Chinese Journal of Electronics
Vol.28, No.5, Sept. 2019

The Cat's Eye Effect Target Recognition Method Based on Visual Attention

WANG Xingbin[1,2], ZHANG Jun[3], WANG Shuaihui[1,4]

(1. *State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Science, Beijing 100093, China*)

(2. *University of Chinese Academy of Sciences, Beijing 100049, China*)

(3. *Hubei University of Arts and Science, Xiangyang 441053, China*)

(4. *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China*)

Abstract — The Cat's eye effect target recognition method based on visual attention (CTRVA) is proposed. The difference image can be processed by a designed second-directional derivative filter at eight directional channels. Morphological method is employed to deal with the filtered image in all directions, which ensures that target can be easily distinguished from background. The salient maps for each channel where the potential targets exist are calculated through the spectral residual approach, and the "target-saliency" map is computed by a designed saliency fusing method. The coarse detection is performed by the adaptive threshold to extract candidate targets from the "target-saliency" map. The real target region is identified by the characteristics of the cat's eye effect target. Experimental results show that the proposed method is efficient and has an outstanding performance for cat's eye effect target detection.

Key words — Cat's eye effect target recognition method based on visual attention (CTRVA), Target recognition, Saliency map.

I. Introduction

Over the past decade, many researchers have paid attention to the detection of cat's eye effect targets[1−9], and great progress has been made in this field. A recognition method based on roundness and eccentricity is put forward[10]. However, the method based only on shape and gray priors cannot precisely recognize cat's eye effect targets from the dynamic background. Another method based on Shape-frequency dual criterions (SFDC) method[11] is put forward to improve the identification probability of the cat's eye effect targets in the dynamic background. By demodulating frequency and identifying the shape feature of targets, the cat's eye effect targets can be found. However, the SFDC method requires a large number of image sequences to demodulate the frequency of illumination, so it is impossible to meet the requirement of real-time applications. Although fast cat-eye effect target recognition based on saliency extraction[12−18] has potential to achieve real-time cat-eye target recognition by parallel mode, it cannot eliminate more false targets after threshold the target saliency map.

In recent years, human vision has selective attention property which can help people focus on the salient target quickly without preceding training from complex unknown scene[13]. Motivated by this, the Cat's eye effect target recognition method based on visual attention (CTRVA) is proposed. Firstly, the difference image which can be acquired by the subtraction of the active image and passive image is decomposed into eight directional channels using a designed Second-directional derivative filter (SDDF) based on a facet model[19]. Then based on the fact that small target energy scatters in all direction, namely it lacks directivity information[20]. Morphological method is employed to deal with the filtered image in all directions, which ensures that target can be easily distinguished from background. Secondly, the salient maps for each channel where the potential targets exist are calculated through the Spectral residual (SR) approach[21], and the "target-saliency" map is computed by a designed saliency fusing method that can fuse all saliency maps from each channel. Thirdly, the

*Manuscript Received Jan. 2, 2018; Accepted June 21, 2018. This work is supported by the Hubei Superior and Distinctive Discipline Group of "Mechatronics and Automobiles" (No.XKQ2019031).

 DOI:10.1049/cje.2019.06.027

coarse detection is performed by the adaptive threshold to extract candidate targets from the "target-saliency" map. Finally, the real target region is identified by the characteristics of the cat's eye effect target, so the true targets are detected and the spurious objects are rejected. Experimental results indicate the effectiveness of the proposed method in this paper.

II. Salient Region Extraction and Segmentation

The laser imaging system for cat's eye effect targets detection can acquire active image sequence with pulsed laser and passive image sequence without pulsed laser respectively.

1. Saliency map computation based on the SDDF

Haralick's cubic facet model assumes that the underlying grey-level intensity surface can be approximated by a bivariate cubic function f[19] in each neighborhood of an image.

