

Infrared small-dim target detection based on Markov random field guided noise modeling



Chenqiang Gao^{a,*}, Lan Wang^a, Yongxing Xiao^a, Qian Zhao^b, Deyu Meng^b

^a Chongqing Key Laboratory of Signal and Information Processing, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

^b Institute for Information and System Sciences and Ministry of Education Key Lab of Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China

ARTICLE INFO

Article history:

Received 23 February 2017

Revised 19 October 2017

Accepted 16 November 2017

Available online 20 November 2017

MSC:

00-01

99-00

Keywords:

Infrared image

Small target detection

Mixture of Gaussians

Markov random field

Variational Bayesian

ABSTRACT

Small target detection is one of the key techniques in infrared search and tracking applications. When small targets are very dim and of low signal-to-noise ratio, they are very similar to background noise, which usually causes high false alarm rates for conventional methods. To address this problem, we novelly treat the small-dim targets as a special sparse noise component of the complex background noise and adopt Mixture of Gaussians (MoG) with Markov random field (MRF) to model this problem. Firstly, the spatio-temporal patch image is constructed using several consecutive frames to utilize the temporal information of the image sequence. Then, the MRF guided MoG noise model under the Bayesian framework is proposed to model the small target detection problem. After that, by variational Bayesian, the small target component can be effectively separated from complex background noise. Finally, a simple adaptive segmentation method is used to extract small targets. Several series of experiments are done to evaluate the proposed method and the results show that the proposed method is robust for real infrared images with complex background.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Infrared small target detection is a key technique in many areas, including space surveillance systems, early-warning systems, object tracking systems, etc. Its task is to localize targets, e.g., boats in the sea, airplanes in the sky, vehicles in the land, etc., as shown as Fig. 1 [1], in infrared images. Due to the long imaging distance, these targets are usually of very small sizes. Besides, the cloudy clutter, sea clutter or other clutter makes the background very complex and thus the targets are usually of low signal-to-noise ratio (SNR). Although the community has made a good progress on this task in past decades [1–4], it still remains an open problem, due to these challenges.

Up to now, a large number of approaches have been proposed. Some of them use the spatio-temporal cues to detect small targets. There are two representative categories among these methods: detection before track (DBT) [5–7] and track before detection (TBD) [8–10]. DBT can exploit the continuity of target's trajectories to reject the false targets in primary detection results obtained by single frame based detection methods. Thus, the performance of this kind of methods greatly depends on detection results from the sin-

gle frame. In contrast, TBD can enhance the target signal energy by seeking the potential target trajectory and then accumulating the signal energy of the target along the trajectory before detecting targets. In this way, the enhanced target can be more robustly detected. The classical methods include 3D matched (directional) filters [11–13] and other spatio-temporal methods [14–17]. Generally, the methods using spatio-temporal information depend on the assumption of the motion continuity of targets. Whether this assumption is true in practical applications would influence the final detection performance.

Different from previous methods, many other methods just use spatial information to detect targets. It is usually assumed that an infrared image $f_F(x, y)$ can be formulated as a combination of three components, which are a background component $f_B(x, y)$, a target component $f_T(x, y)$, and a noise component $f_N(x, y)$, respectively. Some methods attempted to firstly predict the background component $f_B(x, y)$, and then extract targets from the difference image between $f_F(x, y)$ and $f_B(x, y)$. The representative methods include Top-Hat filtering [18], Max-Median filtering [19] and others methods [20–22]. In contrast, other methods directly model the target component $f_T(x, y)$ based on the infrared small target characteristics, such as GST [3], edge directional 2D LMS filter [23], sparse ring representation [24], modified gaussian function [25]. Since these methods just focus on one aspect of the infrared image compo-

* Corresponding author.

E-mail address: gaocq@cqupt.edu.cn (C. Gao).

