Recognizing cat-eye targets with dual criterions of shape and modulation frequency

Ximing Ren (任熙明)* and Li Li (李 丽)

School of Electronics and Information Engineering, Beihang University, Beijing 100191, China

**Corresponding author: rxmbit@gmail.com*

Received September 19, 2010; accepted November 19, 2010; posted online March 28, 2011

We present an image recognition method to distinguish targets with cat-eye effect from the dynamic background based on target shape and modulation frequency. Original image sequences to be processed are acquired through an imaging mechanism that utilizes a pulsed laser as active illuminator and an industrial camera as detection device. There are two criterions to recognize a target: one exploits shape priors and the other is the active illuminator's modulation frequency. The feasibility of the proposed method and its superiority over the single criterion method have been demonstrated by practical experiments.

OCIS codes: 110.2970, 100.3008, 110.4155, 100.2000.

doi: 10.3788/COL201109.041101.

The cat-eye effect in imaging systems such as eyes, cameras, and optical sights is a special phenomenon of reflection that occurs when the incoming radiation is retro-reflected by a certain optical system. Estimating the position and orientation of cat-eye effect targets is an important technique that can be widely applied in the recognition of passive optical systems, seeker identification, and free-space communication, etc. Some researchers have conducted investigations exploiting the cat-eye effect[1−4]. One method for detecting the cat-eye effect target is image object recognition. Currently, there are three strategies of major theory for image object recognition[5,6]: through geometric correspondence and pose consistency, through template matching via classifiers, and through correspondence search to establish the presence of suggestive relations between templates. Therefore, we cannot efficiently identify a cat-eye effect target in images without *a priori* knowledge of them. Under the condition of laser illumination, there are two features of the cat-eye effect target. One is that the cat-eye effect target's region is brighter than its surroundings; the other is that its shape is almost circular, especially for most optical lenses. A recognition method based on roundness and eccentricity was put forward in Ref. [7]. However, the method based only on shape and gray priors cannot precisely recognize cat-eye effect targets from the dynamic background. Referring to modulation information, cat-eye modulation retro-reflectors combine passive optical retro-reflectors with electro-optic modulators to allow long-range, free-space optical communication with a laser and pointing-acquisition-tracking system[3]. Referring to the modulating retro-reflector link of the above system, active illumination modulation information is added to the traditional shape-and-gray-priors-based recognition method in order to detect cat-eye effect targets in this letter.

The object recognition method based on shape priors and modulation frequency is proposed in this letter. To demonstrate the feasibility of this method, an imaging system consisting of a pulsed laser and an industrial camera is developed. The imaging mechanism of the system utilizes the cat-eye effect principle to detect the binocular. We use the pulsed laser as modulation illuminator and utilize the industrial camera to acquire image sequences. The time sequence of the recognition process is shown in Fig. 1.

The received energy per pixel can be described as

$$E_m = \frac{\gamma \exp(-\alpha z) S_{\mathrm{r}}}{z^2} P\big(t - \frac{2z}{c}\big) G(t), \tag{1}$$

where γ is the target reflectivity, m is a certain label in the image sequence, z is the detection range, c is the speed of light, S_{r} is the area of the receiver aperture, $\exp(-\alpha z)$ is the atmospheric attenuation, $P(t)$ is the laser illuminating power in the field of view per pixel, and $G(t)$ is the periodic camera shutter, which has a 50% duty cycle and 60-Hz square signal, and can be regarded as the synchronous Nyquist sampling sequence.

A sequence of E_m represents a series of discrete intensities for a pixel in the laser-covering region. Hence, the distribution of E_m sequences in the timeline can indirectly indicate the laser pulse frequency, called modulation frequency, mirroring the frequency feature of the laser pulse.

In the experiments, we set the modulation frequency f_{L} of the pulsed laser much less than the frequency f_{C} of the camera shutter to follow the Nyquist condition (i.e., $f_{\mathrm{L}} = 1$ Hz, $f_{\mathrm{C}} = K f_{\mathrm{L}} = 60\ f_{\mathrm{L}}$, $K = 60$). Active and passive image sequences can then be collected during the high level and low level of the periodic camera shutter, respectively. Therefore, the number of images acquired

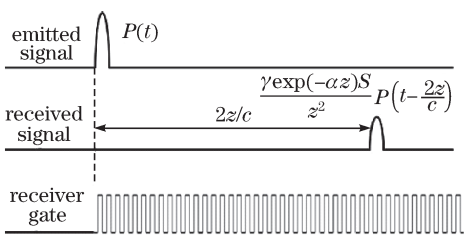

Fig. 1. Time sequence of detection process.