Under the cubic facet model, each surface facet centered about a given pixel is approximated by the bivariate cubic function in canonical form[19,22,12]

$$\begin{aligned} f(r,c) =& K_1 + K_2 r + K_3 c + K_4 r^2 + K_5 rc \\ &+ K_6 c^2 + K_7 r^3 + K_8 r^2 c + K_9 r c^2 + K_{10} c^3 \end{aligned} \tag{1}$$

where K_i, $i = 1, \cdots, 10$ are coefficients for the bivariate cubic function expressed in discrete orthogonal polynomials. The coefficients $K_1, \cdots, K_N$ are determined by the least-squares surface fitting and the orthogonal property of the polynomials

$$K_i = \frac{\sum\limits_{(r,c)\in S} g_i(r,c) I(r,c)}{\sum\limits_{(r,c)\in S} {g_i}^2(r,c)} \tag{2}$$

Eq.(2) shows that each fitting coefficient K_i can be computed individually as a linear combination of the intensity values $I(r,c)$. The weight associated with each $I(r,c)$ for the i-th coefficient is determined by

$$W_i = \frac{\sum\limits_{(r,c)\in S} g_i(r,c)}{\sum\limits_{(r,c)\in S} {g_i}^2(r,c)} \tag{3}$$

Evaluate the second row and column partial derivatives at the neighborhood center $(0,0)$, that is $r = 0$ and $c = 0$, yields the second-directional derivatives

$$\frac{\partial^2 f(r,c)}{\partial r^2} = 2K_4, \frac{\partial^2 f(r,c)}{\partial r \partial c} = K_5, \frac{\partial^2 f(r,c)}{\partial c^2} = 2K_6 \tag{4}$$

Therein, K_i are the coefficients which can be computed independently by convolving the image with the corresponding weight kernel computed by using Eq.(3).

According to the extremum theory, if the following conditions are satisfied

$$\begin{aligned} D_1 =& \frac{\partial^2 f}{\partial r^2} = 2K_4 < 0 \\ D_2 =& \frac{\partial^2 f}{\partial r^2}\frac{\partial^2 f}{\partial c^2} - \frac{\partial^2 f}{\partial r \partial r} = 4K_4K_6 - K_5^2 > 0 \end{aligned} \tag{5}$$

then the corresponding pixel $f(r,c)$ is the maximal extremum points, which means the possible targets pixels. Eq.(4) can be deduced from the formula of SDDF along direction vector l at pixel (x_0, y_0) as follows[23]:

$$\begin{aligned} \frac{\partial^2 f(x,y)}{\partial l^2}|_{(x_0,y_0)} =& [f_{xx}(x,y)\cos^2\alpha \\ &+ 2f_{xy}(x,y) \times \cos\alpha\cos\beta \\ &+ f_{yy}(x,y)\cos^2\beta]|_{x_0,y_0} \\ =& 2K_4(x_0,y_0)\cos^2\alpha \\ &+ 2K_5(x_0,y_0) \times \cos\alpha\cos\beta \\ &+ 2K_6(x_0,y_0)\cos^2\beta \ldots \end{aligned} \tag{6}$$

where α is the angle between l and the x-axis; β is the angle between l and the y-axis.

Since the difference images may be brought about disturbance for the saliency detection after the SDDF, we need to amend the SDDF maps by morphological processing[24]. Along with the increase of the distance between imaging system and target, the cat's eye effect target in the image usually has the properties of low SNR, small size and unavailable shape information.

2. The directional saliency map computation using SR approach

The SR method[21] has low computational complexity among most saliency detection approaches. And it needn't the prior knowledge for saliency detection. Then it is suitable to compute saliency maps for SDDF maps.

The spectral residual contains the innovation of an image. It serves as the compressed representation of a scene. Using Inverse Fourier Transform, we can construct the output image called the saliency map in spatial domain which contains primarily the nontrivial part of the scene. For better visual effects, we smooth the saliency map with a Gaussian filter $g(x,y)(\sigma = 3)$. In this letter, we briefly express the procedure of computing saliency maps using PFT as follows[21]:

$$\begin{aligned} A(f) =& \Re(F[I(x,y)]), \\ P(f) =& \Im(F[I(x,y)]), \\ L =& \log(A(f)), \\ R(f) =& L(f) - h_n(f) * L(f), \end{aligned}$$

$$S(x,y) = g(x,y) * F^{-1}[\exp(R(f) + P(f))] \quad (7)$$

where $I(x,y)$ denotes the image, F and F^{-1} denote the Fourier transform and the inverse Fourier transform, respectively; $g(x,y)$ is a 2-D Gaussian filer ($\sigma = 3$); and $S(x,y)$ is the saliency map.