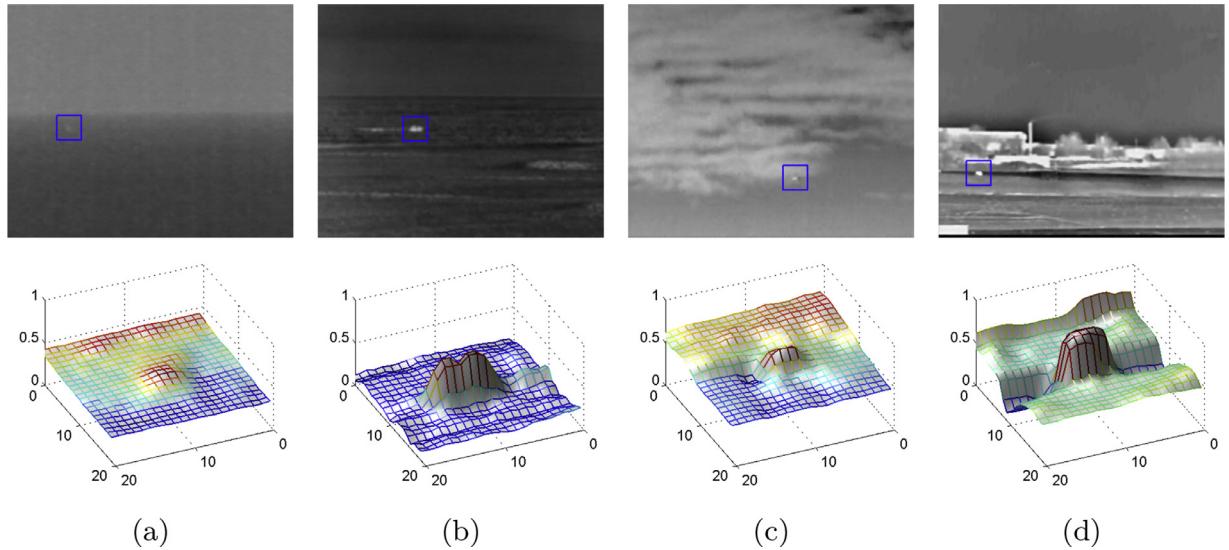


Fig. 1. Representative targets (upper) and the corresponding 3-D surfaces (lower) in different backgrounds (normalized) [1]. (a) A dim small ship target in sea-sky background. (b) A bright ship target in sea-sky background. (c) A dim aeroplane target in sky cloud background. (d) A bright vehicle target in sky-ground background.

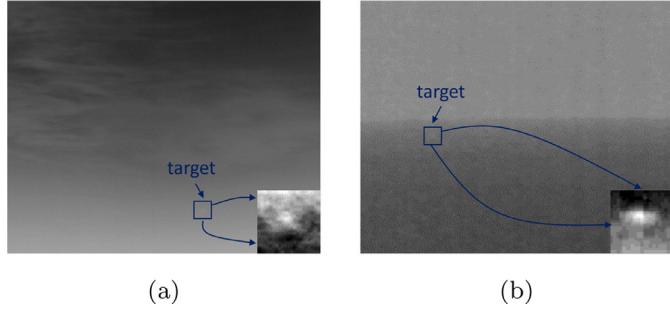


Fig. 2. Two representative small-dim target images. (a) A dim small ship target in sea-sky background. (b) A dim aeroplane target in sky cloud background.

ments with slightly strong assumptions, they are usually suitable for specific applications while not generalizing well to others.

Inspired by the recent advances in low-rank matrix analysis [26–29], the state-of-the-art method [1] is to jointly consider all three components, namely $f_B(x, y)$, $f_T(x, y)$ and $f_N(x, y)$, in the low-rank framework. With an effective image construction method, the background and target components are approximately transformed into a low-rank matrix and a sparse matrix, respectively, and thus an infrared image can be seen as a combination of a low-rank matrix, a sparse matrix and a noise matrix. By applying the accelerated proximal gradient (APG) approach [30], the target and background components can be concurrently and effectively recovered.

However, for the case of complex background noise, the current assumption for noise is simple, which slightly ignores the influence of noise for the small target detection task. As a result, the model could not well match the practical problem with heavy noise and this would influence the robustness of small target detection.

In this paper, we focus on the problem of small target detection in the case that targets are not only small, but also dim. These characteristics make their SNRs so low that targets almost approximate to noise, as shown as Fig. 2. In this situation, it is difficult to model separately the small target and noise components. The conventional methods usually have high false alarm rates on this task since there would be a lot of noise/clutter residual in the target image. To address this challenging problem, we do not explicitly discriminate the target from noise. Instead, we model the target component $f_T(x, y)$ and noise component $f_N(x, y)$ together, and as-

sume that the target is a component of complex noise. Then we adopt the mixture of Gaussians (MoG) noise model [29] to model the complex noise. Due to the sparse property of the small target, the corresponding component is generally significant different from the rest components of the MoG. Thus, the small target can be separated from the complex noise. Besides, the adjacent pixels of small targets are usually dependent each other, while the noise pixels are random. Thus, we adopt the Markov random field (MRF) model to guide the separation of small targets from noise, which makes the detected small targets full shapes. In order to tackle the challenge of the low SNR, the spatio-temporal information is utilized and this is different from the state-of-the-art work [1] which is just based on spatial information in a single image.

The remainder of this paper is organized as follows. Section 2 describes the proposed method in detail, including the spatio-temporal patch image model, problem formulation, solution and the small target extraction framework. Experiments and comparisons between the proposed method and the baseline methods are provided in Section 3. Conclusions are given in Section 4.

2. The proposed method

In this section, we first introduce the construction and reconstruction method of a spatio-temporal patch image. Then, we describe the formulation of small target detection problem based on MoG and its solution in detail. After that, we present small target extraction and introduce the full framework of the proposed method, including the implementation steps. Finally, we analyze the computational complexity of the proposed method.

2.1. Spatio-temporal patch image

Similar to our previous work [1], given an image sequence f_1, f_2, \dots, f_d and a cubic sliding window with a size of $w \times h \times d$, we can obtain a series of cubic patches. Then, a 2D matrix can be constructed by orderly vectorizing cubic patches as its columns, as shown in Fig. 3. Contrariwise, as shown in Fig. 4, after being processed, the constructed 2D matrix can be reconstructed into an image sequence with d frames by the inverse processing with a minor modification. Namely, for the pixel with overlap patches, its value is determined by pooling multiple different values into one. In this

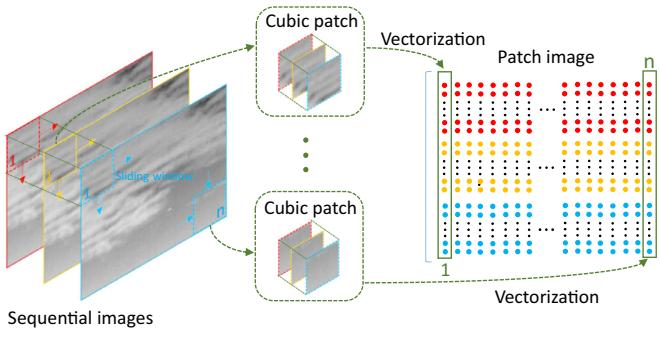


Fig. 3. The construction of the spatio-temporal patch image with sequential frames.

