Infrared Physics & Technology 76 (2016) 148–156

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Infrared small moving target detection using sparse representation-based image decomposition

Hanlin Qin [a,*], Jiaojiao Han [a], Xiang Yan [a], Qingjie Zeng [a], Huixin Zhou [a], Jia Li [a,b], Zhimin Chen [c]

[a] School of Physics and Optoelectronic Engineering, Xidian University, Xi'an 710071, China
[b] Science Institute, Air Force Engineering University, Xi'an 710051, China
[c] Joint Laboratory of Flight Vehicle Ocean-based Measurement and Control, Wuxi 214400, China

HIGHLIGHTS

- Target is separated by the morphology differences between target and background.
- Random projection is used to reduce the redundant information in time domain.
- The RX filter is adopted to extract the target trajectory from clutter background.
- The proposed method works efficiently and robustly with low false alarm.

ARTICLE INFO

Article history:
Received 25 November 2015
Available online 24 February 2016

Keywords:
Infrared image
Moving target detection
Image decomposition
Sparse representation
Random projection

ABSTRACT

Infrared small moving target detection is one of the crucial techniques in infrared search and tracking systems. This paper presents a novel small moving target detection method for infrared image sequence with complicated background. The key points are given as follows: (1) since target detection mainly depends on the incoherence between target and background, the proposed method separate the target from the background according to the morphological feature diversity between target and background; (2) considering the continuity of target motion in time domain, the target trajectory is extracted by the RX filter in random projection. The experiments on various clutter background sequences have validated the detection capability of the proposed method. The experimental results show that the proposed method can robustly provide a higher detection probability and a lower false alarm rate than baseline methods.

1. Introduction

Infrared search and tracking system (IRST) is widely used in modern defense, where infrared small moving target detection technology plays a crucial role [1]. Generally, small moving target appears as a point without any shape and structure information, since it is far away from the imaging system. Moreover, infrared small target is often submerged in complicated background with low signal-to-clutter ratio (SCR). Thus, the accurate detection of infrared small moving target is considered as a difficult and challenging work. To accurately detect small target, various methods have been proposed over the past few years and they are mainly divided into the single-frame detection method and the sequential detection method [2,3].

Single-frame detection methods could be further classified into background estimation based methods and target extraction based methods. For the former methods, they focus on designing a filter to well estimate background, and the target is detected by subtracting the estimated background from the original image, such as max-mean/max-median filter [4], two-dimensional least mean square filter [5], top-hat transform [6], hit-or-miss transform [7], and toggle contrast operator [8]. In addition, Gu et al. [9] introduced a kernel-based nonparametric regression method to estimate the background. Bae et al. [10] worked to predict clutter background by using an edge directional 2D least-mean squares filter. These methods use the fixed-scaled masks to predict the background, which makes they cannot efficiently detect the targets with changing sizes in real cases. What's more, these methods are also sensitive to the various heavy clutter background. For the latter methods, they directly detect the targets via extracting targets feature. Sun and Kwak [11] proposed the center-surround difference operation with adaptive threshold to detect targets.

* Corresponding author.
E-mail address: hlqin@mail.xidian.edu.cn (H. Qin).

http://dx.doi.org/10.1016/j.infrared.2016.02.003
1350-4495/© 2016 The Authors. Published by Elsevier B.V.

Inspired by the contrast mechanism of human vision system, a local contrast method for small infrared target detection has been presented in [12]. Zhao et al. [13] proposed an infrared small target detection method via using the over-complete target dictionary produced by two-dimensional Gaussian model to represent sparsely target. Gao et al. [14] proposed a sparse ring represent method to detect infrared small targets, which based on local self-similarity descriptor according to the local characteristics of small target. Furthermore, a new method based on second-order directional derivative filter and Boolean map visual theory is applied to achieve small target detection [15]. These methods can usually work well, but may become less effective for the complicated background because of high false alarm rates.

In sequential detection methods, in general, they deal with several frames to estimate backgrounds or extract targets according to the continuity and regularity of moving targets in temporal domain. Many typical methods have been widely used in application, such as triple temporal filter [16], Mexican hat continuous wavelet transform [17], infinite impulse response [18] and three-dimensional matched filtering [19]. With the development of geometric theory, a lot of new methods have been proposed. Such as, Liu et al. [20] proposed a small moving target detection method using the temporal profile based on the connecting line of the stagnation points; Li et al. [21] employed tensor locality preserving projection to detect the moving point target in the infrared image sequence. These methods can obtain relative good effect under static background. However, they fail to detect the targets on a non-stationary background and also still have higher false alarm rates. To further improve the detection performance under complicated background, some new detection methods based on spatial–temporal information have been proposed. Such as, a moving point target detection method based on three-dimensional spatiotemporal anisotropic diffusion model is presented in [22]; Chen et al. [23] proposed a novel spatial–temporal detection method based on bi-dimensional empirical mode decomposition and time-domain differential filtering; Bae [24] introduced a spatial and temporal bilateral filter for target detection, which uses spatial bilateral filter to extract spatial target information and applies temporal bilateral filter to extract temporal target information. These methods are more effective compared with the traditional detection methods. Nevertheless, the detection results of these methods are hard to satisfy, when they handle the images with heavy clutter background or non-moving target.

