Infrared Small Target Detection Based on Visual Attention

Hao Chena*, Hong Zhanga, Feiyang Chenga, Ding Yuanb, Mingui Sunb

aImage Processing Center, Beihang University
bLaboratory for Computational Neuroscience, University of Pittsburgh

ABSTRACT

Detecting dim and small target in infrared images and videos is one of the most important techniques in many computer vision applications, such as video surveillance and infrared imaging precise guidance. In this paper, we proposed a real-time target detection approach in infrared imagery. This method combined saliency detection technology and local average filtering. First, we compute the log amplitude spectrum of infrared image. Second, we find the spikes of the amplitude spectrum using cubic facet model and suppress the sharp spikes using local average filtering. At last, the detection result in spatial domain is obtained by reconstructing the 2D signal using the original phase and the filtered amplitude spectrum. Experimental results of infrared images with different types of backgrounds demonstrate the high efficiency and accuracy of the proposed method to detect the dim and small targets.

Keywords: Infrared target detection, visual attention, regular patterns, local average filtering, saliency detection

1. INTRODUCTION

The ability to detection small targets in infrared images or video has a major impact on applications such as video surveillance and infrared imaging precise guidance. However, infrared images have a complex background with low Signal-to-Clutter Ratio (SCR) [1]. All these factors make this problem far from being solved.

There are many automatic target detection techniques. For example, conventional infrared target detection methods based on morphological wavelet [2], directional wavelet [3] and morphology [4]. Recently, a simple and fast algorithm, called the spectrum residual, was proposed based on the Fourier Transform [5]. The paper argues that the spectrum residual corresponds to saliency, and then this model attracted a wide-spread attention in the field of target detection and recognition [6, 7]. The detection is more difficult when targets are small and background clutter is heavy. Dim target detection for infrared image is challenging because of the low contrast between the target and background. It is important for infrared target tracking and recognition system to detect targets accurately and rapidly.

Here we propose a real-time infrared target detection algorithm based on saliency detection and local average filtering. Experimental results of infrared images with different types of backgrounds demonstrate the high efficiency and accuracy of the proposed method to detect the dim and small targets.

2. SALIENCY DETECTION

Many researchers have proposed models of saliency, which invariably then require the detection of salient regions. These regions are described as irregular patches which possess a distinct feature distribution when compared with the rest of the image. In this paper, instead of searching for these irregular patterns, we model regular patterns that do not attract much attention by our visual system.

For convenience, we take a 1D periodic signal \( f(t) \) as an example. Suppose \( f(t) \) can be represented by \( f(t) = \sum_{\omega=-\infty}^{\infty} F(n)e^{i\omega t} \). Then the Fourier transform is given by \( \Gamma(w) = 2\pi \sum_{\omega=-\infty}^{\infty} F(n)\delta(w-n\omega) \).

* Corresponding author Tel: 86-010-82338991, Fax: 86-010-82338991, E-mail address: fpgadsp@163.com
This work was supported by the National Natural Science Foundation of China under Grant No.61272351 and National Institutes of Health grants R01CA165255 and R21CA172864 of the United States.
Figure 1 provides an illustration of this point. Figure 1(a) shows three signals with a different number of repeated patterns (cycles), while Figure 1(b) shows their corresponding log amplitude spectrum. We observe that the larger the number of repeated cycles, the sharper the spikes (labeled by dashed circle) in the spectrum. In [8], the authors illustrate that the spikes in the log amplitude spectrum turn out to correspond to regular patterns in an image.

3. INFRARED TARGET DETECTION

The proposed method can be implemented as follows. First, we compute the log amplitude spectrum of an input infrared image. Second, we find the spikes in the log amplitude spectrum using cubic facet model and suppress the sharp spikes using local average filtering. At last, the target-saliency map in spatial domain is obtained by reconstructing the 2D signal using the original phase and the filtered log amplitude spectrum.

3.1 Log Amplitude Spectrum

Given an infrared image \( f(x, y) \) of size \( M \times N \), it was transformed into the frequency domain:

\[
F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \exp(-j2\pi(ux/M+vy/N))
\]  

(1)

As in the analysis of complex numbers, we find it convenient sometimes to express \( F(u, v) \) in polar coordinates:

\[
F(u, v) = |A(u, v)| \exp(-jP(u, v))
\]  

(2)

Because the zero-frequency term dominates the values of the amplitude spectrum, the dynamic range of other intensities in the displayed image are compressed. To bring out those details, we perform a log transformation. The log amplitude spectrum is obtained:

\[
L(u, v) = \log(A(u, v))
\]  

(3)

3.2 Find the Sharp Spikes

To find the sharp spikes in the log amplitude spectrum, we put forward a new method based on the Haralick’s cubic facet model [9].

