Optics Communications 350 (2015) 33–39

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

Fast cat-eye effect target recognition based on saliency extraction

Li Li, Jianlin Ren*, Xingbin Wang

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

ARTICLE INFO

Article history:
Received 4 March 2015
Received in revised form
24 March 2015
Accepted 25 March 2015
Available online 27 March 2015

Keywords:
Cat-eye effect target
Visual attention
Saliency map
Parallel processing

ABSTRACT

Background complexity is a main reason that results in false detection in cat-eye target recognition. Human vision has selective attention property which can help search the salient target from complex unknown scenes quickly and precisely. In the paper, we propose a novel cat-eye effect target recognition method named Multi-channel Saliency Processing before Fusion (MSPF). This method combines traditional cat-eye target recognition with the selective characters of visual attention. Furthermore, parallel processing enables it to achieve fast recognition. Experimental results show that the proposed method performs better in accuracy, robustness and speed compared to other methods.

1. Introduction

Imaging systems such as cameras and optical sights, when illuminated by a laser, could often reflect back the incoming radiation much higher than the diffuse targets [1–3]. This is the phenomenon of “Cat-eye Effect”. Cat-eye effect targets recognition is an important issue that can be widely applied in the recognition of passive optical systems, seeker identification [4], and free-space communication [5], etc.

Over the past decade, many researchers have paid attention to the recognition of cat-eye effect targets, and then great progress has been made in this field. A recognition method based on roundness and eccentricity was put forward in Ref. [6]. The method based only on shape and gray priors work well in static background but proves less efficient in recognizing cat-eye targets from dynamic background. Another method based on shape-frequency dual criterions (SFDC) method [7] was put forward to improve the identification probability of the cat-eye effect targets in complicated backgrounds. By demodulating frequency and identifying the shape feature of targets, the cat-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. The compressive sensing (CS) method for recognizing cat-eye effect targets was presented in [8] with less data storage size. However, this approach is not applicable in engineering field since CS itself is far from a well-developed theory. The error rate of CS reconstruction remains a negative factor to small object recognition for it increases the probability of false recognition.

Over the past twenty years, using human visual characteristics to recognize target is under research. In 1998, Itti first proposed a visual attention model [9]. Subsequently, much effort has been devoted to detecting the target from visible images by visual attention mechanism [10,11]. Cat-eye targets, usually emerging in images as tiny, circular and bright objects, are “salient” compared to the backgrounds. Motivated by this, we put forward a multi-channel saliency map fusion method to recognize cat-eye targets. The initial image pairs are processed in eight directional channels to obtain directional saliency maps, followed by a new process for fusion. After thresholding, the targets are recognized. MSPF can run in either serial mode or parallel mode. It should be noted that our method is superior in processing High Resolution (HR) images. HR images carry more information and is useful in increasing recognition probability. However, it will unavoidably be more time-consuming as the pixels in images increase. So parallel processing is implemented to shorten running time.

In Section 2, the related work about saliency extraction will be shown. In Section 3, our recognition method is described in detail. In Section 4, experimental results are displayed.

2. Spectral residual method for saliency extraction

Visual attention is an important characteristic of human visual system (HVS), and it can make targets stand out relatively to its neighbors and thus capture our attention.

Among the applications developed from visual attention, saliency map extraction which was first put forward by Itti [9] is the most representative one. Saliency measures the contrast between

* Corresponding author.
E-mail address: renjianlin3493@163.com (J. Ren).

http://dx.doi.org/10.1016/j.optcom.2015.03.065
0030-4018/

one region to its surroundings. When a region differs significantly to its environment, it has high saliency. A great many approaches are proposed to extract salient regions in an image and most use some basic features like intensity, color to construct saliency maps. Researchers also put forward other significant analysis method such as graphic theory and spectral residual [12] (SR) approach.SR has low computational complexity, and does not require a prior knowledge while extracting target saliency maps. Therefore, we select SR to extract saliency maps of every channel in this paper. Firstly the theorem of SR approach will be described briefly.

Given an input image, the information required to be processed is as follows:

$$H(R(f)) = H(L(f)|A(f)), \tag{1}$$

where $A(f)$ denotes the general shape of $L(f)$ (log spectra), which is prior information. $R(f)$, defined as spectral residual of an image, denotes the statistical singularities similar to the input image.

To approximate the shape of $A(f)$, a local average filter is introduced. It is an $n \times n$ matrix defined by:

$$h_n(f) = \frac{1}{n^2}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix},$$

Therefore, we get:

$$A(f) = h_n(f)_* L(f), \tag{2}$$

$$R(f) = L(f) - A(f). \tag{3}$$

Using inverse Fourier transform, we can in spatial domain construct the output image namely the saliency map. The saliency map contains primarily the nontrivial part of the scene. The value at each pixel in a saliency map is then squared to indicate the estimation error. For better effects, a Gaussian filter $g(x, y)$ is used to smooth out the sharpness. It can be interpreted as follows [12]:

$$S(x, y) = g(x, y)_* F^{-1}[\exp(R(f) + P(f))]^2. \tag{4}$$

where $S(x, y)$ denotes saliency map of an image; F^{-1} denotes the inverse Fourier transform; $P(f)$ is the phase spectrum of the image, which is preserved during the process.

