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:

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

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-

smooth and non-uniform scenes, the low-rank assumption will cause a severe blur of estimated background. Some important detail of background, like edges and corners, will disappear. This degradation of accuracy can greatly influence the detection result. For example, some evident edges of background region could be considered as "sparse" parts, which will introduce clutter to detection result.

To realize a more precise detection of infrared target, the TV regularization term is brought into PCP model to preserve edges and corners of background region. We have the following TV-PCP model:

$$\min_{T,B} \|B\|_* + \lambda_1 TV(B) + \lambda_2 \|T\|_1, \quad \text{s.t. } I = T + B + N, \ \ \|N\|_F \le \delta. \tag{6}$$

Where $TV(\cdot)$ denotes the TV norm, N is the noise term. λ_1 and λ_2 are positive constants to balance three terms in the optimization problem. To have an unified format in the proposed model, we use Eq. (2) to replace $TV(B)$:

$$\min_{T,B} \|B\|_* + \lambda_1 \sum_i \|D_i B\|_2 + \lambda_2 \|T\|_1, \quad \text{s.t. } I = T + B + N, \ \ \|N\|_F \le \delta. \tag{7}$$

Here D_i is the gradient operator. We can see that despite the nuclear norm, TV norm is also applied on background patch model B, simultaneously. This will help to reduce clutters in detection results, as well as obtaining a good estimation of background region. After solving Eq. (7), we can get the infrared target image. By adding simple post-processing on target image, target information will be obtained. Fig. 1 shows the flow chart of proposed method.

2.3. Optimization algorithm

To solve the proposed TV-PCP model, we derive an optimization algorithm by decomposing the proposed model into several sub-problems. First of all, Eq. (7) is reformulated into an equivalent problem by splitting the correlated variables into independent variables:

$$\begin{aligned} \min_{Z_1,Z_2,Z_3} \ & \|Z_1\|_* + \lambda_1 \sum_i \|z_i\|_2 + \lambda_2 \|Z_3\|_1 \\ \text{s.t.} \ \ & Z_1 = B \\ & Z_2 = [z_1; z_2; \ldots; z_{mn}], \quad z_i = D_i B \\ & Z_3 = T \\ & I = T + B + N, \quad \|N\|_F \le \delta. \end{aligned} \tag{8}$$

The augmented Lagrangian function of Eq. (8) is

$$\begin{aligned} L_A = {} & \|Z_1\|_* + \lambda_1 \sum_i \|z_i\|_2 + \lambda_2 \|Z_3\|_1 + \langle Y_1, Z_1 - B\rangle + \frac{\beta}{2}\|Z_1 - B\|_F^2 \\ & + \sum_i \left(\langle y_i, z_i - D_i B\rangle + \frac{\beta_i}{2}\|z_i - D_i B\|_F^2 \right) + \langle Y_3, Z_3 - T\rangle + \frac{\beta}{2}\|Z_3 - T\|_F^2 \\ & + \langle Y_4, I - T - B - N\rangle + \frac{\beta}{2}\|I - T - B - N\|_F^2. \end{aligned} \tag{9}$$

Here $\langle \cdot \rangle$ denotes the inner product of matrixes. Y_i $(i = 1,2,3,4)$ are the Lagrange multipliers, where $Y_i (i = 1,3,4) \in R^{mn \times 1}$, $Y_2 = [y_1, y_2, \ldots, y_{mn}] \in R^{2 \times mn}$. β can be considered as a penalty parameter. For the convenience of calculation, Z_1, Z_3, T, B, N in Eq. (9) are vectorized to corresponding column vectors. This problem can be solved by alternating direction multiplier (ADM) method [27–28]. In each iteration, we firstly solve Z_i $(i = 1,2,3)$ sub-problems. Secondly, we solve N, B and T sub-problems. Finally, the Lagrange multipliers are updated. Note that when optimizing one variable, other variables are fixed.

For Z_1 sub-problem, it can be translated to

$$\begin{aligned} Z_1^{k+1} \leftarrow {} & \arg\min_{Z_1} L_A\left(Z_1, Z_2^k, Z_3^k, B^k, T^k, N^k\right) \\ = {} & \arg\min_{Z_1} \|Z_1\|_* + \left\langle Y_1^k, Z_1 - B^k \right\rangle + \frac{\beta}{2}\left\|Z_1 - B^k\right\|_F^2 \\ = {} & \arg\min_{Z_1} \|Z_1\|_* + \frac{\beta}{2}\left\|Z_1 - \left(B^k - \frac{Y_1^k}{\beta}\right)\right\|_F^2. \end{aligned} \tag{10}$$

The above problem can be solved by singular value thresholding [29]

$$Z_1^{k+1} = SVT_{\frac{1}{\beta}}\left(B^k - \frac{Y_1^k}{\beta}\right). \tag{11}$$

Where $SVT_\mu(\cdot)$ is the singular thresholding operator:

$$\begin{aligned} & SVT_\mu(Y) = U diag\left[(\sigma - \mu)_+\right] V^T \\ & (\sigma - \mu)_+ = \begin{cases} \sigma - \mu & \sigma > \mu \\ 0 & otherwise. \end{cases} \end{aligned} \tag{12}$$

Here $X = U\Sigma V^*$ denotes the singular value decomposition of X, $\Sigma = \mathrm{diag}(\{\sigma_i\}_{1 \le i \le r})$ are the singular values of X.