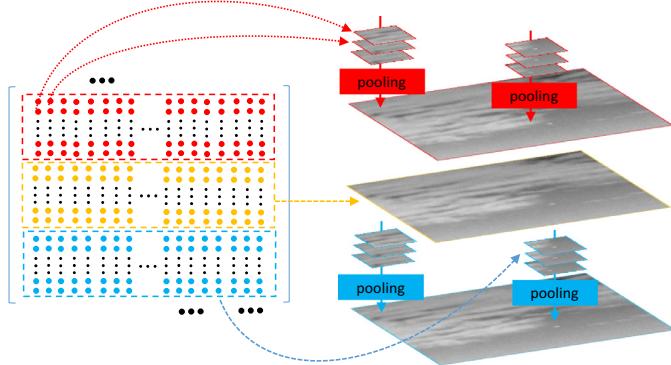


Fig. 4. The reconstruction of the temporal-spatial patch image.

paper, we experimentally choose the *median pooling* method as did in our previous work [1]. Here, our spatio-temporal patch image is a generalization of the spatial patch image in our previous work [1], in order to utilize the temporal information of small targets.

2.2. Formulation

Assume that an image sequence (f_1, f_2, \dots, f_d) containing small targets is constructed into a 2D patch image matrix F with a size of $m \times n$ by the construction method introduced in Section 2.1. F can be seen as a combination of background B and a complex noise E as follows:

$$F = B + E. \quad (1)$$

Here the small target component T is integrated into the noise E , since the targets can be seen as a component of noise in our problem discussed previously. We consider the Eq. (1) as a generative model to obtain the B and E .

Background component B modeling: According to the infrared imaging mechanism, thermal radiation is affected by the atmospheric scattering and diffraction when transmitting in the atmosphere from a long distance. Therefore, the energy of the thermal radiation captured by the infrared imaging sensor is low and the background image is usually blurred. Though there are heavy local fluctuations in sky and sea-sky backgrounds, the background still changes slowly on the whole and the non-local background blocks are usually similar as discussed in [1]. Thus, the cubic patches are approximately correlated to each other. As a result, B can be considered as a low rank matrix. We model B as follows [28]:

$$B = UV^T = \sum_{r=1}^R u_r v_r^T, \quad (2)$$

where $U \in \mathbb{R}^{m \times R}$, $V \in \mathbb{R}^{n \times R}$, R is the initial rank value of B , and u_r (v_r) is the r th column of U (V).

To guarantee the low rank nature of B , we need to ensure that U and V are of column sparsity. This goal can be achieved by imposing the following priors on U and V :

$$u_r \sim \mathcal{N}(u_r | \mathbf{0}, \gamma_r^{-1} \mathbf{I}_m), v_r \sim \mathcal{N}(v_r | \mathbf{0}, \gamma_r^{-1} \mathbf{I}_n), \quad (3)$$

where \mathbf{I}_m denotes a $m \times m$ identity matrix. The conjugate priors on each γ_r is:

$$\gamma_r \sim \text{Gam}(\gamma_r | a_0, b_0), \quad (4)$$

where $\text{Gam}(\gamma_r | a_0, b_0)$ is a gamma distribution parameterized by a_0 and b_0 . It has been validated that such a modeling could lead to large precision values of some γ s, and hence result in a good low-rank estimate of B [28].

Noise component E modeling: As discussed previously, the small-dim targets in infrared images can be regarded as a special kind of noise. Therefore, we can model the small targets and background noise together. This can be achieved by using the mixture of Gaussians, which is a universal approximator to any continuous probability distribution [31] and has been verified to be effective in modeling the complex noise [29,32]. Thus, the noise E can be modeled as a mixture of K Gaussians:

$$E_{ij} \sim \sum_{k=1}^K \pi_k \mathcal{N}(x | \mu_k, \tau_k^{-1}), \quad (5)$$

where E_{ij} is the element of E at the position (i, j) and $\mathcal{N}(x | \mu_k, \tau_k^{-1})$ represents the Gaussian distribution with mean μ_k and precision τ_k . The precision is normally defined as the reciprocal of the variance. π_k is the mixing proportion with $\pi_k \geq 0$ and $\sum_{k=1}^K \pi_k = 1$. The precision τ_k determines the probability distribution shape and thus can constrain the sparsity degree of the noise. For example, when τ_k is large, most of the elements of the k th noise component would approach zeros. As a result, this noise component can be considered as a sparse noise matrix.

For the convenience of inference, Eq. (5) can be equivalently expressed as a two-level generative model by introducing the indicator variables z_{ijk} [29]:

$$E_{ij} | Z_{ij} \sim \prod_{k=1}^K \mathcal{N}(E_{ij} | \mu_k, \tau_k^{-1})^{z_{ijk}}, \\ Z_{ij} \sim \text{Multinomial}(Z_{ij} | \pi), \quad (6)$$

where $Z_{ij} = (z_{ij1}, \dots, z_{ijk}) \in \{0, 1\}^K$, $\sum_{k=1}^K z_{ijk} = 1$ and Z_{ij} follows a multinomial distribution parameterized by π . Besides, the priors of the parameters in the MoG are brought in the model to further complete the Bayesian framework:

$$\mu_k, \tau_k \sim N(\mu_k | \mu_0, (\beta_0 \tau_k)^{-1}) \text{Gam}(\tau_k | c_0, d_0), \\ \pi \sim \text{Dir}(\pi | \alpha_0), \quad (7)$$

where $\text{Dir}(\pi | \alpha_0)$ is a dirichlet distribution parameterized by $\alpha_0 = (\alpha_{01}, \dots, \alpha_{0K})$.