3. Computing the "target-saliency" map by the saliency fusing method

After generating the directional saliency maps of different channels, the fused saliency map becomes a crucial step. In order to compute the fused saliency map effectively and quickly, an effective saliency fusing method to fuse all the directional saliency maps into a final "target-saliency" map is put forward.

Finally, the "target-saliency" map S_{end} is obtained as

$$S_{\text{end}} = \frac{2}{n} \sum_{i}^{\frac{n}{2}} S_i \cdot (\perp S_{n-i+1}) \quad (8)$$

where n represents the number of the directional saliency maps; S_i is the directional saliency map; $\perp S_{n-i+1}$ is the orthogonal directional saliency map of S_i; and S_{end} is the final fused map. In order to get the fine fused map, eight directional saliency maps are used to fuse the final saliency map. The effective saliency fusing method can inhibit the orientational clutters and noise, and heighten the isotropic cat's eye effect target.

4. Object segmentation using "target-saliency" maps

After the saliency map S_{end} is computed, we will segment it to extract the salient regions and the regions of interest, where the suspicious targets exist. With background suppressed and target enhanced by the above processing, the cat's eye effect target has the maximum intensity within its local region in an image. Given that, the threshold segmentation can be applied to detect real target. Two steps are designed to accomplish this:

Step 1: search the maximal value $S_{\text{end\_max}}$ and compute the average value $S_{\text{end\_mean}}$ of the saliency map S_{end}, then take them as references to set the threshold T:

$$T = \frac{kS_{\text{end\_max}} + S_{\text{end\_mean}}}{k+1} \quad (9)$$

where k is a constant. After a great of experiments, k is selected in the range $[2,6]$ for the image.

Step 2: As shown in Eq.(10), if the pixel is the brightest one among its 7×7 neighborhood W and its value is higher than T, the value is set as 1; otherwise as 0.

$$S_{\text{end}}(x,y) = \begin{cases} 1, & \text{if } (S_{\text{end}}(x,y) \geq S_{\text{end}}(x_w,y_w) \text{ and} \\ & \quad S_{\text{end}}(x,y) \geq T, (x_w,y_w) \in W) \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

It is worth pointing out that the size of W is determined by the size of the targets that are detected in real-life images. Generally speaking, the size of cat's eye effect targets in our imaging system is smaller than 100 pixels. Consequently, we should specify an average size of W as 7×7 in our experiments. After segmenting the "target-saliency" map, the candidate targets can be extracted.

5. Cat's eye effect target identification

After the candidate targets are selected, we should decide whether they belong to the target region or not. The salient region whose characteristics do not accord with the cat's eye effect target will be discarded. In that way we need to select the features of the cat's eye effect target to identify the salient regions in the difference images. Each of the features is calculated within a local region, using the dual rectangular window for every pixel in the difference image. An approximate target size is predetermined via prior knowledge of the range and field of view information. Subsequently, we take the center of the target region as center and expand the window circumferentially to get a larger region.

1) Feature extraction

Contrast mean difference (feature 1): Suppose that P represents the set of pixels of target region, $I(x,y)$ represents the difference image. Denote N_p and M_p as the number of the pixels and the mean pixel value of the target region in the difference image, respectively. Then we can get[25] :

$$M_p = \frac{1}{N_p} \sum_{(x,y) \in P} I(x,y) \quad (11)$$

Let B be the set of pixels of the background region, we can get the mean pixel value of such region:

$$M_B = \frac{1}{N_B} \sum_{(x,y) \in B} I(x,y) \quad (12)$$

where N_B is expressed as the number of the pixels of the background region. Contrast means the difference between target intensity and background intensity, and it is usually defined as follows:

$$F^1_{x,y} = C_t = \frac{M_P - N_B}{M_P + N_B} \quad (13)$$

where C_t is the contrast threshold, and it must be greater than zero as well. That is to say, the average gray value of the target region in the difference images is larger than the one of the background region. After a great experiments, C_t can be set 0.06 as a threshold. When C_t is smaller than the threshold, candidate target can be discarded.