For Z_2 optimization problem, it can be translated as

$$\begin{aligned} Z_2^{k+1} \leftarrow {} & \arg\min_{Z_2} L_A\left(Z_1^k, Z_3^k, B^k, T^k, N^k\right) \\ = {} & \arg\min_{Z_2} \sum_i \left(\|z_i\|_2 + \left\langle y^k{}_i, z_i - D_i B^k \right\rangle + \frac{\beta_i}{2}\left\|z_i - D_i B^k\right\|_F^2 \right). \end{aligned} \tag{13}$$

Note that this is an l_2 optimization problem with constrain. It can be solved by a 2-D shrinkage-like formula

$$z_i = \max\left\{ \left\| D_i B - \frac{y_i}{\beta_i} \right\|_2 - \frac{1}{\beta_i}, 0 \right\} \cdot \frac{\left(D_i B - \frac{y_i}{\beta_i}\right)}{\left\| D_i B - \frac{y_i}{\beta_i} \right\|_2}. \tag{14}$$

For more details about this formula, please see [30].

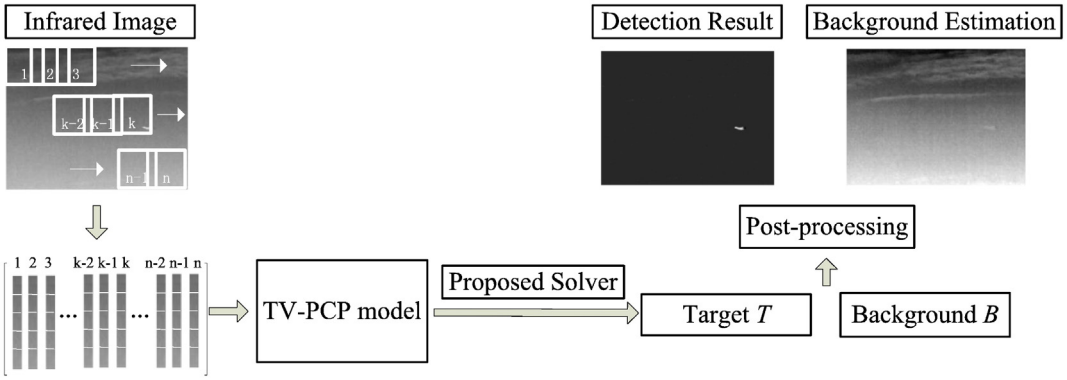

Fig. 1. Flow chart of the proposed method in this paper.

For Z_3 sub-problem, our goal is to solve

$$\begin{aligned} Z_3^{k+1} &\leftarrow \arg\min_{Z_3} L_A\left(Z_1^k, Z_2^k, B^k, T^k, N^k\right) \\ &= \arg\min_{Z_3} \lambda_2\|Z_3\|_1 + \left\langle Y_3^k, Z_3 - T^k\right\rangle + \frac{\beta}{2}\left\|Z_3 - T^k\right\|_F^2 \\ &= \arg\min_{Z_3} \lambda_2\|Z_3\|_1 + \frac{\beta}{2}\left\|Z_3 - \left(T^k - \frac{Y_3^k}{\beta}\right)\right\|_F^2. \end{aligned} \tag{15}$$

The solution of Eq. (15) is given by

$$Z_3^{k+1} = Th_{\frac{\lambda_2}{\beta}}\left(T - \frac{Y_3^k}{\beta}\right) \tag{16}$$

$$Th_\varepsilon(W) = \begin{cases} w-\varepsilon & w>\varepsilon \\ w+\varepsilon & w<-\varepsilon \\ 0 & otherwise \end{cases} \tag{17}$$

Where $Th_\varepsilon(\cdot)$ is the thresholding operator [31].

In the second step, we solve N, B and T optimization problems. For N sub-problem, it can be reformulated as

$$\begin{aligned} N^{k+1} &\leftarrow \arg\min_N \left\langle Y_4^k, I - T^k - B^k - N\right\rangle + \frac{\beta}{2}\left\|I - T^k - B^k - N\right\|_F^2 \\ &= \left\|N - \left(I - T^k - B^k + \frac{Y_4^k}{\beta}\right)\right\|_F^2, \qquad \text{s.t. } \|N\|_F \le \delta. \end{aligned} \tag{18}$$

Then the solution to Eq. (18) is given by a projection operator

$$N^{k+1} = P_\Omega\left(I - T^k - B^k + \frac{Y_4^k}{\beta}\right). \tag{19}$$

Ω denotes the sphere of $\|\cdot\|_F \le \delta$. P_Ω denotes the projection on the corresponding sphere.

For B sub-problem, it can be solved by

$$B^{k+1} \leftarrow \frac{\partial L_A}{\partial B} = 0. \tag{20}$$

Eq. (20) is a linear equation, which can be rewritten as

$$\begin{aligned} -\frac{\partial L_A}{\partial B} &= Y_1^k + \beta\left(Z_1^{k+1} - B\right) \\ &\quad + \sum_i \left[D_i^T y_i + \beta_i D_i^T (z_i - D_i B)\right] + Y_4^k + \beta\left(I - T^k - N^{k+1}\right) \end{aligned} \tag{21}$$

$$\begin{aligned} B^{k+1} = &\left(2\beta + \sum_i \beta_i D_i^T D_i\right)^{-1} \\ &\cdot \left[Y_1^k + Y_4^k + \sum_i \left(D_i^T y_i + \beta_i D_i^T z_i\right) + \beta\left(Z_1^{k+1} + I - T^k - N^{k+1}\right)\right]. \end{aligned} \tag{22}$$

