Image and Vision Computing 63 (2017) 1–9

Contents lists available at ScienceDirect

Image and Vision Computing

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

Infrared dim target detection based on total variation regularization and principal component pursuit☆

Xiaoyang Wang, Zhenming Peng *, Dehui Kong, Ping Zhang, Yanmin He

School of Opto-Electronic Information, University of Electronic Science and Technology of China, Chengdu 610054, China

ARTICLE INFO

Article history:
Received 6 April 2016
Accepted 20 April 2017
Available online 29 April 2017

Keywords:
Infrared images
Dim target detection
Total variation regularization
Principal component pursuit

ABSTRACT

Robust detection of infrared dim and small target contributes significantly to the infrared systems in many applications. Due to the diversity of background scene and unique characteristic of target, the detection of infrared targets remains a challenging problem. In this paper, a novel approach based on total variation regularization and principal component pursuit (TV-PCP) is presented to deal with this problem. The principal component pursuit model only considers the low-rank feature of background images, which will result in poor detection ability in non-uniform and non-smooth scenes. We take into account the total variation regularization term to thoroughly describe background feature, which can achieve good detection result as well as good background estimation result. Firstly, the input infrared image is transformed to a patch image model. Secondly, the TV-PCP model is presented on the patch image. An effective optimization algorithm is proposed to solve this model. Experiments on six real datasets show that the proposed method has superior detection ability under various backgrounds, especially with good background suppression performance and low false alarm rate.

1. Introduction

Infrared target detection plays an important part in various applications, especially in military area such as accurate guidance, anti-missile technique and early warning systems. These application areas are always facing complex and changeable scenes. Considering the long imaging distance and various backgrounds, infrared targets in aforesaid applications always share some common characteristics like small size and low signal-to-clutter ratio (SCR). These targets are lack of shape information, texture information and color information. Meanwhile, background clutter will have severe impact on the detection result of infrared dim target. Thus, infrared dim target detection technique remains a challenging problem.

Recently, intensive researches have focused on developing efficient methods to solve the problem of infrared dim and small targets detection from various scenes. Generally speaking, the existing methods can be divided into two types: detection before track (DBT) method and track before detection (TBD) method. As the name indicates, DBT method concentrates more on detecting target from a single frame, using approaches like image filtering, pattern recognition and so on. These methods can be considered as spatial feature based methods. Classic image filtering based methods contains two dimensional least mean square(TDLMS) method [1], top-hat filtering method [2], wavelet transformation method [3] and so on. Pattern recognition based method uses different features in spatial domain to divide pixels into background class and target class [4]. Another type of method to detect dim and small target is TBD method, which considers the inter-frame information as a crucial feature in target detection. Typical TBD methods include 3D matched filtering [5–6], dynamic programming method [7–8], multistage hypothesis testing method [9–10] and maximum likelihood method [11]. These methods make up the deficiency of DBT methods by processing continuous frames.

The implicit idea in aforesaid methods is to use different features of background and target in order to describe and separate these two parts. There exist some empirical assumptions on infrared backgrounds and targets. The method in [12] treats infrared targets as circular shaped signal, which is described by the 2D Gaussian function. For statistic background, background modeling method can be used to realize background subtraction and target enhancement [13,14]. Given some prior information of background and target, the detection procedure can also be treated as a classification problem, dividing pixels into background class and target class [15]. However, these methods heavily rely on the prior information and proper assumption of target feature and background characteristic, which is not very convenient to obtain in practical real-time system.

Some novel features are also introduced into the detection approach of infrared dim and small target. Despite traditional spatial and

☆ This paper has been recommended for acceptance by Stefanos Zafeiriou.

* Corresponding author.
E-mail addresses: xywang1211@outlook.com (X. Wang), zmpeng@uestc.edu.cn (Z. Peng), kongdehui2013@gmail.com (D. Kong), pingzh@uestc.edu.cn (P. Zhang), heyanmin@uestc.edu.cn (Y. He).

http://dx.doi.org/10.1016/j.imavis.2017.04.002
0262-8856/

temporal features, new researches focus on extracting features of target and background in diverse domain, combining with new theories in image and signal processing. One example is the human visual system based methods, which makes use of the contrast of target region and local background [16–17], or the difference of Gabor filter [18] and so on. These methods perform well in some certain scenes. Line-based feature is also introduced to suppress the background clutter [19]. Along with the development of matrix decomposition and completion theory, the low-rank feature of background model and sparse feature of target model have been attracting more and more attention [20,21]. One research shows that some of the infrared backgrounds have the non-local self-correlation feature [20]. This theory considers uniform and blurred background as low-rank matrix, using an infrared patch image (IPI) model, while small target can be considered as sparse matrix. Target detection problem is transferred into low-rank and sparse matrix separation problem under this assumption. This problem can be solved by principal component pursuit (PCP). Similar approach can also been done in compressive domain using rank estimation [22]. However, this model can only reveal one side of the background characteristics. Utilizing PCP method, we cannot separate background and targets completely, especially in some non-smooth and non-uniform scenes. Extracting more inherent features of infrared images which can fit various working scene is crucial in reaching a better performance of dim target detection. To summarize, most of the research focused on the single assumption of target or background, using matrix decomposition methods like PCP to separate infrared target from background. The method of adding novel regularization terms to construct a more comprehensive model has not been widely researched.