In the cubic facet model, each facet centered about a given pixel of the log amplitude spectrum can be approximated by the bivariate cubic function in canonical form. In this paper, let \( R \) be defined as \( R = \{-1,0,1\} \), and \( C \) be defined as \( C = \{-1,0,1\} \). The bivariate cubic function \( f(r, c) \), expressed using discrete orthogonal polynomials, is

\[
f(r, c) = a_1 + a_2 r + a_3 c + a_4 (r^2 - 2/3) + a_6 c + a_7 (c^2 - 2/3) + a_8 r(c^2 - 2/3)
\]  

(4)
Evaluating the second row and column partial derivatives at the neighborhood center \((0,0)\), \(i.e. \ r = 0 \ and \ c = 0\) yields the second directional derivatives
\[
\frac{\partial^2 f(r,c)}{\partial r \partial c} = 2a_s, \quad \frac{\partial^2 f(r,c)}{\partial r^2} = a_s, \quad \frac{\partial^2 f(r,c)}{\partial c^2} = 2a_s
\] (5)

Each \(a_i\) can be computed independently by convolving the log amplitude spectrum with the corresponding weight kernel. The three weight kernels are obtained as follows:
\[
w_1 = \begin{bmatrix} -1 & -2 & 1 \\ -1 & -2 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 1 & 1 \end{bmatrix}, \quad w_3 = \begin{bmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{bmatrix}
\]

According to extremum theory, if
\[
D_1 = 2a_i < 0; \quad D_2 = 4a_s a_s - a_i^2 < 0
\] (6)

Then the corresponding \(f(r,c)\) is a spike. According to many experiments, there are many spikes satisfying the conditions, but we only smooth the sharp spikes.

3.3 Target Detection

In order to detection infrared targets, we need to suppress the sharp spikes which lead to regular patterns. Using cubic facet model, we find the regions \(S_i (i = 1, 2, 3 \ldots)\) which contain the sharp spikes of \(L(u,v)\) and the rest are called B. The traditional log amplitude spectrum can be formulated as: \(L(u,v) = B + \sum_{i=1}^{n} S_i\). Then, we can suppress the spikes through convoluting the \(S_i\):
\[
S_i' = S_i \ast h_s(f)
\] (7)

where \(h_s(f)\) is an \(n \times n\) matrix defined by:
\[
h_s(f) = \frac{1}{n^2} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}
\]

In our experiments, \(n = 3\). The filtered log amplitude spectrum is:
\[
L(u,v) = B + \sum_{i=1}^{n} S_i'
\] (8)

Based on the above discussion, we define the target-saliency map hereafter as:
\[
g(x, y) = F^{-1}\{L(u,v)e^{i\rho(u,v)}\}
\] (9)

Figure 2 shows the diagram illustrating our novel infrared small target detection method.
4. EXPERIMENTAL RESULT AND EVALUATION

In order to validate the performance and robustness of our proposed algorithm, experiments are performed on typical real infrared images representing several kinds of complex backgrounds. Meanwhile, we compare our algorithm with four state-of-the-art approaches, including morphological detection method (Top-hat), adaptive Butterworth high-pass filter method (BHPT), facet-based method (Facet), and phase spectrum of the Fourier Transform method (PFT). For fair comparison, we adopt one common evaluation indicator, i.e., SCR Gain defined as follows:

$$\text{SCR} = \left| \mu_t - \mu_b \right| / \sigma_b$$  \hspace{1cm} (10)

where $\mu_t$ is the average pixel value of the target, $\mu_b$ and $\sigma_b$ are the average pixel value and the standard deviation of the pixel values in neighboring area around the target, respectively. The experimental results, including SCR and run time, are listed in Table I.

<table>
<thead>
<tr>
<th>Method</th>
<th>Evaluation Indicators</th>
<th>Image (a)</th>
<th>Image (b)</th>
<th>Image (c)</th>
<th>Image (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top-hat</td>
<td>SCR Gain</td>
<td>1.4145</td>
<td>7.7131</td>
<td>11.8914</td>
<td>17.4571</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>0.0073</td>
<td>0.0424</td>
<td>0.0142</td>
<td>0.0151</td>
</tr>
<tr>
<td>BHPT</td>
<td>SCR Gain</td>
<td>1.2268</td>
<td>4.6845</td>
<td>9.0171</td>
<td>15.0385</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>0.1295</td>
<td>1.4586</td>
<td>0.3198</td>
<td>0.3089</td>
</tr>
<tr>
<td>Facet</td>
<td>SCR Gain</td>
<td>1.9799</td>
<td>8.6069</td>
<td>15.5852</td>
<td>28.4302</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>0.0515</td>
<td>0.1217</td>
<td>0.0702</td>
<td>0.0608</td>
</tr>
<tr>
<td>PFT</td>
<td>SCR Gain</td>
<td>3.2283</td>
<td>8.2377</td>
<td>42.4939</td>
<td>55.2414</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>0.0103</td>
<td>0.0666</td>
<td>0.0175</td>
<td>0.0170</td>
</tr>
<tr>
<td>Our method</td>
<td>SCR Gain</td>
<td>4.7978</td>
<td>31.2548</td>
<td>46.8349</td>
<td>73.6772</td>
</tr>
<tr>
<td></td>
<td>Time (s)</td>
<td>0.0934</td>
<td>0.9664</td>
<td>0.2181</td>
<td>0.1903</td>
</tr>
</tbody>
</table>
Table I shows that our method has good performance in SCR Gain. Considering this fact, our method actually outperforms other methods in most cases. This fact can also prove that the proposed method is more robust in such harsh circumstances. In spite that our method is more time consuming than Top-hat, it has an acceptable order of magnitude just as traditional method BHPT and can be improved by parallel computing.

5. CONCLUSION

In this paper, we proposed a real-time target detection approach in infrared imagery. This method combined saliency detection technology and local average filtering. First, we compute the log amplitude spectrum of infrared image. Second, we find the spikes of the amplitude spectrum using cubic facet model and suppress the sharp spikes using local average filtering. At last, the detection result in spatial domain is obtained by reconstructing the 2D signal using the original phase and the filtered amplitude spectrum. Experimental results of infrared images with different types of backgrounds demonstrate the high efficiency and accuracy of the proposed method to detect the dim and small targets.

REFERENCES