3. Cat-eye target recognition method based on saliency extraction

The laser imaging system is utilized to obtain the active image and passive image. Its schematic is shown in Fig. 1. The block diagram of the proposed algorithm is presented in Fig. 2. Then we will describe the procedures of our recognition method in detail.

Three frames of image sequences, consisting of one active image and two passive images (or the opposite case), are required. Suppose I_{p1}, I_{a2}, I_{p3} stand for three frames of image sequences(in time order) acquired by our device. I_{a2} is an active image, I_{p1} and I_{p3} are passive images. We do: $I_{d1} = I_{a2} - I_{p1}, I_{d2} = I_{a2} - I_{p3}$. Thus we get difference image pairs I_{d1} and I_{d2}.

Since the image has two dimensions, an intensity surface of the difference image can be created by mapping the (r, c) coordinates of a pixel to a specific gray level. According to Haralick's cubic facet model, the underlying gray-level intensity surface can be approximated by a bivariate cubic function f in each neighborhood of an image. It is expressed by the bivariate cubic function in canonical form [13].

$$f(r, c) = K_1 + K_2 r + K_3 c + K_4 r^2 + K_5 rc + K_6 c^2 + K_7 r^3 + K_8 r^2 c + K_9 rc^2 + K_{10} c^3 \tag{5}$$

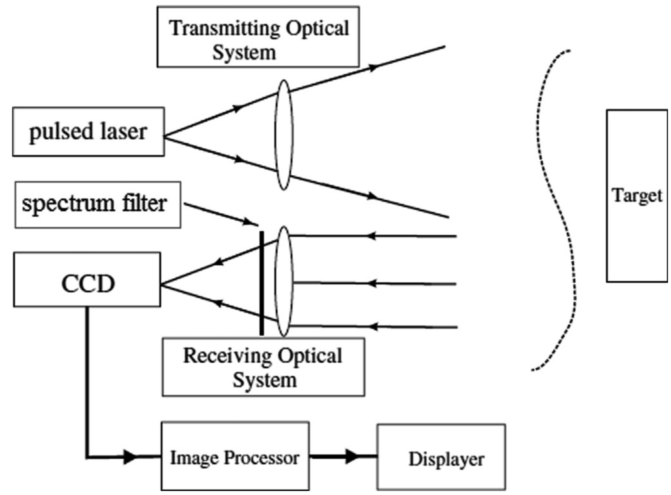

Fig. 1. The cat-eye effect target detection laser imaging system. The system includes a camera lens (Computar Co., H10Z0812DC), a CCD (Watec Co. Ltd., Frame frequency=25 frames/s), a narrow spectrum filter (808 nm),a semiconductor laser with a center wavelength of 808 nm (average power 1 W), and a field-of-view of 6°. The spectrum filter is set between the camera lens and the CCD.

where K_i, $i = 1, \ldots, 10$ are coefficients for the bivariate cubic function expressed in discrete orthogonal polynomials. The second order directional derivative of pixel (x_0, y_0) along vector/can be expressed as [13]

$$\begin{aligned} \frac{\partial^2 f(x, y)}{\partial l^2}\Big|_{(x0,y0)} &= [f_{xx}(x, y)\cos^2\alpha + 2f_{xy}(x, y) \\ &\times \cos\alpha \cos\beta + f_{yy}(x, y)\cos^2\beta]|_{(x0,y0)} \\ &= 2K_4(x_0, y_0)\cos^2\alpha + 2K_5(x_0, y_0) \\ &\times \cos\alpha \cos\beta + 2K_6(x_0, y_0)\cos^2\beta \end{aligned} \tag{6}$$

where α is the angle between l and the x-axis; β is the angle between l and the y-axis. Hence, we get the second order directional derivative filter (SDDF). We select eight directional channels: Channel 1: ($\alpha = 0°$, $\beta = 90°$), Channel 2: ($\alpha = 180°$, $\beta = 90°$), Channel 3: ($\alpha = 90°$, $\beta = 0°$), Channel 4: ($\alpha = 270°$, $\beta = 180°$), Channel 5: ($\alpha = 45°$, $\beta = 45°$), Channel 6: ($\alpha = 225°$, $\beta = 135°$), Channel 7: ($\alpha = 135°$, $\beta = 45°$), Channel 8: ($\alpha = 315°$, $\beta = 225°$) operating on difference images to get SDDF maps. In our algorithm, d_1 is calculated in Channels 1–4 and d_2 is calculated in Channels 5–8. The SDDF maps from different directional channels are denoted as $I_i(x, y)$, $i \in 1, 2, \ldots 8$

Along with the increasing detection distance between imaging system and target, the cat-eye effect target in the image usually has the properties of low signal-to-noise ratio (SNR), small size and unavailable shape information. In addition, target energy scatters in all directions have the characteristic of point spread. Consequently, small target is not sensitive to direction. Therefore morphological method [13,14] can be exploited to deal with the filtered image (SDDF maps) in each direction in order to further eliminate residual clutter and noise.

The SR approach as a saliency map extraction algorithm derived from visual attention theory is used to highlight the targets, as described in Section 2. The saliency maps of SDDF maps are established. In our algorithm, the saliency maps are denoted as $S_i(x, y)$, $i \in 1, 2, \ldots, 8$ respectively.

After that, in order to fuse the 8 saliency maps effectively and quickly, a new saliency fusing approach to fuse the 8 directional saliency maps into a final target-saliency map is put forward. It is divided into two steps:

Step 1: calculate the fused maps using equation: