. For example, an input image containing *N* pixels is supposed to be divide into *K* superpixel blocks, then each block has has approximately *N/K* pixels, and the distance between two adjacent blocks is about $S=\sqrt{N/K}$. At the beginning of the SLIC algorithm, the input image usually need to be converted from RGB color space to CIELAB color space. Then the color value and the coordinates of a pixel are be represented as $[l\ a\ b]^T$ and $[x\ y]^T$ respectively. But for infrared images, there is no need to do color space conversion. The similarity between two pixels is measured by both color distance d_c and spatial distance d_s . The distance calculation method is as follows.

$$d_c=\sqrt{(l_j-l_i)^2}$$

$$d_s=\sqrt{(x_j-x_i)^2+(y_j-y_i)^2}$$

$$D=\sqrt{(\frac{d_c}{m})^2+(\frac{d_s}{S})^2} \tag{4}$$

Where *D* represents the complete distance between the two pixels, m is a constant to balance the color information and spatial information. The smaller the value of m, the less spatial proximity is emphasized. We use the SLIC algorithm to over-segment the input infrared image into K_1 superpixel blocks:

$$SP=\{r_i \mid i=1,2,\ldots,K_1\} \tag{5}$$

2.3 Texture features extraction

A. Gabor wavelet transform feature

The Gabor wavelet transform is very similar to the visual stimuli of simple cells in the mammalian basic visual cortex, so it can achieve good results in spatial local information and frequency domain information extraction [12]. In this paper, multi-channel filtering technology is used to apply a set of Gabor wavelet transforms with multi-directional and multi-scale to the image. Then each channel can obtain a kind of local characteristics of an input image, which are combined to obtain final Gabor feature of the image. In the space domain, the 2D-Gabor filter can be regarded as a Gaussian kernel function to modulate a sinusoidal plane wave. That is, the result expressed in the frequency domain is the product of the Gaussian kernel function and the complex sinusoidal plane wave. The definition of the kernel function is as follows:

$$\psi_{u,v}(x,y)=\frac{k^2}{\sigma^2}\exp(-\frac{k^2(x^2+y^2)}{2\sigma^2})\left[\exp(ik(\frac{x}{y}))-\exp(-\frac{\sigma^2}{2})\right] \tag{6}$$

Where *u* means direction, *v* means scale, (x,y) is the coordinates of a pixel, and $k=\begin{pmatrix}k_x\\k_y\end{pmatrix}=\begin{pmatrix}k_v\cos\varphi_u\\k_v\sin\varphi_u\end{pmatrix}$, $k_v=2^{-\frac{v+2}{2}\pi}$,

$\varphi_u=\frac{u\pi}{K}$, $K=8$, $\sigma=2\pi$, $u\in[0,1,\ldots,7]$, $v\in[0,1,2]$.

In this paper, the input image is convoluted with Gaussian kernels of three scales in eight directions, and 24 Gabor feature maps are extracted. By averaging them, the final Gabor feature map is obtained.

B. Fractal feature

The concept of fractal was originally proposed by the mathematician Benoit B. Mandelbrot. Fractal theory uses fractal dimensions (usually a decimal) to describe objective things in nature, which usually have self-similarity. There are many different definitions and methods to compute fractal dimension, including self-similar dimension, Hausdorff dimension, box count dimension, power spectrum dimension, and structure function dimension [13]. This paper uses a method based on the quadratic pyramid model [14] to obtain the local fractal feature of an image. The process of obtaining the fractal dimension is given below.

1) Selecting 8×8 image block centered on pixel $I(x, y)$. Then taking a grid of different sizes (1x1, 2x2, 4x4 and 8x8) to cover the block.
2) Taking the average gray level of the four vertices of the grid as the high value, the grid as the bottom constitutes a quadrangular pyramid, and calculating the total surface area of all quadrangular pyramids S_1, S_2, S_3, S_4 under the corresponding grid size, the point pairs $(\log 1, \log S_1), (\log 2, \log S_2), (\log 4, \log S_3), (\log 8, \log S_4)$ can be obtained.
3) The least square method is used to fit a straight line. The absolute value of whose slope is the fractal dimension corresponding to the pixel $I(x, y)$. And the final Fractal feature map $F(x, y)$ is obtained.