Average gradient strength (feature 2): The gradient strength feature is used because man-made objects tend

to show sharper internal detail than natural objects. The average gradient strength of the target region in difference image is calculated as[25]:

$$F_{x,y}^2 = \frac{1}{n_i n} \sum_{(x,y)\in N_{in}(x,y)} \left(G_{in}^x(x,y) + G_{in}^y(x,y)\right) \quad (14)$$

where n_{in} is the number of pixels in $N_{in}(x,y)$, and $N_{in}(x,y)$ is the target region. $G_{in}^x(x,y)$ and $G_{in}^y(x,y)$ are the gradient of the difference image in the x-level and y-level direction, respectively.

Local variation (feature 3): While the gradient-strength feature is used to find rapid intensity variations among the immediate neighboring pixels, the local-variation feature is used to detect slow intensity variations over the local region. The local variation of the target region is calculated as[25]:

$$\begin{aligned} F_{x,y}^3 =& \frac{1}{n_{in}} \sum_{(x,y)\in N_{in}(x,y)} \Big| I(x,y) \\ & - \frac{1}{n_{in}} \sum_{(x,y)\in N_{in}(x,y)} I(x,y) \Big| \end{aligned} \quad (15)$$

Normalization of features: Each feature is normalized across the image, so that the feature value at each pixel represents the number of standard deviations that the pixel stands apart from the values for the same feature across the image. Thus the feature image for the m-th feature is normalized as

$$\begin{aligned} F_{x,y}^{m,N} =& \frac{F_{x,y}^m - \mu_m}{\sigma_m} \\ \mu_m =& \frac{1}{M} \sum_{n=1}^{M} F_{x,y}^{m,n} \\ \sigma_m =& \frac{1}{M} \sum_{n=1}^{M} (F_{x,y}^{m,n} - \sigma_m)^2 \end{aligned} \quad (16)$$

where M is the numbers of the candidate target, μ_m and σ_m are the mean value and variance.

2) Feature-based fusion and target identification

After normalization, the features from the difference image are adaptively combined by the expression:

$$G_{x,y}^n = \sum_{m=1}^{3} \omega_m F_{x,y}^{m,n}, n \in M \quad (17)$$

where ω_m is the weight of the m-th feature and is determined so that it is proportional to the significance for detection of the m-th feature. After a great experiments, ω_1, ω_2, ω_3 can be set as 0.2, 0.4, 0.4, respectively. In order to decide whether the candidate targets belong to the real target region or not, the $G_{x,y}^n$ with the maximum value is chosen as the real cat's eye effect target. Then the n-th ($n \in M$) candidate target can be selected and marked with red rectangle.

III. The Proposed Method for the Identification of Cat's Eye Effect Target

The block diagram of the proposed method is presented in Fig.1. The difference image is decomposed into eight directional channels by the SDDF, and directional saliency maps are computed by SR; then the "target-saliency" map is obtained by the saliency fusing method; after adaptive threshold segmentation, candidate targets from the "target-saliency" map can be extracted. Finally, the real target is identified by the characteristics of the cat's eye effect target. The detailed processing steps of CTRVA are described as follows:

