Infrared Physics & Technology 62 (2014) 100–109

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Infrared dim and small target detecting and tracking method inspired by Human Visual System

Xiabin Dong, Xinsheng Huang *, Yongbin Zheng, Lurong Shen, Shengjian Bai

College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha, Hunan 410073, PR China

HIGHLIGHTS

- A novel method which is inspired by Human Visual System is proposed in this paper.
- This method combines three mechanisms of HVS together.
- DOG filter is used to simulate the contrast mechanism.
- A Gaussian window is added at a point to simulate the visual attention.
- The PID algorithm is first introduced to simulate the eye movement of human being.

ARTICLE INFO

Article history:
Received 10 July 2013
Available online 20 November 2013

Keywords:
Infrared small target tracking
Human Visual System
DOG filter
Gaussian window
PID algorithm

ABSTRACT

Detecting and tracking 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. Recently, more and more algorithms based on Human Visual System (HVS) have been proposed to detect and track the infrared dim and small target. In general, HVS concerns at least three mechanisms including contrast mechanism, visual attention and eye movement. However, most of the existing algorithms simulate only a single one of the HVS mechanisms, resulting in many drawbacks of these algorithms. A novel method which combines the three mechanisms of HVS is proposed in this paper. First, a group of Difference of Gaussians (DOG) filters which simulate the contrast mechanism are used to filter the input image. Second, a visual attention, which is simulated by a Gaussian window, is added at a point near the target in order to further enhance the dim small target. This point is named as the attention point. Eventually, the Proportional-Integral-Derivative (PID) algorithm is first introduced to predict the attention point of the next frame of an image which simulates the eye movement of human being. Experimental results of infrared images with different types of backgrounds demonstrate the high efficiency and accuracy of the proposed method to detect and track the dim and small targets.

1. Introduction

The ability to detect and track dim and small targets in infrared (IR) images or videos has a major impact on applications such as video surveillance and IR imaging precise guidance in military use. In these applications targets are always far away from the imaging equipment, as a result, targets in the sensed IR images or videos are usually very dim, small and shapeless. Besides, IR images have a low Signal-to-Noise ratio (SNR) and the background of the target is complicated and chaotic. All these factors make this problem far from being solved. Recently, there has been a surge of interest in dim and small target detecting and tracking [1–19]. The leading approaches for this problem can be divided into two classes: one is to detect targets using only a single frame of images. The target is tracked as long as it has been detected successfully in every single frame of images. This kind of algorithms is simple and only suitable for targets with homogeneous background; the other uses videos to detect and track targets. Because of utilizing more information, this kind of algorithms can detect and track targets under complicated backgrounds.

In demand of a robust infrared small target detecting and tracking algorithm under much more complicated background, even under a background with lots of traffics, we have proposed a novel algorithm which can deal with such complicated scenarios. The experiments are carried out in IR images with different types of complicated backgrounds and the results demonstrate the high efficiency and the accuracy of the proposed method.

2. Related work

There are many target detection methods based on a single frame of images proposed in the open literature. Kim et al. [3]

* Corresponding author. Tel.: +86 15111003874.
E-mail address: dxb_nudt@163.com (X. Huang).

1350-4495/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.infrared.2013.11.007

proposed a small target detection method based on the contrast mechanism of HVS which used Laplacian of Gaussian (LOG) filters to deal with the input image. The LOG filter is capable of increasing the small target intensity as well as suppressing background clutters and noises, which is a kind of best simulation of the contrast mechanism. But this method is not suitable for infrared images with complicated backgrounds. Shao et al. [4] improved the method of Kim by adding a morphological processing which ensures that a target can be easily distinguished from the background. Then local threshold segmentation is used to obtain the true target by considering the characteristic that a target is the brightest among its local domain. In 2012, Kim and Lee [5] proposed a novel method of Tune-Max of SCR (TM-SCR) in scale-space to solve both the scale problem and the clutter rejection problem motivated by the properties of HVS. The key idea is to simultaneously enhance the target information and minimize the background clutter in scale-space. This method works well if there is a horizontal background. Itti et al. [6] and Elazary and Itti [7] proposed a saliency-based visual attention system which extracted three kinds of visual feature maps from input images. They are colors feature maps, intensity feature maps and orientations feature maps. Their methods combine these three kinds of feature maps together to locate the target of interest. But this model of visual attention performs badly when detecting an IR dim and small target, because this kind of target is colorless and shapeless. Wang et al. [8] presented an automatic detection algorithm based on visual attention for IR dim target. This method firstly uses the DOG filters to compute the saliency map. Then salient regions where potential targets exist in are extracted by searching through the saliency map with a control mechanism of winner-take-all (WTA) competition and inhibition-of-return (IOR). At last, according to the characteristics of dim IR targets, true targets are detected, and spurious objects are rejected by examining these salient regions.

