Infrared Physics & Technology 69 (2015) 7–12

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Learning to detect small target: A local kernel method

Kai Xie, Tao Zhou, Yu Qiao, Chenjie Ge, Jie Yang *

Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, 200240, China
Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China

HIGHLIGHTS

- This paper proposes a novel method to detect infrared small target in heavy clutter.
- The method could preserve heterogeneous area while removing target region.
- Different regions could be differentiated using the weighted quadratic model.
- The kernel trick makes our local learning method more general.

ARTICLE INFO

Article history:
Received 31 July 2014
Available online 7 January 2015

Keywords:
Small target detection
Local learning framework
Kernel trick
Heavy clutter

ABSTRACT

Small target detection is a critical problem in the Infrared Search And Track (IRST) system. Although it has been studied for years, there are some challenges remained, e.g. cloud edges and horizontal lines are likely to cause false alarms. This paper proposes a novel local learning framework to detect infrared small target in heavy clutter. First, we propose a quadratic cost function to learn the parameters in the weighted local linear model. Second, we introduce the kernel trick to extend the linear model to the non-linear model. Finally, small targets are detected in the residual image which subtracts the estimation image from original input. Our method could preserve heterogeneous area while removing target region. Experimental results show our method achieves satisfied performance in heavy clutter.

1. Introduction

Small target detection is a critical problem in the Infrared Search And Track (IRST) system. Although it has been studied for years [1–7], there are some challenges remained, e.g. cloud edges and horizontal lines are likely to cause false alarms. The reasons are as follows: first, features such as texture and color are unavailable for small targets when they are far away from the infrared sensor. Second, heterogeneous areas such as cloud edges and sky-sea lines may be falsely detected as small targets.

Background estimation based small target detection method is widely studied in recent years [8–11]. These methods detect small targets from the residual image, i.e. subtracting the estimated background image from the original input. The detection performance relies highly on the quality of the estimated background. A good background estimation method should preserve the heterogeneous area as much as possible and exclude the target region. In other word, it requires the method to effectively distinguish the target region and the heterogeneous area. The 2-D least mean square (TDLMS) method [8] minimizes the difference between an input image and a background image that is estimated by the weighted average of neighboring pixels. The TopHat method [9] estimates background by a morphological opening operator with structure element. These methods are less effective because they could not distinguish the target region and the heterogeneous region, and the residual image contains much clutter.

Existing background suppression methods are mainly based on the filtering methods [12,13]. The LS-SVM [13] method uses filter templates which can suppress most part of the correlative background but may be easily interfered because of the strong fluctuation of background clutters. Yang et al. [12] provide an adaptive Butterworth high pass filter by suppressing low frequency components to achieve the purpose for enhancing targets. This method makes use of the diversity between small targets and backgrounds in frequency domain. Insufficiently, it is inoperative to noises or clutter fluctuation which also possesses high frequency in the image.

These methods [8,9,12,13] may achieve good results on simple background, but not on heavy clutter background. The main reason is they cannot distinguish small targets from heavy clutter

* Corresponding author at: Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, 200240, China.
E-mail address: jieyang@sjtu.edu.cn (J. Yang).

http://dx.doi.org/10.1016/j.infrared.2014.11.014
1350-4495/

effectively. In this paper, we focus on removing target while preserving the high-frequency component in the background. We propose a novel local learning framework to detect infrared small target in heavy clutter. First, we propose a quadratic cost function to learn the parameters in the weighted local linear model. Second, we introduce the kernel trick to extend the linear model to the nonlinear model. Finally, small targets are detected in the residual image which subtracts the estimation image from original input. The proposed framework achieves better performance, when compared with TDLMS [8], TopHat [9], LS-SVM [13], especially in the heavy clutter.

Fig. 1 shows the proposed small target detection system. It consists of kernel based local learning, generating residual image, thresholding and 8-NN based clustering.

2. The proposed method

Generally, the IR image model can be formulated as:

$$f = f_T + f_B + n, \tag{1}$$

where f, f_T, f_B and n are IR image, target image, background image and random noise, respectively. n is assumed to follow Gaussian distribution with mean 0 and variance σ^2 [6].