Similar to Eqs. (21)–(22), T can be solved by

$$T^{k+1} \leftarrow \frac{\partial L_A}{\partial T} = 0 \tag{23}$$

$$T^{k+1} = \frac{Y_3 + \beta Z_3^{k+1} + Y_4 + \beta\left(I - B^{k+1} - N^{k+1}\right)}{2\beta} \tag{24}$$

In the third step, we upgrade Lagrange multipliers by

$$\begin{aligned} Y_1^{k+1} &\leftarrow Y_1^k + \gamma\beta\left(Z_1^{k+1} - B^{k+1}\right) \\ Y_2^{k+1} &\leftarrow Y_2^k + \gamma\beta\left(Z_2^{k+1} - DB^{k+1}\right) \\ Y_3^{k+1} &\leftarrow Y_3^k + \gamma\beta\left(Z_3^{k+1} - T^{k+1}\right) \\ Y_4^{k+1} &\leftarrow Y_4^k + \gamma\beta\left(I - T^{k+1} - B^{k+1} - N^{k+1}\right). \end{aligned} \tag{25}$$

Where $\gamma > 0$ is a pre-set step length, which can avoid the upgrade of β in each iteration and lead to a more stable solution. Thus, an efficient solver to the proposed TV-PCP model is presented based on the above analysis. This optimization algorithm is outlined in Algorithm 1.

Algorithm 1. Solving algorithm to TV-PCP model.

Input:
Infrared patch model $I \in R^{m\times n}$, β, λ_1, λ_2, δ
Output:
B, T
Initialize: $B^k = zeros(m,n)$, $T^k = zeros(m,n)$, $N^k = zeros(m,n)$
$Z_1 = Z_3 = Z_4 = zeros(m,n)$, $Y_1 = Y_3 = Y_4 = zeros(m,n)$
In the following iterations, the above matrices are vectorized as column vector.
$Z_2 = zeros(mn,2)$, $Y_2 = zeros(2,mn)$;
while not converged **do**
  $Z_1^{k+1} = D_{\frac{1}{\beta}}\left(B^k - \frac{Y_1^k}{\beta}\right)$
  Z_2^{k+1} is calculated by a 2-D shrinkage solver shown in (14)
  $Z_3^{k+1} = Th_{\frac{\lambda_2}{\beta}}\left(T - \frac{Y_3^k}{\beta}\right)$
  $N^{k+1} = P_\Omega\left(I - T^k - B^k + \frac{Y_4^k}{\beta}\right)$
  B_{k+1} is solved by (22)
  $T^{k+1} = \frac{Y_3 + \beta Z_3^{k+1} + Y_4 + \beta\left(I - B^{k+1} - N^{k+1}\right)}{2\beta}$
  Upgrade $Y_i (i = 1,2,3,4)$ as shown in (25)
  $k = k + 1$
end while

3. Experimental results

3.1. Parameter setting

The aforementioned TV-PCP model has several parameters which influence the performance of proposed model. Parameter $\lambda_i > 0$, $i = 1, 2$ plays a tradeoff between low-rank term, sparse term and TV regularization term. λ_1 is an empirical value, which should be set around 0.01. In our experiments, $\lambda_1 = 0.005$. For λ_2, it balance the value of nuclear norm and l_1 norm, while considering the influence of TV term. Here we set $\lambda_2 = \frac{1}{\sqrt{\min(m,n)}}$, where m and n are the size of input infrared patch image model. β is a penalty parameter in Lagrange function, which is set to 0.025. $\gamma = 1.5$ is an empirical value. In the proposed algorithm to solve TV-PCP model, we define tolerance error

$$tol = \frac{\left\|I - T^k - B^k - N^k\right\|_F}{\|I\|_F}. \tag{26}$$

Here k is the number of iteration. It is considered as converged when tolerance error $tol < 10^{-5}$ or max iteration number $MaxIter = 500$ is reached.

3.2. Infrared dim target detection result.

We test the proposed algorithm on 6 infrared sequences, which contain various scenes like sky, cloud, sea and ground. Detailed description of each image sequence is shown in Table 1.

Table 1
Detailed description of 6 infrared sequences.

No.	Background description	Target description
1	Sky scene with multilayer cloud. The brightness of background changes over a wide range.	Small round-shape target. Relatively low contrast to the surrounding region.
2	Bright cloudy background. Rich details in some part of the image.	One target with irregular shape and rotation during the whole sequence. The edges of target are not clear
3	Sea background. Have artificial architecture near the coast.	Strip-shaped target with a bright spot. Moving slowly during the sequence.
4	Cloudy changing background.	An airplane with changing size and shape. Random motion trajectory.
5	Heavy cloudy background. Non-smooth and non-uniform scene.	An airplane with long imaging distance. Size changes over a wide range.
6	Ground scene with heavy clutter. Bright interference in the background.	Person with small size. Move back and forth in the forest.

To evaluate the effectiveness of proposed method, we compare it with four baseline methods: Tophat, Maxmean, Maxmedian, and the IPI method. The filter size used in Maxmean and Maxmedian method are both 15×15. Parameters of IPI method are same to [20]. Detection results of representative images are shown in Fig. 2. The first column of Fig. 2 is the representative images of each test sequence. The second to the fifth column show the detection results of four baseline methods, and the last column shows the result of proposed method.