by the industrial camera during T is $f_{\rm C}T$, and they are expressed as follows:

$$\{I(x_i, y_j, t_0),\ I(x_i,\ y_j,\ t_1), \cdots, I(x_i,\ y_j,\ t_{n-1}),$$
$$i \in (0,\ 1, \cdots, N-1),\ j \in (0,\ 1, \cdots, M-1)\},$$

where n is the total number of images; N and M are the maximum length and width of pixels per image, respectively. Gray-scale images are represented as $I(x_i, y_j, t)$, with x representing the column number of the image and y representing the line number of the image (a pixel in the image expressed as $(x,\ y)$), and t representing different acquisition moments. Since the number of 50% duty cycle laser pulse during T is $T/(1/f_{\rm L})$, then the number of frames per laser pulse is $Tf_{\rm C}/[T/(1/f_{\rm L})] = K$, and the proportion of the images corresponding to high level and low level occupies a half, respectively. Therefore, two images whose interval is $(\frac{1}{2}+u)K$ frames must be active and passive images reciprocally. Note that u is a positive integer.

The proposed method is called shape-frequency dual criterions (SFDC) method using the acquired image sequences as original images for processing because the shape priors are the shape criterion, and the modulation frequency is the frequency criterion. The details are given as follows. And the relevant flowchart is shown in Fig. 2.

Step 1: Frame m subtracts Frame $m + (\frac{1}{2}+u)K$, resulting in the first difference image. Frame m subtracts Frame $m - (\frac{1}{2}+u)K$, resulting in the second difference image. An adaptive threshold is used to change the first difference image to be the first binary image. Another adaptive threshold is used to change the second difference image to be the second binary image. After that, the first binary image AND the second one results in a dynamic-background-subtracted binary image. Specifically, each adaptive threshold[8,9] satisfies

$$A_{\rm th} = \frac{\sum_{x=1}^{N}\sum_{y=1}^{M} I_{\rm th}(x,y)}{N_{\rm th}}, \tag{2}$$

where $I_{\rm th}(x,y)$ represents the gray value in excess of $A_{\rm th}$, $N_{\rm th}$ represents the corresponding number of pixels, and $A_{\rm th}$ is an initial threshold determined by

$$A_{\rm th} = \frac{\sum_{x=1}^{N}\sum_{y=1}^{M} e_{xy} I(x,y)}{\sum_{x=1}^{N}\sum_{y=1}^{M} e_{xy}}, \tag{3}$$

where e_{xy} represents the maximum one between vertical gradient and horizontal gradient per pixel.

CLOSE operation of gray-scale morphology is applied to remove some stray pixels in the dynamic-background-subtracted binary image. All segmentation regions whose total number is S are labeled by different numbers. Note that $S \ll N \times M$.

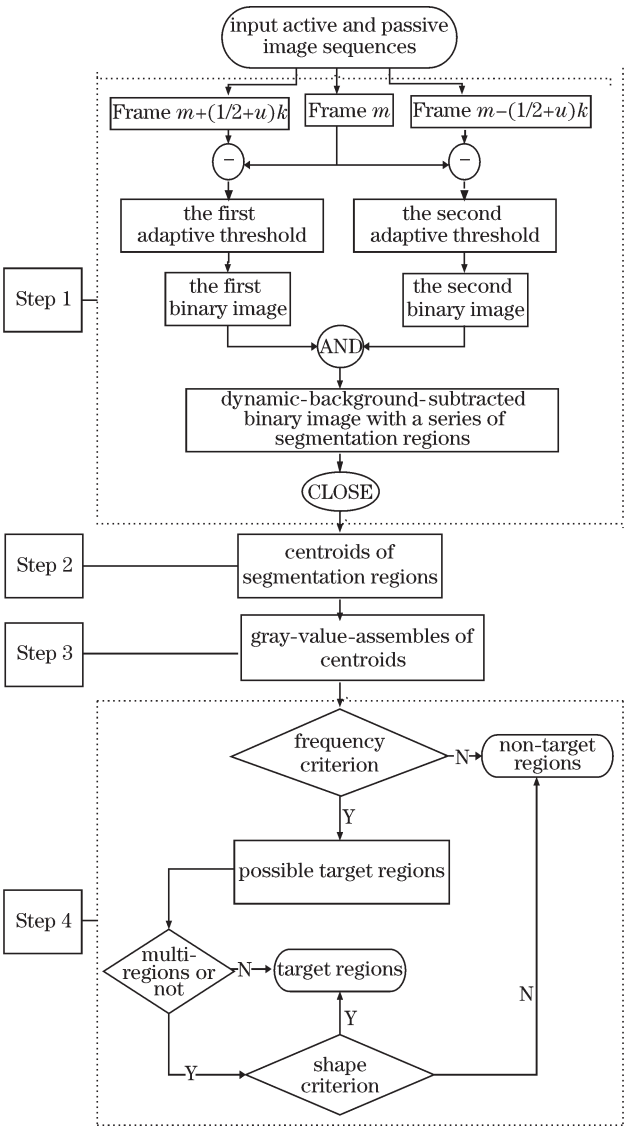

Fig. 2. Flowchart of the recognition method.