Obviously, compared to current research works, our method has more feasibility since our noise modeling can handle more complex situations. To further match the practical problem, we will embed the spatio-temporal continuity prior into our model in the following section.

2.3. MRF guided MoG

Considering the Markov property of the image pixels, the characteristics of the pixel is mainly affected by the adjacent pixels. If a pixel belongs to a small target, its adjacent pixels will also probably belong to this target. Similarly, if a pixel is noise, its adjacent pixels have higher probability to be noise, too. In order to effectively utilize this property to help separate the small target component from other noise components, we introduce a

four-neighborhood Markov Random Field (MRF) [33] into the MoG model as follows:

$$\begin{aligned} E_{ij}|Z_{ij} &\sim \prod_{k=1}^K \mathcal{N}(E_{ij}|\mu_k, \tau_k^{-1})^{z_{ijk}}, \\ Z_{ij} &\sim \text{Multinomial}(Z_{ij}|\pi) \\ \frac{1}{C} \prod_k \prod_{(p,q) \in N(i,j)} &\exp(\lambda(2z_{ijk} - 1)(2z_{pqk} - 1)), \end{aligned} \quad (8)$$

where $N(i, j)$ denotes the four-neighborhood of the pixel (i, j) , and C is a normalization constant. λ is the parameter of the MRF, and can be tuned to adjust the influence of the adjacent pixels. If the central pixel and adjacent pixels are generated by the same Gaussian component, the bigger weight will be assigned to the multinomial distribution. This can increase the probability that the central pixel is generated from this Gaussian component.

Combining the Eqs. (4)–(8), given F , our goal is to infer the posterior of all involved variables:

$$p(U, V, \mathcal{Z}, \mu, \tau, \pi, \gamma | F), \quad (9)$$

where $\mathcal{Z} = \{Z_{ij}\}$, $\mu = (\mu_1, \dots, \mu_K)$, $\tau = (\tau_1, \dots, \tau_K)$ and $\gamma = (\gamma_1, \dots, \gamma_R)$.

2.4. Solution by variational Bayesian

As in [29], we use the variational Bayesian (VB) method [31] to infer the posterior of Eq. (9). Here, we just briefly describe the inference method and the details can be found in [29]. VB computes the posterior distribution approximation $q(X)$ by minimizing the Kullback-Leibler(KL) divergence to $p(X|D)$:

$$\begin{aligned} q^*(X) &= \min_{q \in \Omega} KL(q||p) \\ &= \min_{q \in \Omega} \left\{ - \int q(X) \ln \left\{ \frac{p(X|D)}{q(X)} \right\} dX \right\}, \end{aligned} \quad (10)$$

where Ω denotes the set of probability densities with certain restrictions to make the minimization tractable. Using mean field approximation, we employ the posterior factorization $q(X) = \prod_i q_i(x_i)$ such that the posterior distribution of each unknown is estimated by holding the others fixed and using their most recent distributions:

$$q_j^*(x_j) = \frac{\exp \left\{ \langle \ln p(X, D) \rangle_{X \setminus x_j} \right\}}{\int \exp \left\{ \langle \ln p(X, D) \rangle_{X \setminus x_j} \right\} d_{x_j}}, \quad (11)$$

where $\langle \cdot \rangle$ denotes the expectation and $X \setminus x_j$ denotes the set of X with x_j removed.

According to the mean field theory, the variational approximation distribution of Eq. (9) can be presented as:

$$\begin{aligned} q(U, V, \mathcal{Z}, \mu, \tau, \pi, \gamma) &= \prod_i q(u_i) \prod_j q(v_j) \\ &\quad \prod_{ij} q(z_{ij}) \prod_k q(\mu_k, \tau_k) q(\pi) \prod_r q(\gamma_r), \end{aligned} \quad (12)$$

where u_i, v_j denotes $i-th(j-th)$ row of $U(V)$. Since all distributions in MoG noise model are in the conjugate exponential family, the form of each posterior approximation can be found without major difficulties.

Estimation of noise component: The parameters involved in the noise component are μ, τ, \mathcal{Z} and π . Based on the prior imposed in Eq. (7) and its conjugate property, we can get the following update equation for each μ_k, τ_k ($k = 1, \dots, K$):

$$q(\mu_k, \tau_k) = \mathcal{N}(\mu_k | m_k, (\beta_k \tau_k)^{-1}) \text{Gam}(\tau_k | c_k, d_k), \quad (13)$$

where

$$\begin{aligned} \beta_k &= \beta_0 + \sum_{ij} \langle z_{ijk} \rangle, \\ m_k &= \frac{1}{\beta_k} \left(\beta_0 \mu_0 + \sum_{ij} \langle z_{ijk} \rangle (f_{ij} - \langle u_i \rangle \langle v_j \rangle^T) \right), \\ c_k &= c_0 + \frac{1}{2} \sum_{ij} \langle z_{ijk} \rangle, \\ d_k &= d_0 + \frac{1}{2} \left\{ \sum_{ij} \langle z_{ijk} \rangle ((f_{ij} - \langle u_i \rangle \langle v_j \rangle^T)^2) + \beta_0 \mu_0^2 \right. \\ &\quad \left. - \frac{1}{\beta_k} \left(\sum_{ij} \langle z_{ijk} \rangle (f_{ij} - \langle u_i \rangle \langle v_j \rangle^T) + \beta_0 \mu_0 \right)^2 \right\}. \end{aligned}$$