Background estimation based small target detection method aims to estimate f_B. We propose a novel local learning framework to solve this problem. Our kernel based local learning framework trains a local regression model for estimating the background pixel only based on its neighboring pixels which are considered to be most related. We note that it has a big difference from the conventional method, for we only consider related pixels in the neighborhood. We apply our method on IR image using sliding window manner, then we can obtain a background estimation image of the same size of original input.

We denote the estimation pixel's neighborhood (excluding itself) by $\Omega = \{1, 2, \ldots, n\}$. Where n is the total number of pixels in the neighborhood. We treat each pixel $i \in \Omega$ as a data point denoted by x_i. x_i can be set as features extracted for pixel i. In this paper, we denote x_i as a 3D vector:

$$x_i = \begin{pmatrix} coord_r \\ coord_c \\ 1 \end{pmatrix} \tag{2}$$

where $coord_r$ and $coord_c$ are normalized coordinates of pixel i. The local learning framework can be then formulated as follows: Given a data point set $\chi = \{x_i\}_{i \in \Omega}$, data point of center pixel x_0, intensity values of each pixel $\{y_i\}_{i \in \Omega}$ and intensity value of center pixel y_0, our goal is to compute the intensity value of estimation pixel y^* through learning method.

We firstly introduce a weighted quadratic cost function. Through minimizing the cost, we can achieve a linear solution. Then the linear solution is extended to being nonlinear with the kernel trick.

2.1. The local linear model

For any pixel $i \in \Omega$, we assume that its intensity value y_i can be predicted by a linear model:

$$y_i = a^T x_i, \quad i \in \Omega \tag{3}$$

A weighted quadratic cost function is proposed to estimate a:

$$\min_a \sum_{i \in \Omega} (a^T x_i - y_i)^2 * w_i + \lambda ||a||_2^2 \tag{4}$$

where λ is a tradeoff-factor between the two terms. In this paper, λ should be small, for the second term only enhances numerical stability of our framework.

We select the pixels in a $N \times N$ local patch centered at estimation pixel as the neighbors. In practice, the neighborhood is divided into inner and outer area, shown in Fig. 2. M is the size of inner window. The inner window should be big enough to cover the reference object. The w_i can be calculated by intensity value of center pixel and pixel $i : w_i = e^{-c*||y_0 - y_i||_2^2}$, where c is a constant. We notice that the more similar intensity values y_0, y_i are, the larger w_i is. w_i of the pixel $i \in \Omega$ in inner window could directly be set to 0.

By introducing some new denotations: matrix X through stacking $\{x_i\}_{i \in \Omega} : X = [x_1, x_2, \ldots, x_n]$, matrix W, where

$$W = \begin{bmatrix} w_1 & 0 & 0 & \cdots & 0 \\ 0 & w_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & w_n \end{bmatrix} \tag{5}$$

and matrix Y through stacking $\{y_i\}_{i \in \Omega}$, where

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \tag{6}$$

we can rewrite Eq. (4) in a more concise format:

$$\min_a (X^T a - Y)^T W (X^T a - Y) + \lambda ||a||_2^2 \tag{7}$$

According to Eq. (7), we can easily achieve the closed-form solution of the linear model.

$$a = (XWX^T + \lambda I_{3 \times 3})^{-1} XWY \tag{8}$$

2.2. Extending to nonlinear model

We notice that W can be divided into two equal matrix W', where

$$W' = \begin{bmatrix} \sqrt{w_1} & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{w_n} \end{bmatrix} \tag{9}$$

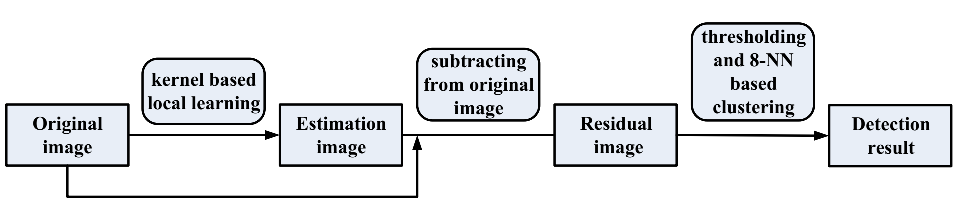

Fig. 1. Pipeline of our detection system.