Step 2: The centroid formulas are written as

$$x_{\rm c} = \frac{\sum_{x=1}^{N'}\sum_{y=1}^{M'} x}{N'M'}, \tag{4}$$

$$y_{\rm c} = \frac{\sum_{x=1}^{N'}\sum_{y=1}^{M'} y}{N'M'}, \tag{5}$$

where N' and M' are the maximum length and width pixels of each segmentation region, respectively. The centroid of each segmentation region is calculated and their coordinate points are recorded as $(x_{\rm c_1},\ y_{\rm c_1})$, $(x_{\rm c_2},\ y_{\rm c_2}), \cdots, (x_{\rm c_S}, y_{\rm c_S})$.

Step 3: Gray values corresponding to $(x_{\rm c_1},\ y_{\rm c_1})$, $(x_{\rm c_2},\ y_{\rm c_2}), \cdots, (x_{\rm c_S},\ y_{\rm c_S})$, which are acquired from $f_{\rm C}T$ images, constitute S gray-value-assembles. They are recorded as

$$A_1 = \{I_i(x_{\rm c_1},\ y_{\rm c_1}),\ i \in f_{\rm C}T\},$$

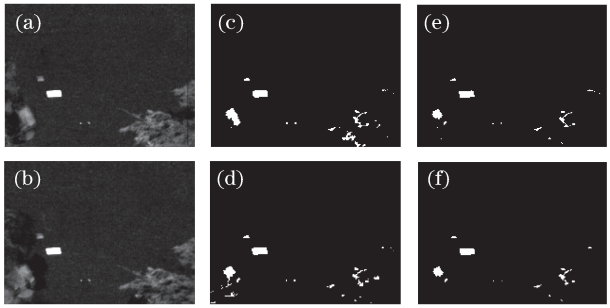

Fig. 3. Two difference images (640×480 pixels): (a) the first difference image: (b) the second difference image. Two binary images (640×480 pixels); (c) the first binary image based on $A_{\mathrm{TH}} = 22$; (d) the first binary image based on $A_{\mathrm{TH}} = 21$. Processing effect of binary image (640×480 pixels): (e) original dynamic-background-subtracted binary image; (f) processed dynamic-background-subtracted binary image.

$$A_2 = \{I_i(x_{c_2},\ y_{c_2}),\ i \in f_C T\},\ \cdots,$$
$$A_S = \{I_i(x_{c_S},\ y_{c_S}),\ i \in f_C T\}.$$

Step 4: Each gray-value-assemble: A_1, A_2,$\cdots$, A_S expanded in the timeline is a gray value's change distribution. If the demodulated frequency of a gray value's change distribution matches the pulsed laser's modulation frequency, f_{L} (i.e., frequency criterion), then the interest target may be in the corresponding region. Other regions are thus jamming ones. If there are more than one regions where the interest target may be in, shape features including roundness and eccentricity are used to take out the optimum regions. The roundness can be written as follows:

$$M_{\mathrm{r}} = \frac{4\pi a}{g^2}, \tag{6}$$

where a represents each segmentation region's area and g represents the perimeter. The eccentricity can be written as:

$$M_{\mathrm{e}} = \frac{Z_{\mathrm{major}}}{Z_{\mathrm{min}}}, \tag{7}$$

where Z_{major} represents the major axis and Z_{min} represents the minor axis in terms of each segmentation region. The comprehensive metric of shape features, namely, shape criterion, is defined as

$$M_{\mathrm{s}} = |1 - M_{\mathrm{r}}| + |1 - M_{\mathrm{e}}|. \tag{8}$$

Following the flowchart shown in Fig. 2, during the recognition, the gray priors are prioritized, i.e., target region with high gray values (in the appropriate detection range, the gray values of target region are almost saturate while it is not prone for the background), followed by the modulation frequency information, then the shape priors. Each obtained adaptive threshold satisfies the gray priors. AND operation is used to remove most of the different dynamic background between two difference images.