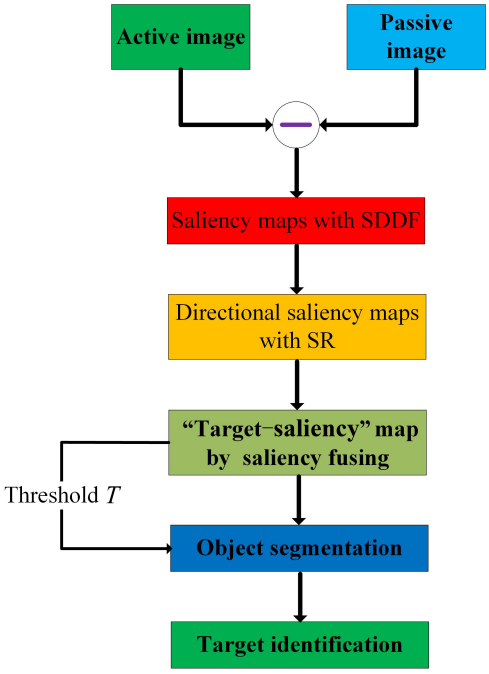

Fig. 1. The block diagram of the proposed method

Step 1: Eight directional channels are selected to process the difference image. Channel one to eight can be set successively as $(\alpha = 90^0, \beta = 0^0)$,$(\alpha = 45^0, \beta = 45^0)$,$(\alpha = 180^0, \beta = 90^0)$,$(\alpha = 225^0, \beta = 135^0)$, $(\alpha = 315^0, \beta = 225^0)$, $(\alpha = 270^0, \beta = 180^0)$,$(\alpha = 135^0, \beta = 45^0)$,$(\alpha = 0^0, \beta = 90^0)$.

Step 2: Calculate K_4, K_5, K_6 by convolving the original image with W_4, W_5, W_6 (using Eq.(3)). With the calculated second directional derivative of the fitted surface, the decision is made whether the neighborhood center pixel is the maximal extremum point or not by using the maximal extremum conditions in Eq.(5). The possible target pixel can be found by this.

Step 3: Compute SDDF maps for each directional channel by the SDDF, and then morphological method is

employed to deal with the filtered image in all directions.

Step 4: Compute directional saliency maps for each channel by the SR method.

Step 5: The "target-saliency" map is acquired by fusing directional saliency map using the designed saliency fusing method.

Step 6: The candidate targets can be extracted by segmenting the "target-saliency" map.

Step 7: The real target region is identified by the characteristics of the cat's eye effect target.

IV. Experiments

To evaluate the detection performance and robustness of the presented method, CTRVA is tested and also compared with two-state-of-the-art method. and multiple scenes gathered under different conditions are selected.

1. Salient region extraction

We carry out experiments to investigate the effect of the proposed salient region extraction approach. The original active and passive images of the two scenes are shown in Fig.2. Therein, the cat's eye effect target is marked with red rectangle, the distance of the cat's eye effect target and the laser imaging system is about 197m and 856m in first scene and second scene, respectively.

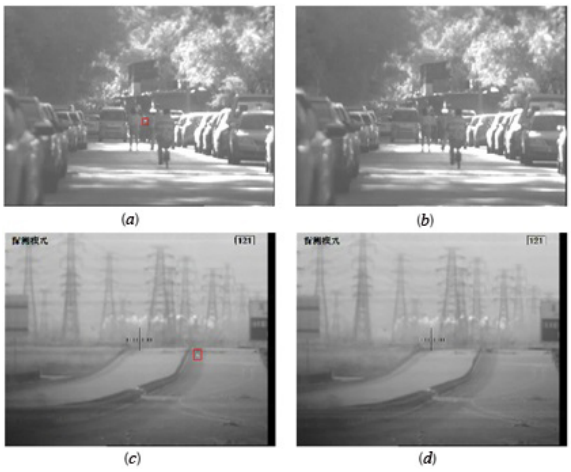

Fig. 2. Original images. (a) and (c) are the original active image; (b) and (d) are the original passive image

The salient region extraction results after binarize are shown in Fig.3 (the spurious objects are marked with green circles). From Fig.3, we can see that the method presented in Ref.[26] has poor salient regions extraction results. That is because this method is easily influenced by the noise existing in the difference image or it will fall if the background is more distinct than the target of interest. Although the background has heavy noise or clutter, our approach obtains better extraction results of salient regions.