Similarly, it is easy to obtain the update equation for mixing proportions π :

$$q(\pi) = \text{Dir}(\pi | \alpha), \quad (14)$$

where $\alpha = (\alpha_1, \dots, \alpha_K)$, $\alpha_k = \alpha_{0k} + \sum_{ij} \langle z_{ijk} \rangle$.

The MoG used in this paper is constrained by the MRF, the effect of which is clearly found in the updating of the indicators \mathcal{Z} . The variational posterior for \mathcal{Z} can be presented as:

$$q(Z_{ij}) = \prod_k r_{ijk}^{z_{ijk}}, \quad (15)$$

where

$$r_{ijk} = \frac{\rho_{ijk}}{\sum_k \rho_{ijk}}, \quad (16)$$

$$\begin{aligned} \rho_{ijk} &= \frac{1}{2} \ln \tau_k + \ln \pi_k + \lambda \sum_{(p,q) \in N(i,j)} \langle z_{pqk} \rangle \\ &\quad - \frac{1}{2} \tau_k \left((f_{ij} - \langle u_i \rangle \langle v_j \rangle^T - \mu_k)^2 \right) - \frac{1}{2} \ln 2\pi. \end{aligned} \quad (17)$$

Estimation of background component: The parameters involved in the background component are U, V and γ . For each row u_i of U , using the factorization (12), we can get

$$q(u_i) = \mathcal{N}(u_i | \mu_{u_i}, \Sigma_{u_i}), \quad (18)$$

with mean μ_{u_i} and covariance Σ_{u_i} given by

$$\begin{aligned} \mu_{u_i}^T &= \Sigma_{u_i} \left\{ \sum_k \langle \tau_k \rangle \sum_j \langle z_{ijk} \rangle (f_{ij} - \langle \mu_k \rangle) \langle v_j \rangle \right\}^T, \\ \Sigma_{u_i} &= \left\{ \sum_k \langle \tau_k \rangle \sum_j \langle z_{ijk} \rangle \langle v_j^T v_j \rangle + \Gamma \right\}^{-1}, \end{aligned}$$

where $\Gamma = \text{diag}(\langle \gamma \rangle)$. Similarly, for each row v_j of V , we have

$$q(v_j) = \mathcal{N}(v_j | \mu_{v_j}, \Sigma_{v_j}), \quad (19)$$

where

$$\begin{aligned} \mu_{v_j}^T &= \Sigma_{v_j} \left\{ \sum_k \langle \tau_k \rangle \sum_i \langle z_{ijk} \rangle (f_{ij} - \langle \mu_k \rangle) \langle u_i \rangle \right\}^T, \\ \Sigma_{v_j} &= \left\{ \sum_k \langle \tau_k \rangle \sum_i \langle z_{ijk} \rangle \langle u_i^T u_i \rangle + \Gamma \right\}^{-1}. \end{aligned}$$