Experiments were performed to evaluate practicability of the proposed cat-eye effect target-recognition method through original images with binocular and dynamic background. The image sequence is acquired as

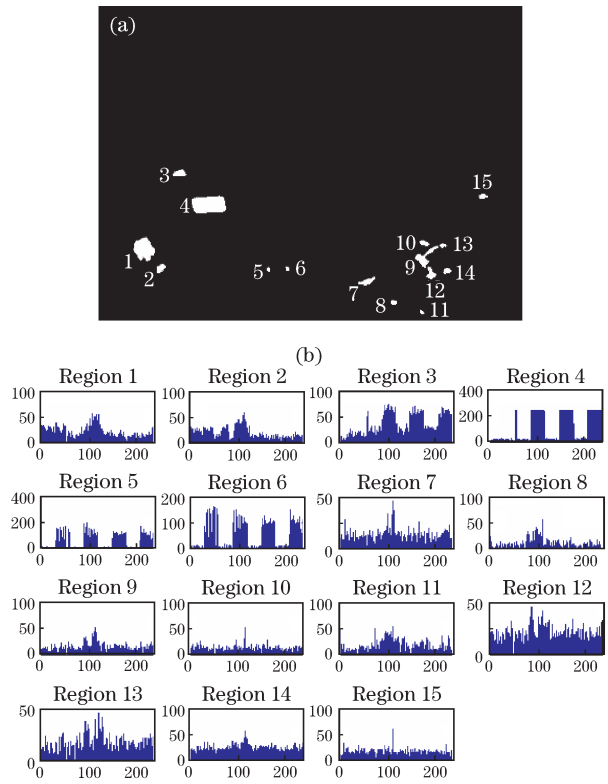

Fig. 4. (a) Processed dynamic-background-subtracted binary image (640×480 pixels) with 15 segmentation regions (Regions 5 and 6 are target regions); (b) Frequency distributions of centroids (x axis represents gray values; y axis, which represents image sequence labels, is equivalent to timeline).

original images. The dynamic background mainly includes shaky trees and an ambulatory person. Based on the above image sequence, Frame 115 subtracts Frame 145, resulting in the first difference image shown in Fig. 3(a). Frame 115 subtracts Frame 85, resulting in the second difference image shown in Fig. 3(b). An adaptive threshold $A_{\mathrm{TH}} = 22$ is used to change the first difference image to be the first binary image shown in Fig. 3(c), and another adaptive threshold $A_{\mathrm{TH}} = 21$ is used to change the second difference image to be the second binary image shown in Fig. 3(d). Segmentation regions in the lower-right and lower-left areas between Figs. 3(c) and (d) are obviously different because trees are shaking and the person is moving in the scene during image acquisition. To abate the jamming dynamic background, the first binary image AND the second one results in the dynamic-background-subtracted binary image shown in Fig. 3(e). CLOSE operation is applied to process the dynamic-background-subtracted binary image. The relevant result is shown in Fig. 3(f). There are 15 segmentation regions shown in Fig. 4(a). Figure 4(b) shows the frequency distributions of the gray value's change corresponding to each segmentation region's centroid. The centroids' coordinates are shown in Table 1.

We then find that the frequencies corresponding to Regions 3, 4, 5, and 6 are identical. They match the modulation frequency of the pulsed laser $f_{\mathrm{L}} = 1$ Hz while those in the rest regions do not. Each image in the image sequence corresponds to each acquisition moment; thus, using the change of the image sequence labels as

Table 1. Shape Feature Values Corresponding to 15 Segmentation Regions

Region	Centroid	Roundness	Eccentricity	Metric
1	(69,371)	0.69	1.07	0.38
2	(94,400)	0.88	1.1	0.23
3	(122,256)	0.75	2.13	1.37
4	(166,303)	0.72	2.27	1.55
5	(257,402)	1.36	1.00	0.36
6	(285,401)	1.18	1.00	0.18
7	(406,420)	0.40	6.00	5.59
8	(446,453)	1.03	1.17	0.19
9	(493,384)	0.24	6.50	6.26
10	(492,363)	0.82	2.40	1.58
11	(488,468)	0.89	1.00	0.11
12	(502,409)	0.46	8.00	7.54
13	(521,366)	0.75	3.33	2.58
14	(527,404)	0.85	1.50	0.65
15	(581,291)	0.60	5.00	4.40

Table 2. Recognition Rates of MF, RE, and SFDC Methods (%)

Experiment	MF	RE	SFDC
1	3	48	97
2	0	19	25
3	12	21	95

timeline in Fig. 5(b) is reasonable. If the modulation frequency is purely used as a criterion to detect the target, Regions 3, 4, 5, and 6 are all determined as possible target regions. However, the two regions belonging to the cat-eye effect binocular cannot be determined among the above four regions. In other words, the criterion based only on frequency feature is invalid.