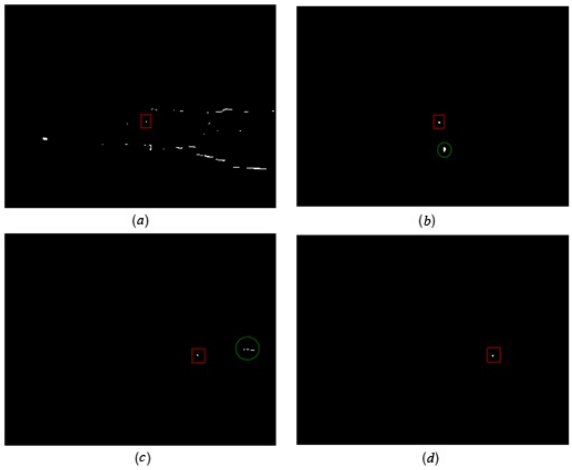

Fig. 3. The salient region extraction results

2. Experiments of the cat's eye effect target identification

For the target detection, finding the potential target regions is the first step of salient region extraction. That is to say, it locates the targets coarsely. Take Fig.3(b) for example, it contains two salient regions. Actually, there is only one real cat's eye effect target in the difference image. Thus, it needs further process to identify the real target from the potential target regions. By using the cat's eye effect target identification approach, the real target can be screened out, and the spurious object can be discarded.

To further validate the detection effect and robustness of the proposed method, many scenes are selected in different conditions to test our method in Fig.4. There are various conditions including illumination, background, distance and weather. All the images are captured in the daytime, at dusk and in the evening. Fig.4(a), Fig.4(b) and Fig.4(c) are taken in the morning, noon and evening respectively. The backgrounds are cars and ambulatory people beside the trees and wall. Obviously, it is raining in Fig.4(c); Fig.4(d) and Fig.4(e) show a crowded street in the daytime. As can be seen from the figures, our method is robust to different laser transmit powers; Fig.4(f) is collected on the road with moving vehicles at dusk. It was a foggy day; Fig.4(g) is captured on the road covered with snow. The distance between detection device and targets varies in different scenes, in the range of 100-900 m.

In addition, we compare it with the three-state-of-the-art methods. Ref.[10] is easily confused by the complex background clutter, while Ref.[11] is too time-consuming, and cannot meet the demand of real-time target detection as well. Although Ref.[12] has better real-time processing capability, there is a relatively high false alarm probability. For instance, the false alarm probability is high in Fig.5. It can be seen that the reflectors and car lights are not removed by the MSPF method in Fig.5(a) and Fig.5(b) respectively. The first

column and the second column display the active images and the recognition results by our method, the third shows the recognition results using MSPF. The true targets are marked in red rectangles and the false in green circles.