The methods mentioned above just use information of the current frame to detect and track targets. In fact, there are a lot of information of the frame sequences we can use when locating the targets at the current frame, such as the intensity, the position and the motion of targets. It seems like that the tracking method based on video could deal with scenarios under more complicated backgrounds. In order to rapidly find the location in the image where the feature histogram is closest to the reference histogram of the target, Mean Shift [10,11] (MS) tracker has been proposed to find such a location by employing MS iterations. The MS tracker [12–14] became a widely used tracker because of its excellent properties of high speed and robustness. But MS tracker can easily deviate from the center of the small target, because this kind of methods just use the feature histogram of the target in the last one frame of images and it can bring accumulated errors. Cross-Correlation (CC) approach [15,16] is a well-known image matching method and can also be used in target tracking. However, CC-based target tracking algorithms are sensitive to changes in rotation and the scale of the target. Laura and Miller [17] introduced Distribution Fields (DFs) into target tracking. This method uses DFs to represent the target instead of the feature histogram, but it can bring accumulated errors the same as MS tracker. Shaik and Iftekharuddin [18] proposed a target detecting and tracking method using Bayesian techniques. In this technique, given the location of the target in the previous frame, the probabilities of the target moving to all possible locations in the next frame is obtained and then the location in the next frame which produces the highest probability is selected. Huang and Mao [19] employed the Kalman filter to predict the location of the target in the next frame and track the target.

Because of the high efficiency and accuracy of the HVS, more and more algorithms based on HVS have been proposed by many researchers [3–9]. In general, HVS concerns at least three mechanisms including contrast mechanism, visual attention and eye movement. However, the algorithms mentioned above only use one of the three mechanisms to simulate the HVS and this can cause some drawbacks. This work presents a novel algorithm which combines the three mechanisms of HVS together to improve the performance of detecting and tracking IR small targets. First, Difference of Gaussians (DOG) filter, which is an approximation to LOG filter is used to filter the input image. Second, Gaussian window, that simulates the visual attention mechanism, is added at a point near the target to enhance the dim small target further, and we call this point as attention point. Last, the Proportional-Integral-Derivative (PID) algorithm is first introduced to predict the attention point of the next frame of an image which simulates the eye movement of human being. The proposed method requires location initialization for each new interest target when it appears in the first frame.

This article is organized as follows: Section 3 gives a detailed description of infrared dim and small targets detecting and tracking method inspired by HVS. The experimental results and the comparisons between the proposed method and the previous methods are shown in Section 4. Section 5 provides the conclusions of this work.

3. The proposed method

In this section a novel algorithm of IR small and dim targets detecting and tracking is proposed. The process of this algorithm includes four steps. Fig. 1 shows the framework of this method.

- Step 1: Filtering the current frame of an image with multi-scale DOG filters and a group of feature maps with different scales are obtained. Then the saliency map is obtained by combining these feature maps together.
- Step 2: Adding a Gaussian window at the visual attention point in the saliency map.
- Step 3: Normalizing the whole image and deciding the location of the target in the current frame of the image.
- Step 4: If it is the last frame, then the program stops. Otherwise it needs to estimate the location of the visual attention point of the next frame with the means of PID, and goes back to step 1.

The following subsections give a detailed description of each part of the algorithm respectively.

3.1. DOG filters inspired by contrast mechanism

When targets with different intensities are set in backgrounds which are also of different intensities, HVS perceives a target's brightness according to the contrast between the target and the background, and considers that the brightness of two targets approximate if they have a similar contrast. So it is the contrast between the target and the background rather than the intensity of the target that plays an important role in HVS. The LOG filter can not only suppress the background noise and clutter but also

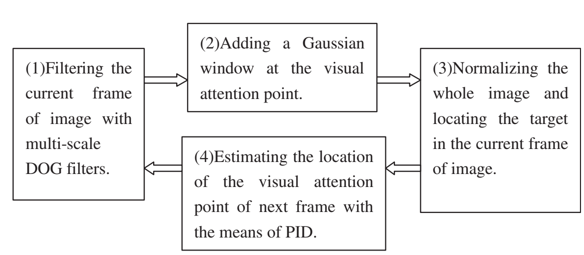

Fig. 1. Framework of IR small and dim targets detecting and tracking method inspired by HVS.