For γ which controls the rank of B , we have

$$q(\gamma_r) = \text{Gam}(\gamma_r | a_r, b_r), \quad (20)$$

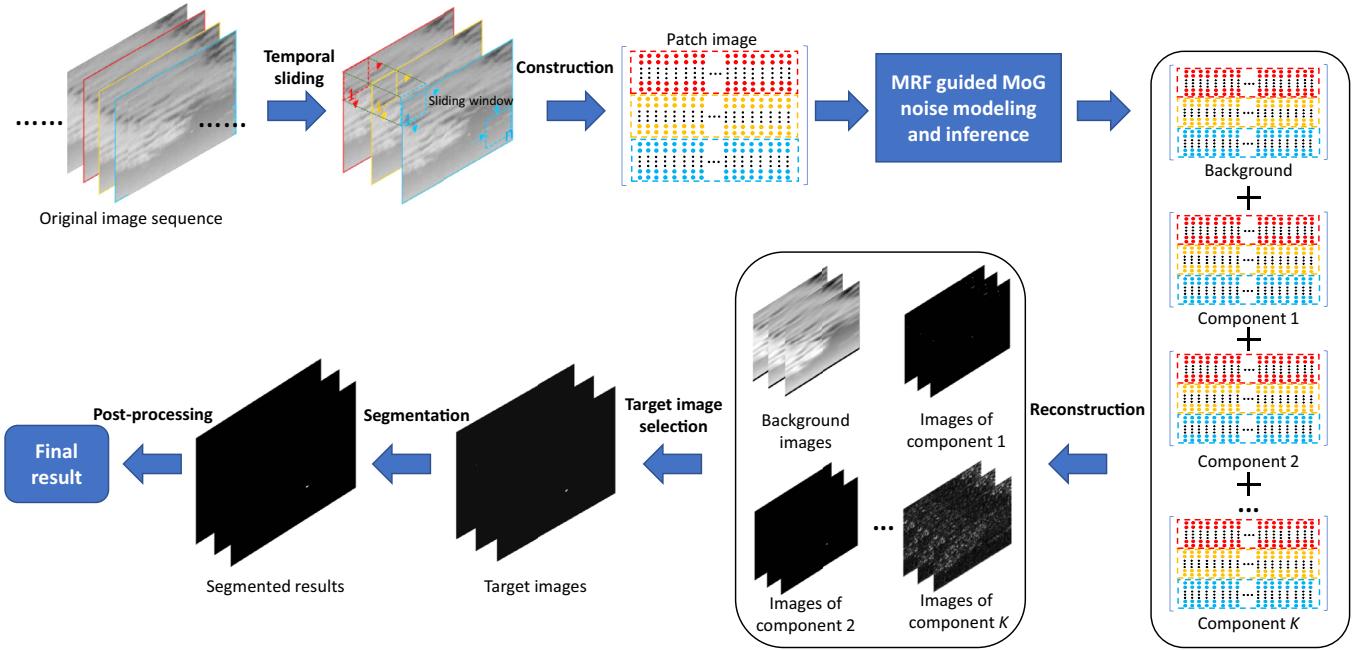


Fig. 5. The framework of the proposed method.

where

$$a_r = a_0 + \frac{m+n}{2}, \quad b_r = b_0 + \frac{1}{2}(\langle u_r^T u_r \rangle + \langle v_r^T v_r \rangle).$$

Then, the optimal U , V and r_{ijk} are obtained after alternatively iteration. Through $E = F - UV^T$, we can finally separate noise component from the background.

In our implementation, we initialize a_0 , b_0 , c_0 , d_0 and $\alpha_{01}, \dots, \alpha_{0K}$ with 10^{-6} , and U and V are initialized through SVD. Besides, a_r , b_r , c_r , d_r , γ_r , τ_k , μ_k , z_{ijk} , π_k are initialized randomly.

2.5. Small target extraction

As r_{ijk} in Eq. (15) presents the probability that the noise components are generated by the k th Gaussian component, the complex noise E can be decomposed into K components $E^1, E^2, \dots, E^m, \dots, E^K$ using the maximum probability criteria as:

$$E_{i,j}^m = \begin{cases} E_{i,j}, & \text{if } m = \arg \max_{k=1,2,\dots,K} (r_{ijk}), \\ 0, & \text{else.} \end{cases} \quad (21)$$

As discussed previously, small targets are contained in one of K components. Once we obtain K components, we need to determine which component contains a small targets. As we discussed previously, the spatial distributions of the K components are different and this is an important cue to select the small target component. In this paper, instead of directly selecting the small target component, we first reconstruct K components into their corresponding image sequences $\bar{E}^1, \bar{E}^2, \dots, \bar{E}^m, \dots, \bar{E}^K$ using the method in Section 2.1. Then, for simplicity, we just calculate their variances and the largest one \bar{E}^i is determined as the small target component as follows:

$$i = \arg \max_{k=1,2,\dots,K} (\text{var}(\bar{E}^k)), \quad (22)$$

where $\text{var}(\cdot)$ means computing variance. Our experimental results show this simple method can work well for different infrared images.

It is inevitable that the reconstructed small target images still contain some residual noise. In order to extract the small target,

a simple segmentation method is applied to each individual image using an adaptive threshold T determined by:

$$T = \max(v_{\min}, \mu + k\sigma), \quad (23)$$

where μ and σ is the mean value and standard deviation of the small target image. k and v_{\min} are constants determined experientially. We set $k = 0.05$ and $v_{\min} = 0.85$ for all test sequences. Here, v_{\min} is to delete false targets with small values.

The framework of our method is shown in Fig. 5 and the detailed steps are summarized as follows:

Step 1: Construct the spatio-temporal patch image F with the given infrared image sequence using the method in Section 2.1.

Step 2: MoG noise model under the Bayesian framework is built by formulating the low rank component B and complex noise E with Eqs. (2) and (5), respectively.

Step 3: VB is utilized to infer the parameters and variables in the proposed model. Then, F is decomposed into B and E based on the results of the inference.

Step 4: Decompose E into K components according to Eq. (21) and reconstruct components into image sequences with the method in Section 2.1.

Step 5: Use Eq. (22) to select the small target images.

Step 6: Segment small target images by Eq. (23) and post-process the segmented result to obtain the detection results.

In Step 6, some post-processing techniques, such as basic region analysis techniques, or morphological techniques, can be used to refine the segmentation results. Generally speaking, the post-processing is still an important step to achieve good detection results in practice. However, for fair comparison with the baseline methods, we do not perform post-processing and directly use the segmentation results to evaluate the performance in our experiments in Section 3.

2.6. Computational complexity

The computational cost of the proposed model consists of three parts: the inference of the MoG noise model, the reconstruction of the spatio-temporal patch image and the segmentation of target.

Assuming that the size of F is $m \times n$, it is clear to see that the cost of variational inference of the parameters of the MoG noise

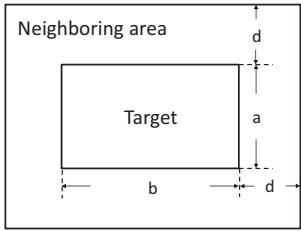


Fig. 6. The neighboring background area of a small target.

model is $O((m+n)R^3 + kmnR)$ per iteration, where $O((m+n)R^3)$ is the computational time of inverting a $R \times R$ matrix in inferring each u_i . and v_j .