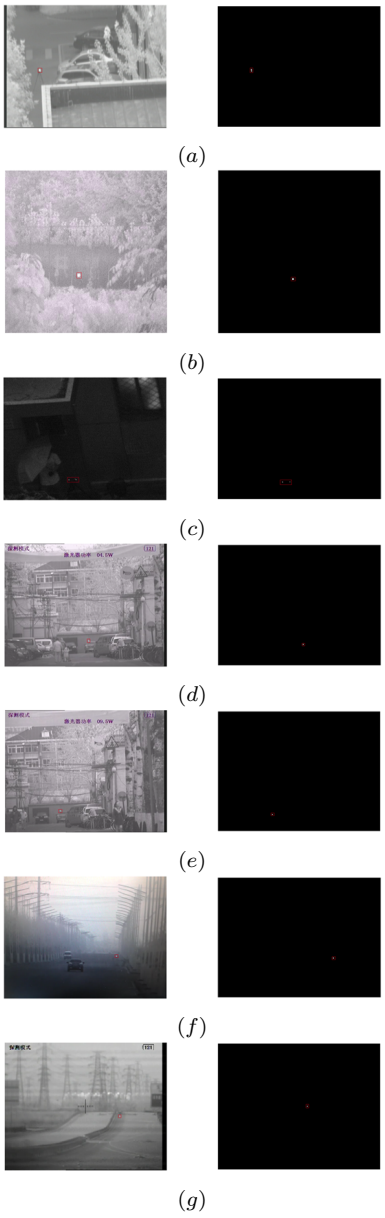

Fig. 4. The active images and processing results by our method

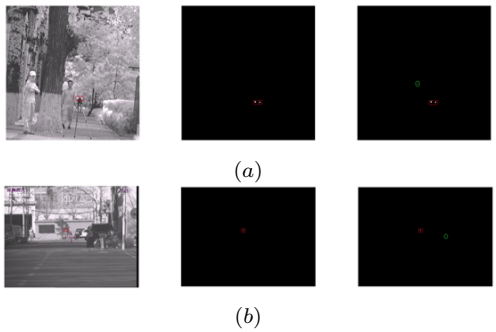

Fig. 5. Comparison results between MSPF and our method

In order to give the quantitative evaluation and further prove the superiority of our method, the comparison among the presented method and other three methods is given in Table 1. It is obvious that our proposed method has higher detection rate and lower false alarm rate.

Table 1. The average false alarm rate and the average detection rate of different detection methods for the multiple scenes

Methods	False alarm rate	Detection rate
The proposed method	1.21 %	96.7 %
Shape discrimination method[10]	8.73 %	43.6 %
SFDC method[11]	4.91 %	71.4 %
MSPF method[12]	5.01 %	93.2 %

The detection probability P_d and P_f false alarm probability are used to measure the performance:

$$P_d = \frac{N_d}{N_t} \tag{18}$$

where N_d denotes the number of true targets detected in image; N_t is the number of true targets existing in image.

$$P_f = \frac{N_f}{N_r} \tag{19}$$

where N_f denotes the number of false targets detected in image; and N_r is the number of targets detected in image.

In addition to the detection performance, the computational costs of our method and other method are reported in Table 2. It is evident that the proposed method yields excellent performance for real-time target detection and is less than the three-state-of-the-art methods in the average detection time. In a word, our method can detect and identify the cat's eye effect target accuracy, robustness and speed compared to other methods.

Table 2. Computational cost comparison among the proposed method and other three methods

Methods	Average detection time (s)
The proposed method	1.939802
Shape discrimination method[10]	2.236531
SFDC method[11]	5.822103
MSPF method[12]	1.274655

V. Conclusions

The cat's eye effect target recognition method based on visual attention is presented in this paper. Parallel processing enables it to make processing whatever LR or HR images fast and suit for hardware design such as FPGA. Experimental results indicate the effectiveness and good performance of the proposed method in this paper. And it achieves satisfying performance including efficiency and the high detection rate and the lower false

alarm rate. Meanwhile, it has lower computational cost and can be used for real-time detection.

Acknowledgement: Thanks to Li Li from Beihang University for providing the experimental data and Rui Hou from Institute of Information Engineering, Chinese Academy of Sciences for helpful comments.

References

[1] C. L. Ge, *et al.*, "Target classification with cat eye effect", *High Power Laser and Partical Beams*, Vol.15, No.7, pp.632–634, 2003.

[2] Y. Zhao, *et al.*, "Three-dimensional analytical formula for oblique and off-axis gaussian beams propagating through a cat's eye optical lens", *Chinese Physics Letters*, Vol.27, No.3, pp.034101, 2010.

[3] C. Lecocq, *et al.*, "Sight laser detection modeling", *Proc. of the SPIE*, Vol.5086, pp.280–286, 2003.

[4] M. Gong and S. F. He, "Periodicity analysis on cat-eye reflected beam profiles of optical detectors", *Optical Engineering*, Vol.56, No.5, pp.053110, 2017.

[5] D. S. Wu, *et al.*, "Detection of cat-eye effect echo based on unit APD", *Proc. of SPIE 10153, Advanced Laser Manufacturing Technology*, Vol.101530J, 2016.

[6] F. Qian, *et al.*, "Fast cat's eye target detection system based on DSP+FPGA architecture," *Journal of Optoelectronics. Laser*, Vol.27, No.8, pp.863–869, 2016.