From Fig. 2 we can see that in the picture of our detection result, the region around target is more smooth and with less clutter. Or, to put it another way, our method can separate background region and target area more completely, which will lead to better detection performance. In particular, sequence No.1, 3, 4 and 5 have complex non-uniform background. Detection results by baseline methods are always with heavy clutter. Our method can enhance the target and reduce the impact of background clutter.

Numerical metrics are also introduced to give qualitative evaluation of detection results. The signal-to-clutter (SCR) is defined as

$$SCR = \frac{|\mu_t - \mu_b|}{\sigma_b}. \tag{27}$$

Where μ_t and μ_b are the average pixel value of target region and a neighborhood area around the target, respectively; σ_b is the corresponding standard deviation of neighborhood area. Similar to Ref. [32], the target area and the local background area is defined in Fig. 3.

SCR represents the degree of target obviousness in a local area. In this experiment, we set $d = 20$. Here we introduce SCR gain, which is defined as

$$G_{SCR} = \frac{SCR_{out}}{SCR_{in}}. \tag{28}$$

Where SCR_{in} and SCR_{out} stand for the SCR of original infrared image and the processed image (columns 2–6 of Fig. 2), respectively. Another evaluation indicator is background suppression factor (BSF)

$$BSF = \frac{C_{in}}{C_{out}}. \tag{29}$$

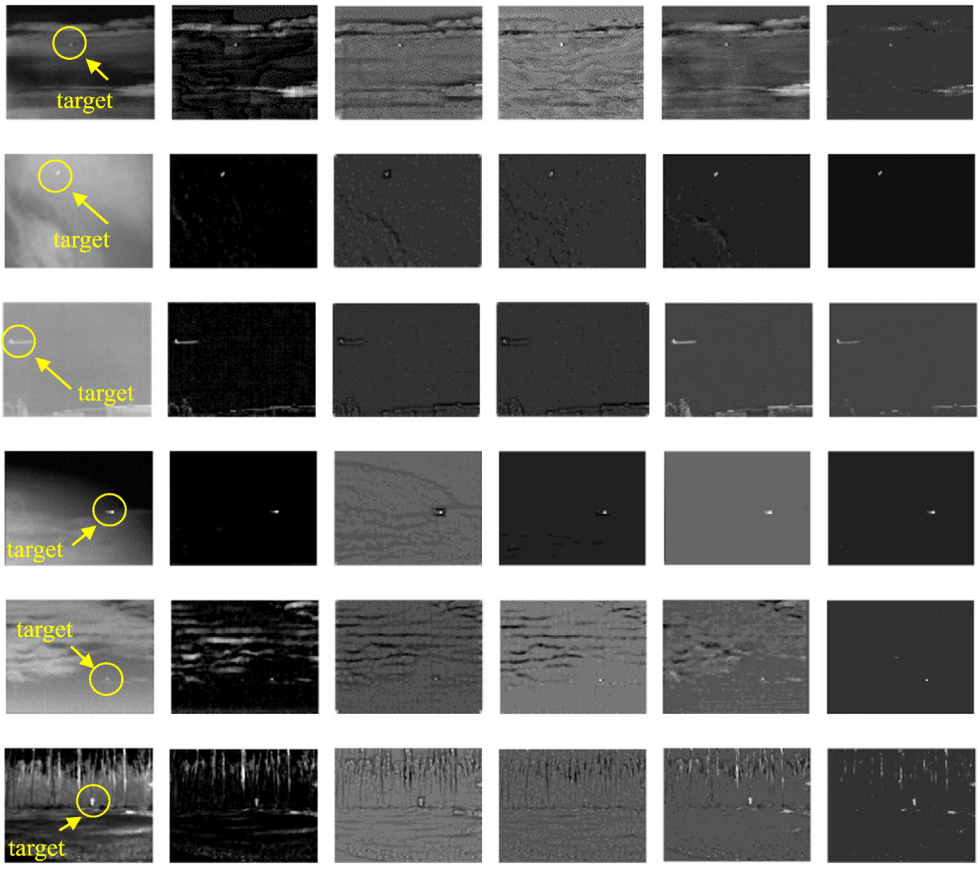

Fig. 2. Detection results of various methods in six test infrared sequences. Each line stands for a test image sequence. The first column is the representative image of test sequence 1 to 6, from the top to the bottom. The second column to the sixth column is the detection results of Tophat method, Maxmean method, Maxmedian method, IPI method and the proposed TV-PCP method, respectively.

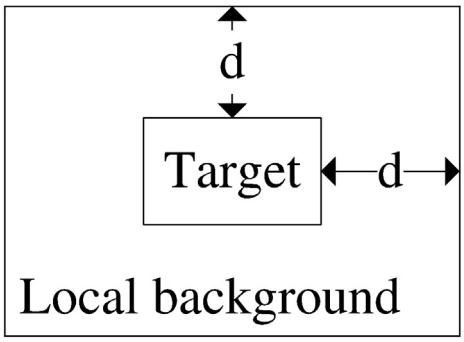

Fig. 3. The infrared target and local background area.

Where C_{in} and C_{out} are the standard deviation of background region in original image and processed image, respectively. Table 2 shows the average G_{SCR} and BSF for each detection method in six image sequences.

In Table. 2, the highest value of each evaluation indicator in each column is marked in bold red numbers, and the second highest value is marked in bold blue numbers. These indicators demonstrate that our method not only have the best ability in enhancing target from local background clutter, but also can suppress background well in the whole image. Despite the proposed method, IPI method has the second-best in three of six sequences, but there is a big gap between the performance of proposed method and IPI method in numerical indicator values. Overall, the proposed method has the greatest ability in target enhancement and background suppression.