The computational cost of the reconstruction is contributed by the median operation. Assuming that w is the overlapping pixel number during the transformation from the spatio-temporal patch image to the reconstruction image, $O(w)$ is needed in the median operation for each pixel. When the size of original image $f(x, y)$ is $g \times h$, the computational cost of the reconstruction is $O(ghw)$.

As to the target segmentation, just a simple comparison operation is implemented and the entire cost of this step is $O(gh)$. In summary, when the iteration times of the inference of the model is N , the whole computational complexity of the proposed model is around $O(ghw + N((m+n)R^3 + kmnR) + gh)$.

3. Experiments

In this section, we firstly introduce the evaluation metrics and baseline methods used in this paper. Secondly, we discuss the key parameters of our method. Finally, we compare the performance of our method with the baseline methods.

3.1. Evaluation metrics and baseline methods

For objective evaluation, the signal to clutter ratio gain (SCRG), background suppression factor (BSF), detection probability P_d and false alarm rate F_a are adopted [24]. SCRG and BSF computed on a single frame can describe the ability of target enhancement, as well as the clutter suppression. They are important indicators to final detection performance. The SCRG is defined as:

$$\text{SCRG} = \frac{\text{SCR}_{\text{out}}}{\text{SCR}_{\text{in}}}, \quad (24)$$

where SCR_{in} and SCR_{out} are respectively the signal to clutter ratio of the original and result images. Here, SCR is defined by:

$$\text{SCR} = \frac{|\mu_t - \mu_b|}{\sigma_b}, \quad (25)$$

where μ_t is the average pixel value of the target, μ_b is the average pixel value of the pixels in neighboring area around the target, and σ_b is the standard deviation of the neighboring area. The size of the neighbour is $(a+2d) \times (b+2d)$, where the size of small target is $a \times b$ as shown as in Fig. 6. We set $d = 15$ in this paper. BSF is defined as follows:

$$\text{BSF} = \frac{C_{\text{in}}}{C_{\text{out}}}, \quad (26)$$

where C_{in} and C_{out} denote the standard deviation of the full image before and after processing, respectively.

In Eqs. (24) and (26), the computation of both metrics includes standard deviation which may be close to zero when the suppressed background is very clean. In this case, these two metrics may approach to infinity. Specifically, this case is prone to happen for SCRG whose standard deviation is calculated based on a local background area. To address this problem, we adopt another metric

to evaluate the ability of target enhancement and clutter suppression, namely contrast gain (CG) which is defined as:

$$\text{CG} = \frac{\text{CON}_{\text{out}}}{\text{CON}_{\text{in}}}, \quad (27)$$

where CON_{in} and CON_{out} are the contrast (CON) of the original and result images, respectively, and CON is defined as:

$$\text{CON} = |\mu_t - \mu_b|, \quad (28)$$

where μ_t and μ_b are the same as those in Eq. (25).

Detection probability P_d and false alarm rate F_a are used to directly evaluate the detection performance and defined as:

$$P_d = \frac{N_a}{N_b}, F_a = \frac{N_f}{N_l}, \quad (29)$$

where N_a is the number of true detections, N_b is the number of actual targets, N_f is the number of false detections and N_l is the number of sequence frames.

The baseline methods chosen in this paper include the IPI model [1], Top-hat [34], MaxMean and MaxMedian filtering [19]. Since the MRF constraint is one of our contributions in our method, we also use the version of our method without MRF constraint as a baseline method, denoted as MoG.

We use five consecutive real image sequences¹ to test the proposed methods. The sequences were recorded with static infrared camera. The five sequences consist of 108, 123, 134, 100, 100 frames, respectively. The representative frames of each sequence are shown in the first column of Fig. 14.

3.2. Effect of parameters

3.2.1. Number of frames

In order to utilize the temporal information, the cubic patch is used. Thus, there are two questions: (1) Is the temporal information helpful? (2) How many frames should be used in a cubic patch? To answer these questions, we test cases of different numbers of frames, including the single frame case, on two real image sequences of sky and sea-sky backgrounds, and the corresponding receiver operating characteristic (ROC) curves are drawn in Fig. 7. Here, the noise component parameter K in the MoG noise model is set as $K = 3$. From Fig. 7, we can observe that the cases of multiple frames are obviously better than the single frame case. This means that the temporal information is really helpful for performance improvement. For the cases of multiple frames, when low false alarm rate, the performance seems to increase with the temporal length, especially when false alarm rate is less than around 0.15. However, when the false alarm rate is more than 0.2, this trend is not obvious. Thus, considering the fact that the complexity of the algorithm will increase greatly with the temporal length, in this paper we just use three frames for cubic patches. Actually, a long temporal length may decrease the correlation between cubic patches.

3.2.2. Number of components

The noise component parameter K in the MoG noise model is an important parameter. On the one hand, the target component may contain noise if K is too small. On the other hand, there may be redundant components if K is too large. This can be easily observed through Figs. 8–10 which correspond to the cases of $K = 2, K = 3, K = 4$, respectively. If we assume two noise components ($K = 2$), the sparse noise may be mixed with the small target component, as shown as “Component1” in Fig. 8. On the contrary,

¹ Most of our sequences are from the provider of research projects. We are waiting for their permission for publicly sharing the images. Once we get the authorization, we will share them on our research homepage for research purpose.