In this case, to lock the target regions, another criterion is required, namely, exploiting the shape feature of the target. In terms of Eqs. (6) and (7), the ideal roundness and ideal eccentricity are both 1, meaning that the ideal comprehensive metric is 0 according to Eq. (8). In other words, the comprehensive metrics of target regions are close to 0, whereas others are apart from 0. Table 1 shows that the roundness, eccentricity, and comprehensive metrics of shape features in Regions 5 and 6 are smaller than those in Regions 3 and 4. Therefore, the location of the cat-eye effect binocular can be determined in Regions 5 and 6 due to the reliability of the dual criterions based on shape and frequency features.

However, if the recognition method based only on roundness and eccentricity is used at the outset, it is manifest to determine Regions 6 and 11 as target regions according to Table 1, due to the relatively smaller comprehensive metrics of shape feature in these two regions. They are not, however, the regions of target. In other words, the criterion based only on shape features is also invalid.

To verify superiority of SFDC, a comparison of the recognition rate of SFDC, roundness and eccentricity (RE in the table), and modulation frequency (MF in the table) method for three experiments (Experiments 1, 2, and 3) is shown in Table 2. Note that, 100 active-and-passive image pairs are used to get 100 difference images during three experiments. The results of Experiments 1 and 3 show that the recognition rate of the SFDC method is much higher than that of the other two methods.

However, the low recognition rate in Table 2 for Experiment 2 illustrates that all the methods based on shape and/or frequency features are invalid. This is because for Experiment 2, there are circular objects with strong reflectivity in the background which are maintained in the difference image. Accordingly, the frequency corresponding to these circular objects is similar to the modulation frequency. Therefore, purely based on modulation frequency as criterion, there are more than two frequency-matching segmentation regions in the binary image. Furthermore, the circular feature of the cat-eye effect binocular is not outstanding. Specifically, roundness and eccentricity of target regions are not superior to those of background regions.

In conclusion, a pulsed laser as active illuminator and an industrial camera as detection device comprise an imaging mechanism to acquire active and passive images as original image sequences. The target reflectivity is higher compared with the background, the gray values of target regions are higher, and the frequency distributions of relevant centroids match the modulation frequency better. The circular shape of the target is also more distinct, and it is easier to lock the target regions among frequency matching regions. Overall, to recognize cat-eye effect target from the dynamic background, experimental results show that the dual-criterion method is more robust than the single-criterion method. However, if circular objects with strong reflectivity exist in the static or dynamic background, both the single-criterion method and the proposed method are inferior. Meanwhile, the proposed method is limited when detecting fast moving cat-eye effect targets because these targets are difficult to capture in image sequences from a fixed field of view. One solution to this problem relies on the expensive ultra-high-speed camera, and the other solution—independent of high performance hardware—is a subject of future study.

References

1. C. Lecocq, G. Deshors, O. Lado-Bordowsky, and J. L. Meyzonnette, Proc. SPIE **5086,** 280 (2003).
2. D. H. Titterton, Proc. SPIE **5615,** 1 (2004).
3. W. S. Rabinovich, P. G. Goetz, R. Mahon, L. Swingen, J. Murphy, M. Ferraro, H. R. Burris, Jr., C. I. Moore, M. Suite, G. C. Gilbreath, S. Binari, and D. Klotzkin, Opt. Eng. **46,** 104001 (2007).
4. R. Mahon, W. S. Rabinovich, M. Plett, H. R. Burris, Jr., M. S. Ferraro, W. T. Freeman, C. I. Moore, J. Murphy, M. Stell, and M. R. Suite, Opt. Eng. **47,** 046002 (2008).
5. D. A. Forsyth and J. Ponce, *Computer Vision*: *A Modern Approach* (Prentice-Hall, Englewood Cliffs, 2001).
6. W. Yuan and X. Bai, Acta Opt. Sin. (in Chinese) **29,** 2158 (2009).
7. L. Tong, X. Jiang, X. Song, X. Wang, and P. Long, Laser Infrared (in Chinese) **39,** 982 (2009).
8. W. Liu and L. Wu, J. Infrared Millim. Waves (in Chinese) **15,** 257 (1996).
9. Z. Pan and Y. Wu, Acta Opt. Sin. (in Chinese) **29,** 2115 (2009).