To make a more intuitive display of the detection result, the 3D gray map of test images and detection results are shown in Fig. 4.

From Fig. 4 we can see that in original infrared images, there always exist background regions whose gray level is near or higher than targets. This leads to the difficulty in detecting small targets from infrared images. By contrast, results of our method have relatively smooth background in 3D gray map. It is easier to determine the location and size of target from the detection results of our method, with fewer false alarms.

Apart from obtaining good detection results, another advantage of the proposed method is the good background estimation performance. From the methodology we know that our method and IPI method have a similar mechanism of getting the estimation of background region. Thus we compare the background estimation result of these two methods. Fig. 5 shows the comparison results.

From Fig. 5, we can see that by adding TV regularization to background model, the background estimation results of proposed method are with high fidelity. This property is mainly demonstrated in well preserved corners and edges. We can also notice that sometimes in the recovered background of proposed method, a dim edge of the target can be seen. This is because of an intrinsic property of TV constrain, which is called staircase effect. Especially in the representative image of the third image sequence, we can see the image of ship body in estimated background image. This is because the target is strip-shaped, with a narrow body. The top and bottom edges of ship body are shown in estimated background due to the staircase effect. Therefore there exists a profile of the ship body. But in the left of this ship, there is a bright spot, which cannot be seen in Fig. 5. To our knowledge, this effect has no influence on the detection result of proposed method, which is also demonstrated in the numerical experiments. This background estimation result can be further used for background modeling, motion detection from changing background and image registration.

To give an overall evaluation of each method dealing with each image sequence, we introduce another two important metrics: the probability-of-detection P_d and the false-alarm rate F_a, which are defined as follows:

$$P_d = \frac{\text{number of true detections}}{\text{number of actual targets}} \tag{30}$$

Table 2
The average SCR gain and BSF of six sequences.

Detective methods	Evaluation indicators	Seq1	Seq2	Seq3	Seq4	Seq5	Seq6
Tophat	$\overline{G_{SCR}}$	0.9075	5.2639	1.0107	**43.3366**	2.5561	1.3781
	$\overline{BSF}$	1.1287	3.2335	0.8044	13.7273	0.6660	1.4532
Maxmean	$\overline{G_{SCR}}$	0.7825	0.6108	0.5703	1.9113	1.1106	0.2679
	$\overline{BSF}$	3.0769	8.9652	2.5080	14.4389	3.5018	2.5119
Maxmedian	$\overline{G_{SCR}}$	**1.6628**	2.8017	**1.4931**	21.6923	3.6981	0.1983
	$\overline{BSF}$	**3.3240**	16.0503	**3.8095**	**34.8797**	3.7469	2.9462
IPI method	$\overline{G_{SCR}}$	0.7924	**12.2334**	1.0723	21.6567	**5.8849**	**2.3603**
	$\overline{BSF}$	2.5291	**23.3861**	2.1022	9.3511	**5.3105**	**3.0878**
Proposed method	$\overline{G_{SCR}}$	**3.8708**	**18.2471**	**1.6735**	**63.2282**	**77.0683**	**3.5340**
	$\overline{BSF}$	**14.5979**	**122.2180**	**5.9564**	**49.2204**	**67.8014**	**5.8583**