2.4 Feature fusion and region merging

A. Feature fusion

To combine the two texture feature maps $G(x, y)$ and $F(x, y)$ obtained in Sec. 2.3. Firstly, the two texture maps are subjected to OTSU [15] adaptive threshold segmentation respectively. Then two binary maps $GB(x, y)$ and $FB(x, y)$ are obtained, the intersection of which is calculated as $B' = \{b'(x, y) = gb(x, y) \,\&\, fb(x, y) \mid gb(x, y) \in GB, fb(x, y) \in FB\}$. At this point, the feature fusion is completed.

B. Superpixel region merging

After feature fusion, the final binary image B' and superpixel segmentation result SP are used for the final step of the superpixel region merging. Firstly, the number of pixels in each superpixel block region n_i and the number of pixels in the binary image whose value is 1 corresponded to the superpixel block region are counted. And then calculating the proportion of numbers of 1 in each superpixel block region.

$$p_i = \frac{\sum b_i'(x, y)}{n_i}, (x, y) \in r_i \tag{7}$$

When p_i exceeds 0.5, the superpixel block r_i is marked as smoke screen area. Finally, all the superpixel blocks marked as smoke screen area are merged and mapped to the original input image to obtain the smoke region extraction result.

3. EXPERIMENTS AND RESULTS

The experimental platform is configured as an Intel(R) Core(TM) i7-4720HQ CPU 2.6GHz, 8G memory, Windows10 operating system and VS2013 development environment. The external dependent library is OpenCV2.4. This section mainly includes three parts of experiments. The first one is the occlusion matching experiment of the ground building, to quantitatively analyze the influence of smoke occlusion on the matching guidance algorithm. The second one is the smoke screen interference discrimination experiment, and the third part is the smoke screen area extraction experiment.

In the occlusion matching experiment of the ground building, the building targets in the experiment images were occluded with three different sizes of smoke screen, which cover 10%, 20% and 30% of the targets respectively. In the meantime, unoccluded target slices are made into reference images. And then the classic HOG template matching algorithm [16] is used to match the reference images and real-time images. The experiment data is based on 200 single-channel building images with resolutions of 320*256, which are firstly made into the highest resolution, 1/2 resolution, and 1/4 resolution according to different projectile distances. Secondly, the target is occluded with three different occlusion proportion. The experimental results are shown in Table 1, which show that the accuracy of target recognition algorithm decreases with the increase of target occlusion proportion.

Table 1. The matching accuracy of HOG template matching algorithm with different occlusion proportions and different target resolutions.

Occlusion proportion \ **Target resolution**	**1**	**1/2**	**1/4**
0	100%	100%	100%
10%	98%	98%	93%
20%	91%	92%	81%
30%	84%	83.5%	64.5%