In this paper, an approach based on total variation regularization and principal component pursuit (TV-PCP) is proposed. Same as PCP-based method, the input infrared image is firstly generalized to a patch image model. Despite the low-rank assumption, total variation (TV) regularization term is introduce to describe the inner smooth and crisp edges of background, which is of great importance in non-smooth and non-uniform background. The task of detecting infrared small target is formed as an optimization problem based on TV-PCP model. With proper solver, the separation of background and target will be achieved while automatically removing the effect of noise interference.

The main contributions of this paper are summarized as follows:

1. Considering the practical application scenario of infrared systems, an infrared dim target detection approach based on total variation regularization and principal component pursuit (TV-PCP) is proposed. By adding TV regularization to background model, this proposed method has a good performance in various scenes.
2. An optimization solver is proposed to solve this TV-PCP model. This solver is derived from alternating direction method (ADM). We implement the optimization of TV regularization term and nuclear term on a same variable, which can also be used in similar models of other applications.
3. In spite of getting good detection result, the proposed method can also get a good estimation of background, simultaneously. This estimation can be further used in motion detection, image registration and so on.

The remainder of this paper is organized as follows: In Section 2, the methodology of proposed method is presented, with detailed mathematical derivation. In Section 3, experiments on real infrared sequences and corresponding results are given, as well as the analysis of results. Conclusions are presented in Section 4.

2. Methodology

2.1. Total variation

Total variation model is proposed by Rudin (1992) to remove noise from gray level image [23]. It is also widely used in image deblurring applications. Without loss of generality, let $X \in R^{m\times n}$ denotes an image, then the TV norm is defined as:

$$\begin{aligned}\mathrm{TV}(X) = &\textstyle\sum_{i=1}^{m-1}\sum_{j=1}^{n-1}\sqrt{\left(x_{i,j}-x_{i+1,j}\right)^2+\left(x_{i,j}-x_{i,j+1}\right)^2}\\ &+\textstyle\sum_{i=1}^{m-1}\left|x_{i,n}-x_{i+1,n}\right|+\sum_{j=1}^{n-1}\left|x_{m,j}-x_{m,j+1}\right|.\end{aligned} \tag{1}$$

Where $x_{i,j}$ are the elements of X. Eq. (1) is called isotropic TV. From the above definition we can see that isotropic TV can be considered as the l_2 norms of image derivatives, if we do not consider the border of image. Let $D_iX \in R^2$ denotes the discrete gradient of X at pixel i, where X is vectorized as a column vector and D_i is the corresponding gradient operator. Then we have the following equation

$$TV(X) = \sum_i \|D_iX\|_2. \tag{2}$$

The TV norm is proved to be capable of preserving important edges and corners of images, which is always used as the regulation term when accurate estimation of discontinuities image parts is required [24]. In other words, TV norm represents the smoothness of a given image. It is also widely used in image decomposition [25], which can decompose an image into two parts: one is uncorrelated random patterns, while the other is sharp edges and piecewise-smooth components [26]. By minimizing the TV norm of an image, the smooth inner surface will be preserved while retaining crisp edges.

2.2. TV-PCP model

Generally, infrared images with small targets can be formulated as three parts: background region, target and noise interference. These three parts form an additive model:

$$f_I(x,y) = f_T(x,y) + f_B(x,y) + f_N(x,y). \tag{3}$$

Where $f_I(x,y)$ is the gray level of a pixel (x,y), $f_T(x,y)$ and $f_B(x,y)$ are the gray level of target and background region, respectively. $f_N(x,y)$ stands for the intensity of noise interference. To separate these three parts, proper features of each part should be extracted. Here we apply a newly proposed infrared patch image (IPI) model, which assumes that local image patches in distant regions of infrared background can be highly correlated with each other. These highly correlated image patches can be seen as low-rank matrix. At the same time, infrared small targets can be considered as sparse matrices. The construction and reconstruction method of IPI model can be found in [20]. The infrared additive model can be written as

$$I = T + B + N. \tag{4}$$

Where B, T N and I are the corresponding IPI model of background region, target area, noise part and original infrared image, respectively.

The infrared small target detection procedure in [20] based on IPI method is formulated as

$$\min_{B,T} \|B\|_* + \lambda\|T\|_1, \quad \text{s.t. } \|I-B-T\|_F \le \delta. \tag{5}$$

Here $\|\cdot\|_*$ is the nuclear norm of a matrix (i.e. the sum of singular values), $\|\cdot\|_1$ is the l_1 norm (i.e. $\|X\|_1 = \sum_{ij}|X_{ij}|$), $\|\cdot\|_F$ is the Frobenius norm (i.e. $\|X\|_F = \sqrt{\sum_{ij}X_{ij}^2}$), λ is a positive regularization parameter. δ represents the noise level of image. This is an optimization problem considering low-rank and sparse feature of background region and infrared small target area, with unavoidable level of noise. This problem can be solved by PCP, which is a state-of-art method of separating low-rank and sparse matrices. However, the performance of this PCP-based detection approach highly relies on the smoothness and uniformity of background region. When dealing with images captured from non-