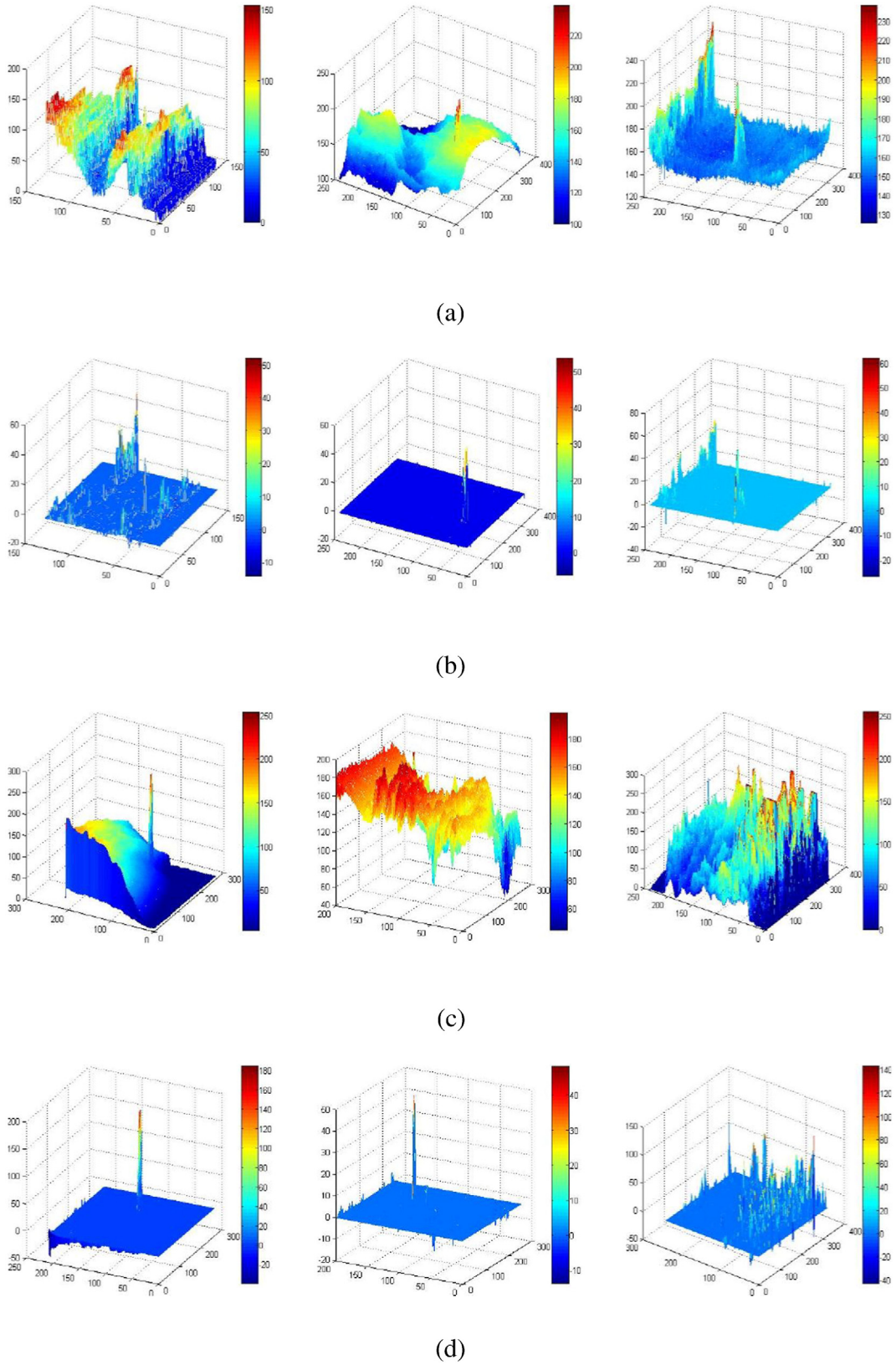

Fig. 4. 3D gray maps of Seq 4–6 and corresponding detection result. (a) Original images of Seq 1–3. (b) Corresponding detection results of proposed method. (c) Original images of Seq 4–6. (d) Corresponding detection results of proposed method.

$$F_a = \frac{\text{number of pixels in false detections}}{\text{number of images}}. \tag{31}$$

Then we give the receiver operation characteristic (ROC) curve of five methods. The ROC curve characterizes the changing of P_d in different level of F_a. Also, we calculate the area under curve (AUC) of each

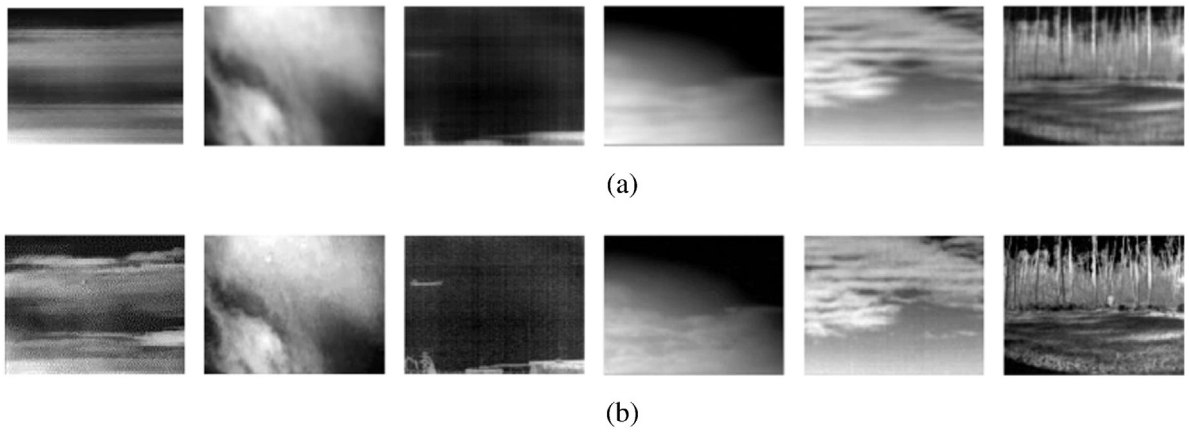

Fig. 5. Estimation of background region by IPI method and proposed method. (a) IPI method; (b) the proposed method.