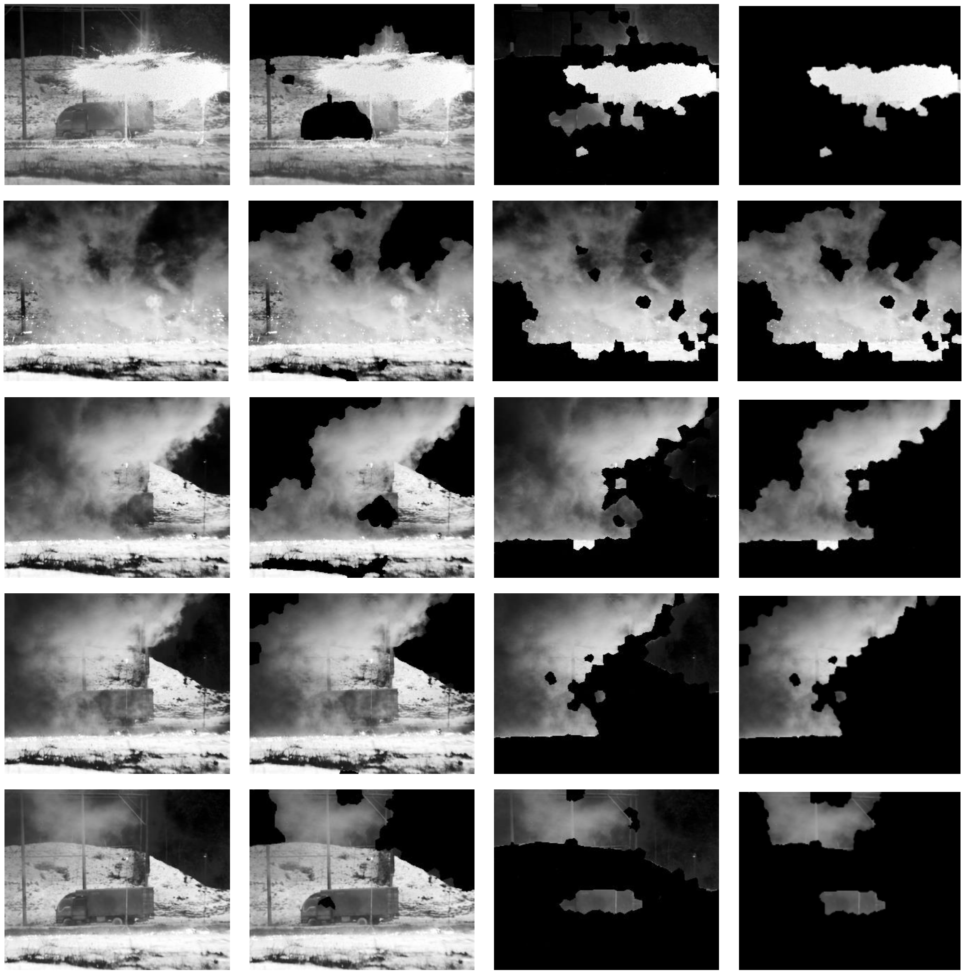
The experiment data of the second experiment and the third experiment are real infrared smoke interference images. We collected two infrared smoke interference image sequences from the shooting range, the data collecting environment of which are windy and windless respectively. This paper intercepts the sequences from the explosion of the smoke bomb to the disappearance of the smoke screen. The infrared smoke interference image sequences have a resolution of 320*256, and image bit depth of 8. In the experiment, the windy sequence contains 590 images with smoke screen and 110 images with no smoke. In the windless sequence, there are 1320 images with smoke screen interference and remaining 80 images with no smoke. In order to verify the accuracy of our method, we use precision, recall and F_1-score to evaluate the results quantitatively. The experiment results of the discrimination of smoke interference and smoke screen area extraction are shown in table 2.

Table 2. The experiment results of the discrimination of smoke interference and smoke screen area extraction.

	Smoke discrimination / %			Smoke area extraction (IOU)/%		
	Precision	Recall	F_1-score	S-F	S-G	Fusion
Windy sequence	100	98.57	99.28	45.32	67.50	80.05
Windless sequence	100	97.14	98.55	44.25	68.02	78.92

As shown in table 2, the accuracy of smoke discrimination is higher when it is windy than it is windless. The reason of the result is that the smoke diffusion speed is slower when it is windless, which is easy to cause misjudgment. The S-F and S-G represent results of single fractal feature and single Gabor feature respectively. Due to space limitations, only several typical smoke screen area extraction examples are listed below in figure 3.

As shown in figure 3, compared with the single texture feature, the fusion feature can effectively improve the accuracy of the smoke screen extraction. The reason is that the fractal texture feature can effectively extract the morphological features of the smoke area, but the ground areas which are very light causes lots of area extraction error. While when combined with the Gabor texture feature, some extraction error can be avoided. Thereby texture feature fusion improve the accuracy of the extraction of the smoke screen area.