To cope with the images with heavy clutter background or non-moving target, a new spatiotemporal detection method using morphological component analysis (MCA) based on sparse representation is proposed in this paper. According to the morphological feature diversity of target and background [25], small target with the isotropic structure is sparsely represented by spectral graph wavelet transform (SGWT) [26], meanwhile complicated piecewise smooth background is efficiently represented by nonsubsampled shearlet transform (NSST) [27]. In the proposed method, the small target is separated from the clutter background for several frames, and a group of target images are obtained. Subsequently, random projection [28] is employed to reduce the redundancy of information in target images, and a small number of target images are generated. Finally, the target trajectory can be extracted by the RX filter [29] from these target images according to the continuity of moving target in temporal domain. Several experimental results show that the proposed method can not only efficiently enhance the small target and suppress the clutter background, but also improve the detection performance for small moving target compared with other existing methods.

The rest of this paper is organized as follows. In Section 2, a sparse decomposition method for infrared image containing small target is described. Section 3 presents the proposed small moving target detection method. Section 4 performs experiments for several infrared image sequences to show the effectiveness of the proposed method. Finally, we conclude this paper in Section 5.

2. Sparse decomposition of infrared image

Generally, the infrared image sequence containing small target and clutter background can be modeled as the linear superposition of target and background:

$$\boldsymbol{x}(i,j,k) = \boldsymbol{x}_T(i,j,k) + \boldsymbol{x}_B(i,j,k), \tag{1}$$

where $\boldsymbol{x}(i,j,k)$ is the intensity of image $\boldsymbol{x}$ at pixel (i,j) in the kth $(k \in \{1,2,\ldots,K\})$ frame. $\boldsymbol{x}$, $\boldsymbol{x}_T$ and $\boldsymbol{x}_B$ denote infrared image, target image and background image, respectively. In this paper, the small moving target detection problem is converted into a sparse decomposition problem of infrared image. Namely, the pixel $\boldsymbol{x}(i,j,k)$ is decomposed into target pixel $\boldsymbol{x}_T(i,j,k)$ and background pixel $\boldsymbol{x}_B(i,j,k)$ based on the diversity of morphological features between target and background.

2.1. Morphological component analysis

To achieve image sparse decomposition, the MCA theory is introduced in this paper. As shown in Fig. 1, the MCA is applied to the original infrared image $\boldsymbol{x}$ to separate the components $\boldsymbol{x}_T$ and $\boldsymbol{x}_B$. In theory, the MCA assumes that each morphological component of an image can be sparse represented by an associated dictionary. Thus, in this paper, the target component $\boldsymbol{x}_T$ and background component $\boldsymbol{x}_B$ can be represented as follows:

$$\boldsymbol{x}_T = \boldsymbol{\Psi}_T \boldsymbol{\theta}_T, \tag{2}$$

$$\boldsymbol{x}_B = \boldsymbol{\Psi}_B \boldsymbol{\theta}_B, \tag{3}$$

where $\boldsymbol{\Psi}_T$ and $\boldsymbol{\Psi}_B$ represent the dictionaries of target component and background component, respectively. $\boldsymbol{\theta}_T$ and $\boldsymbol{\theta}_B$ are the sparse coefficient vectors of target and background in the corresponding dictionaries, respectively. Thus, the dictionary $\boldsymbol{\Psi}$ which can sparsely represent image $\boldsymbol{x}$ can be built by amalgamating target dictionary and background dictionary $[\boldsymbol{\Psi}_T \quad \boldsymbol{\Psi}_B]$.

$$\boldsymbol{x} = \boldsymbol{\Psi}_T \boldsymbol{\theta}_T + \boldsymbol{\Psi}_B \boldsymbol{\theta}_B = [\boldsymbol{\Psi}_T \quad \boldsymbol{\Psi}_B] \begin{bmatrix} \boldsymbol{\theta}_T \\ \boldsymbol{\theta}_B \end{bmatrix} = \boldsymbol{\Psi}\boldsymbol{\theta}, \tag{4}$$

where $\boldsymbol{\theta}$ is a sparse vector consisting of the vectors $\boldsymbol{\theta}_T$ and $\boldsymbol{\theta}_B$. In general, the morphological features of the target and the background are different. Thus, they can be distinguished by different dictionaries. But these dictionaries must be incoherent. Specifically, if $\boldsymbol{x}$ is a target, it cannot be sparsely represented by the background dictionary $\boldsymbol{\Psi}_B$. In this case, $\boldsymbol{\theta}_B$ is a zero vector and $\boldsymbol{\theta}_T$ is a sparse vector; and vice versa [25,30]. Generally, the nonzero coefficients in the sparse vector $\boldsymbol{\theta}$ contains key information of infrared image $\boldsymbol{x}$.

However, the dictionary $\boldsymbol{\Psi} = [\boldsymbol{\Psi}_T \quad \boldsymbol{\Psi}_B]$ provides an overcomplete representation of $\boldsymbol{x}$. In this case, the underdetermined system $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{\theta}$ can be solved and morphological components $\boldsymbol{x}_T$ and $\boldsymbol{x}_B$ can be recovered by solving the following constrained optimization problem:

$$\min_{\boldsymbol{\theta}_T,\boldsymbol{\theta}_B} \|\boldsymbol{\theta}_T\|_1 + \|\boldsymbol{\theta}_B\|_1 + \rho TV(\boldsymbol{\Psi}_B\boldsymbol{\theta}_B) \quad \text{s.t.} \quad \boldsymbol{x} = \boldsymbol{\Psi}_T\boldsymbol{\theta}_T + \boldsymbol{\Psi}_B\boldsymbol{\theta}_B, \tag{5}$$

where $\|\cdot\|_1$, i.e., l_1-norm, denotes the sum of absolute value of nonzero elements. In essence, the total variation of an image is the l_1-norm of the gradient. Thus, a total variation (TV) penalty is added to the background component force the component $\boldsymbol{x}_B$ to have a sparse gradient, and to be closer to a piecewise smooth component. Here, ρ is a regularization factor controlling the smoothness degree of the TV correction.