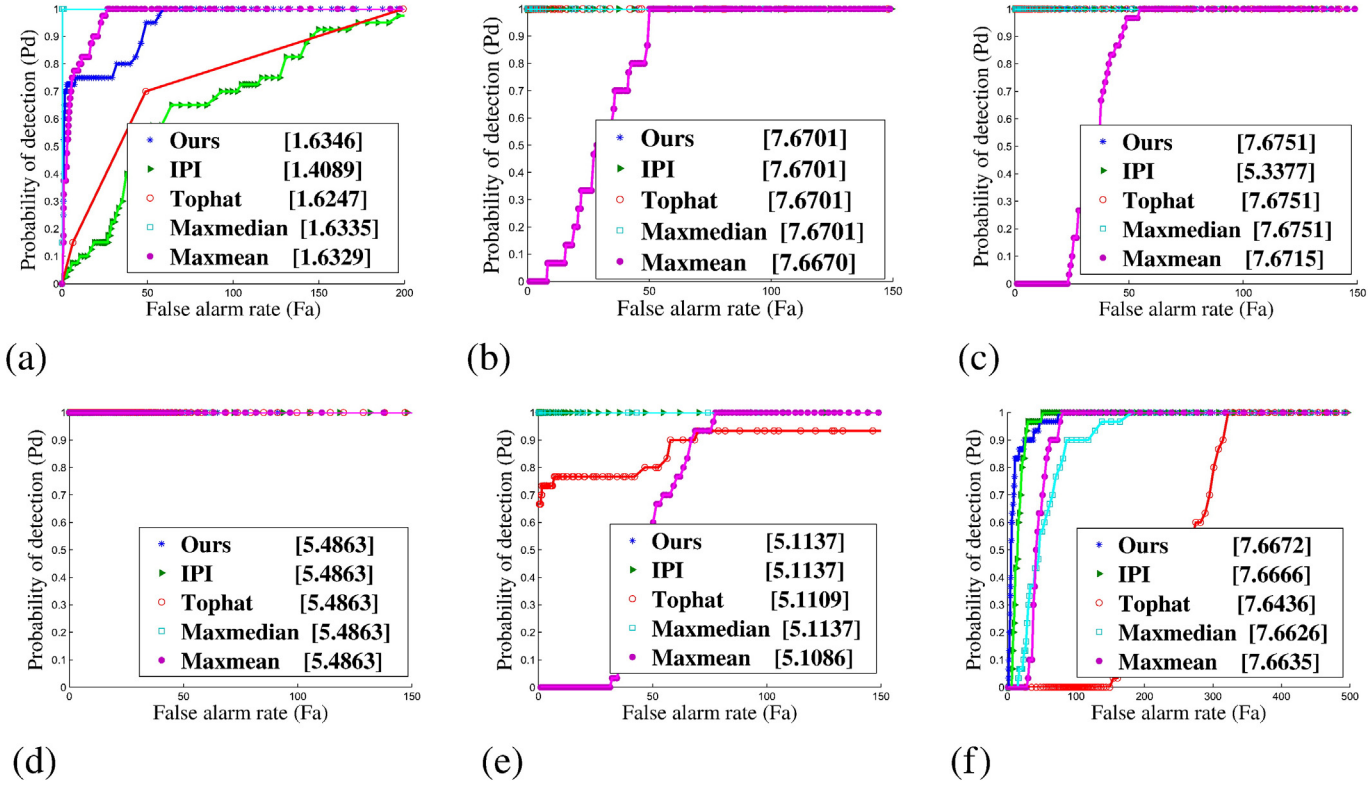

Fig. 6. ROC curve and AUC value ($\times 10^4$) of image sequence 1 to sequence 6.

method, and the AUC values are shown in the figures. The larger AUC value means a better target detection performance. The ROC curve of six image sequences with AUC values are shown in Fig. 6.

In Fig. 6 (a), we can see that the proposed method reaches $P_d = 1$ a little slower than Maxmedian method, but still much faster than other three methods. This can demonstrate the contribution of adding TV regularization term to PCP model. In Fig. 6 (b), these methods have similar detection performance in ROC curve except Maxmean method. In Fig. 6 (c), the Maxmean method is the last one reaching $P_d = 1$. But from the AUC value, we can see that IPI method has a bad performance comparing with other four methods. In Fig. 6 (d), the performance of Tophat method and Maxmean method are not very good. The proposed method and IPI method have similar detection performance. In Fig. 6 (e), the detection performance distinguish degree of five methods are not very high. In Fig. 6 (f), the P_d of proposed method reaches 0.9 at a very low level of F_a, which is much better than other four methods. Although the IPI method reaches $P_d = 1$ faster than the proposed method, the overall detection ability is worse than the proposed method. The AUC value of proposed method is the highest in all the six ROC curves. This means that considering both P_d and F_a of each image, the proposed method has the best detection performance. These six image sequences are with non-smooth and non-uniform background, the ROC results further explained the reasonability and effectiveness of TV-PCP model when dealing with scenes with abundant background region details. This ability of proposed method makes it more useful in practical applications. We should also note that ROC curve only describes the relationship of F_a and P_d in each detection result. All indicators are of equal importance in method evaluation. Synthesizing the above evaluation metrics, we know that the proposed TV-PCP method has the best detection performance comparing to baseline methods.

4. Conclusion

This paper proposes a TV-PCP model to detect dim target in infrared images, especially in non-smooth and non-uniform images. The main idea is to bring TV regularization term to traditional PCP model. Despite the empirical low-rank feature of non-local self-correlation background, the TV regularization term can describe the inner smooth and crisp edges of non-uniform background. This TV-PCP model can not only hold the overall distribution of background pixels, but also retain detailed edges and corners. With a proper optimization algorithm, the dim target can be detected through an optimization problem. By accurately recovering the background, the performance of target detection can be improved. It has been shown through numerical examples that the proposed TV-PCP model provides an improved SCR gain as well as improved detection ability compared to four baseline methods under diverse non-smooth and non-uniform backgrounds.

Acknowledgement

This work is supported by the National Science Foundation of China (61571096, 41274127, 41301460, 61308102) and the Key Laboratory Fund of Beam Control, Chinese Academy of Sciences (2014LBC002).