(a) (b) (c) (d)

Figure 3. The examples of smoke screen area extraction. (a) The original input images. (b) Region merging results of single fractal feature(S-F). (c) Region merging results of single Gabor feature(S-G). (d)Smoke screen extraction results of our method.

4. CONCLUSION

In this paper, an infrared smoke detection method based on superpixel segmentation and fusion texture feature is proposed. Firstly, quantitative experiment is conducted to verify the influence of smoke interference on target matching recognition algorithm. And then smoke screen interference discrimination experiment and smoke screen area extraction experiment are carried out respectively, whose results demonstrate that our method can effectively improve the detection effect of infrared smoke interference sequences. Especially achieving a higher accuracy of smoke interference discrimination. While the effect of smoke screen area extraction is not so satisfying. In our future work, we will reduce remaining area misjudgement, and distinguish thin smoke and thick smoke, which actually cause different occlusion effects on targets.

REFERENCES

[1] X. Yang, C. Li, "Research on the jamming effect evaluation of smoke screen in opto-electric imaging guidance," Laser and Infrared, Vol. 42, No. 2, pp. 170-174 (2012).

[2] Z. Zhang, Q. Dong, Z. Jin, "Application of Smoke Screen on Countermining IR and laser Guided Weapons," Modern Defence Technology, Vol. 35, No.3, pp. 17-20 (2007).

[3] P. Xue, W. Dong, W. Luo, "Research on the jamming effects of smoke screen on target identification capability of photoelectric guidance," Laser and Infrared, Vol. 48, No. 3, pp. 374-378 (2018).

[4] J. Xu, B. Liu, J. Yang, et al. "Anti-Jamming Strategy for Smoke-Screen Jamming and Composite Seeker," Vol. 53, No. 6, pp. 62-68 (2016).

[5] C. Yu, "The study of recognition of the fire smoke video images based on optical flow and multi-information fusion detection algorithm," PhD dissertation. University of Science and Technology of China (2010).

[6] D. Yang, "Low-illumination indoor smoke detection method based on video," Master thesis. Huazhong University of Science and Technology (2012).

[7] J. Yuan, "Method of active infrared smoke video detection based on fusion texture feature," University of Science and Technology of China (2015).

[8] F. Yuan, L. Zhang, X. Xia, et al. "Deep Smoke Segmentation," Vol. 357, pp. 248-260 (2019).

[9] R. Achanta, A. Shaji, K. Smith, et al. "SLIC superpixels," EPFL Technical Report, No. 149300 (2010).

[10] H. Yun, L. Xu, X. Sun, et al. "Moving Target Detection Algorithm Based on Superpixel Spatiotemporal Saliency," Infrared Technology, Vol. 37, No. 5, pp. 599-608 (2015).

[11] R. Achanta, A. Shaji, K. Smith, et al. "SLIC superpixels compared to state-of-the-art superpixel methods," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 34, No. 11, pp. 2274-2282 (2012).

[12] X. Wang, Y. Lu, S. Song, et al. "Face recognition based on Gabor wavelet transform and modular PCA," Computer Engineering and Applications, Vol. 48, No. 3, pp. 176-178 (2012).

[13] B. Li, K. Yang, Y. Liu, "Fractal Theory and Calculation Method of Fractal Parameter," Tool Engineering, Vol. 38, No. 12, pp. 80-84 (2004).

[14] D. Bai."Research on forest fire image segmentation based on local fractal," Master thesis. Heilongjiang University (2015).

[15] J. Liu, W. Li, "The Automatic Thresholding of Gray-Level Pictures via Two-Dimensional OTSU Method," Acta Automatica Sinica, Vol. 19, No. 1, pp. 101-105 (1993).

[16] C. Gao, F. Wang, "Algorithm of License Plate Recognition Based on Template Matching and Local HOG Feature," Computer Systems and Applications, Vol. 26, No. 1, pp. 122-128 (2017).