[7] M. H. Liu, Q. S. Chen, X. Y. Li and X. Li, "Comparative study on the numeric models of cat's eye echo power," *Laser Journal*, Vol.38, No.6, pp.12–15, 2017.

[8] Y. X. Cai, M. Z. Ouyang and Y. G. Fu, "Research on the mechanism of laser active detection", *Laser & Infrarel*, Vol.48, No.4, pp.451–457 2018.

[9] D. P. Casasent, *et al.*, "SAR ship detection using new conditional contrast box filter", *Proc. of the International Society for Optical Engineering*, Vol.3721, pp.274-284, 1999.

[10] L. Tong, X. Jiang and X. Song, "Target detection based on laser imaging with cat eye effect", *Laser & Infrared*, Vol.39, pp.982–985, 2009.

[11] X. Ren and L. Li, "Recognizing cat's eye targets with dual criterions of shape and modulation frequency", *Chinese Optics Letters*, Vol.9, No.4, pp.041101, 2011.

[12] L. Li, J. Ren and X. Wang, "Fast cat-eye effect target recognition based on saliency extraction", *Optics Communications*, Vol.350, No.1, pp.33–39, 2015.

[13] L. Itti, C. Koch and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.20, No.11, pp.1254–1259, 1998.

[14] B.C. Ko and J. Nam, "Object-of-interest image segmentation based on human attention and semantic region clustering", *J. Opt. Soc.Am. A*, Vol.23, No.10, pp.2462–2470, 2006.

[15] Y. Xu, *et al.*, "Salient target detection based on pseudo-Wigner-Ville distribution and Rényi entropy", *Opt. Lett.*, Vol.35, No.4, pp.475–477, 2010.

[16] R. Rensink. "Seeing, sensing, and scrutinizing", *Vision Research*, Vol.40, No.10-12, pp.1469–1487, 2000.

[17] D. Walther and C. Koch, "Modeling attention to salient proto-objects", *Neural Networks*. Vol.19, No.9, pp.1395–1407, 2006.

[18] D. Walther, L. Itti, M. Riesenhuber, *et al.*, "Attentional selection for object recognition a gentle way", *Lecture Notes in Computer Science*, Vol.2525, No.1, pp.472–479, 2002.

[19] R. Haralick, "Digital step edges from zero crossing of second directional derivatives", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. PAMI-6, No.1, pp.58–68, 1984.

[20] X.Q. Luo and X.J. Wu, "A novel fusion detection algorithm for infrared small targets", *Proc. of 2009 Third International Conference on Intelligent Information Technology Application*, Shanghai, China, pp.427–430, 2009.

[21] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach", *Proc. of 2007 IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, MN, USA, pp.1–8, 2007.

[22] G. D. Wang and Ch. Y. "Chen and X. B. Shen, Facet-based infrared small target detection method", *Electronics Letters*, Vol.41 No.22, pp.1244–1246, 2005.

[23] S. X. Qi, J. Ma and C. Tao, "A robust directional saliency-based method for infrared small-target detection under various complex backgrounds", *IEEE Geoscience and Remote Sensing Letters*, Vol.10, No.3, pp.495–499, 2012.

[24] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, New York, USA, 1982.

[25] Kwon, Der and Nasrabadi, "Adaptive multisensory target detection using feature-based fusion", *Opt. Eng.* Vo.41, No.1 pp.69–80, 2002.

[26] R. Achanta, S. Hemami, F. Estrada, *et al.*, "Frequency-tuned salient region detection", *Proc. of 2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, USA, pp.1597–1604, 2009.

WANG Xingbin was born in 1987. He received the B.E. degree in electronic engineering from Beihang University. He is a Ph.D. candidate of Institute of Information Engineering, Chinese Academy of Science. His research interests include computer architecture, machine learning, AI accelerator and secrity. (Email: wxbbuaa2011@163.com)

ZHANG Jun (corresponding author) was born in 1985. He received the Ph.D. degree in computer architecture from Institute of Computing Technology, CAS. He is a assistant professor of Hubei University of Arts and Science. His research interests include computer architecture security and machine learning. (Email: zhangjunhbxf@163.com)