References

[1] M.M. Hadhoud, D.W. Thomas, The two-dimensional adaptive LMS (TDLMS) algorithm, IEEE Trans. Circuits Syst. 35 (5) (1988) 485–494.
[2] M. Zeng, J. Li, Z. Peng, The design of top-hat morphological filter and application to infrared target detection, Infrared Phys. Technol. 48 (1) (2006) 67–76.
[3] G. Davidson, H.D. Griffiths, Wavelet detection scheme for small targets in sea clutter, Electron. Lett. 38 (19) (2002) 1128–1130.
[4] Y. Cao, R.M. Liu, J. Yang, Infrared small target detection using PPCA, Int. J. Infrared Millimeter Waves 29 (4) (2008) 385–395.
[5] I.S. Reed, R.M. Gagliardi, L.B. Stotts, Optical moving target detection with 3-D matched filtering, IEEE Trans. Aerosp. Electron. Syst. 24 (4) (1988) 327–336.
[6] R.W. Fries, Three dimensional matched filtering, OE/LASE '89, Los Angeles 1989, pp. 19–27.
[7] Y. Barniv, Dynamic programming solution for detecting dim moving targets, IEEE Trans. Aerosp. Electron. Syst. 1 (1985) 144–156.
[8] L.A. Johnston, V. Krishnamurthy, Performance analysis of a dynamic programming track before detect algorithm, IEEE Trans. Aerosp. Electron. Syst. 38 (1) (2002) 228–242.
[9] S.D. Blostein, T.S. Huang, Detection of small moving objects in image sequences using multistage hypothesis testing, 1988 International Conference on Acoustics, Speech, and Signal Processing, Seattle 1988, pp. 1068–1071.
[10] S.D. Blostein, T.S. Huang, Detecting small, moving objects in image sequences using sequential hypothesis testing, IEEE Trans. Signal Process. 39 (7) (1991) 1611–1629.
[11] S.M. Tonissen, Y. Bar-Shalom, Maximum likelihood track-before-detect with fluctuating target amplitude, IEEE Trans. Aerosp. Electron. Syst. 34 (3) (1998) 796–809.
[12] K.L. Anderson, R.A. Iltis, A tracking algorithm for infrared images based on reduced sufficient statistics, IEEE Trans. Aerosp. Electron. Syst. 33 (2) (1997) 464–472.
[13] Q. Chen, Y. Wang, A small target detection method in infrared image sequences based on compressive sensing and background subtraction, 2013 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC), Kun-Ming 2013, pp. 1–3.
[14] Y. Liu, Y. Zhao, M. Liu, et al., Research on the algorithm of infrared target detection based on the frame difference and background subtraction method, Signal and Data Processing of Small Targets 2015, San Diego 2015, pp. 1–9.
[15] Z. Li, Z. Dai, H. Fu, et al., Dim moving target detection algorithm based on spatio-temporal classification sparse representation, Infrared Phys. Technol. 67 (2014) 273–282.
[16] C.L.P. Chen, H. Li, Y. Wei, et al., A local contrast method for small infrared target detection, IEEE Trans. Geosci. Remote Sens. 52 (12) (2014) 574–581.
[17] J. Han, Y. Ma, B. Zhou, et al., A Robust infrared small target detection algorithm based on human visual system, IEEE Geosci. Remote Sens. Lett. 11 (12) (2014) 2168–2172.
[18] J. Han, Y. Ma, J. Huang, et al., An infrared small target detecting algorithm based on human visual system, IEEE Geosci. Remote Sens. Lett. (2016).
[19] K. Shang, X. Sun, J. Tian, et al., Infrared small target detection via line-based reconstruction and entropy-induced suppression, Infrared Phys. Technol. 76 (2016) 75–81.
[20] C.Q. Gao, D.Y. Meng, Y. Yang, et al., Infrared patch-image model for small target detection in a single image, IEEE Trans. Image Process. 22 (12) (2013) 4996–5009.
[21] Y. He, M. Li, J. Zhang, et al., Small infrared target detection based on low-rank and sparse representation, Infrared Phys. Technol. 68 (2015) 98–109.
[22] L. Li, L. Hui, L. Tian, et al., Infrared small target detection in compressive domain, Electron. Lett. 50 (7) (2014) 510–512.
[23] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1–4) (1992) 259–268.
[24] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imaging Sci. 2 (1) (2009) 183–202.
[25] W. Yin, D. Goldfarb, S. Osher, A comparison of three total variation based texture extraction models, J. Vis. Commun. Image Represent. 18 (3) (2007) 240–252.
[26] M. Liu, X. Chen, X. Wang, Latent fingerprint enhancement via multi-scale patch based sparse representation, IEEE Trans. Inf. Forensics Secur. 10 (1) (2015) 6–15.
[27] Z. Wang, H. Li, Q. Ling, et al., Robust temporal-spatial decomposition and its applications in video processing, IEEE Trans. Circuits Syst. Video Technol. 23 (3) (2013) 387–400.
[28] C. Gao, J. Zhang, D. Li, et al., Robust Kronecker product video denoising based on fractional-order total variation model, Signal Process. 119 (2015) 1–20.
[29] J. Cai, E.J. Candès, Z. Shen, A singular value thresholding algorithm for matrix completion, SIAM J. Optim. 20 (4) (2010) 1956–1982.
[30] C. Li, An Efficient Algorithm for Total Variation Regularization with Applications to the Single Pixel Camera and Compressive Sensing, Rice University, Houston, 2009.
[31] Z. Lin, M. Chen, Y. Ma, The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-rank Matrices (arXiv preprint arXiv:1009.5055) 2010.
[32] C. Gao, T. Zhang, Q. Li, Small infrared target detection using sparse ring representation, IEEE Aerosp. Electron. Syst. Mag. 27 (3) (2012) 21–30.