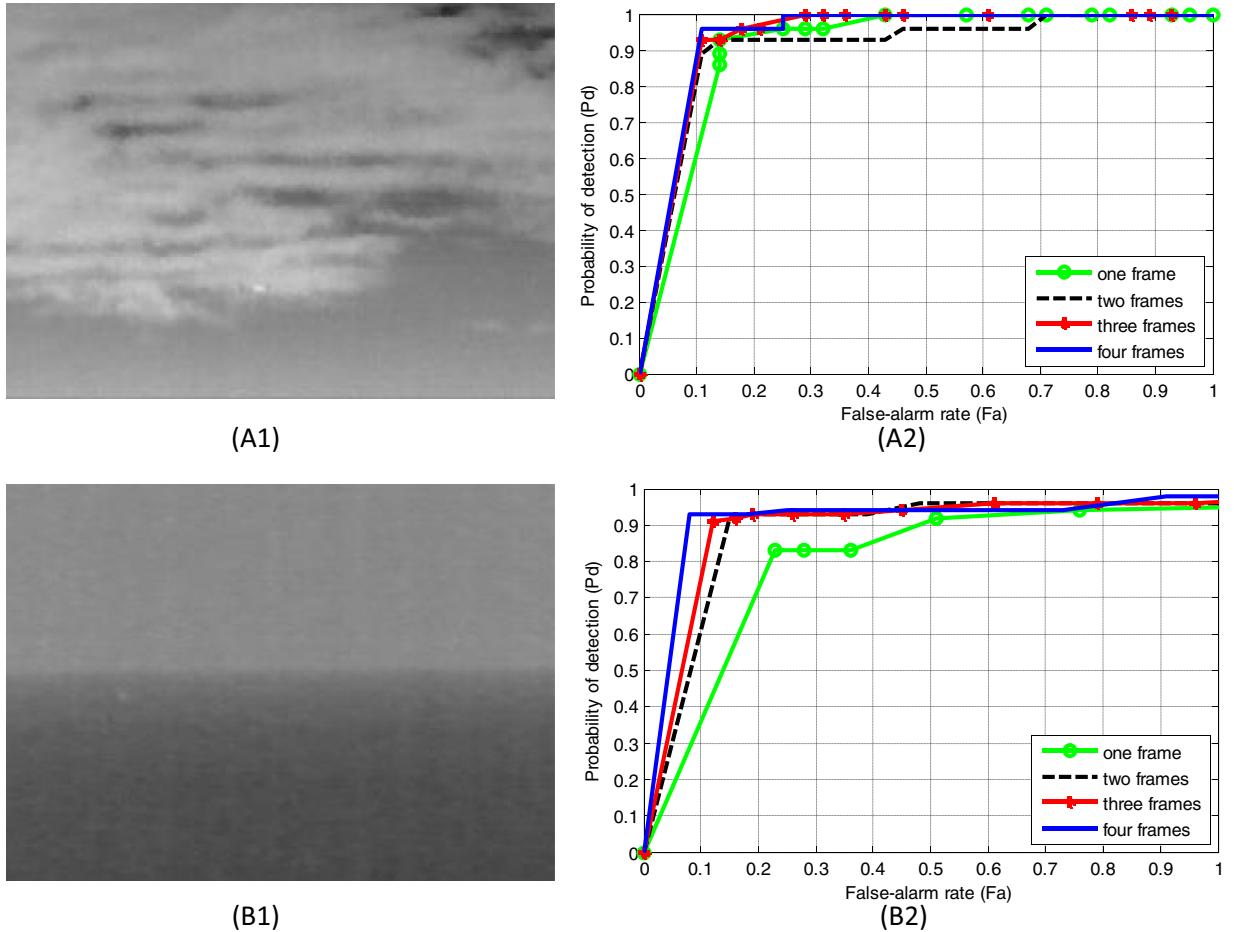


Fig. 7. ROC curves of different temporal lengths on two image sequences. (A1) and (B1) are the representative frames of the two image sequences. (A2) and (B2) are the corresponding ROC curves of the proposed method.

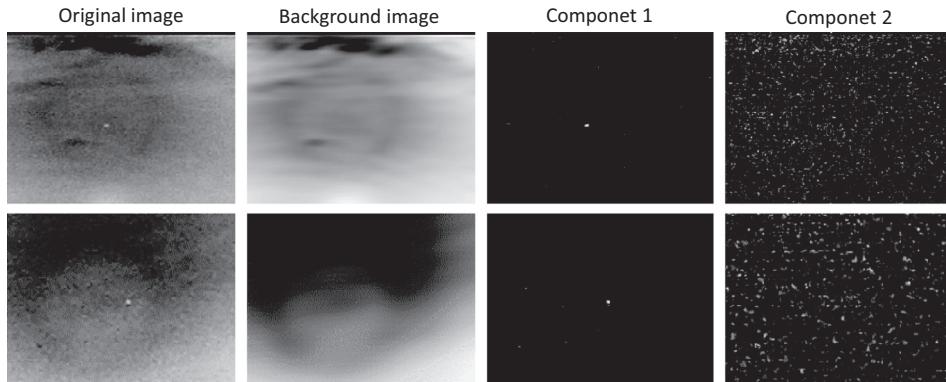


Fig. 8. The results of our method with two Gaussian components.

if we assume four noise components ($K = 4$), there may be a redundant component, as shown as ‘‘Component4’’ in Fig. 10. When $K = 3$, the divided components exactly correspond to small target, sparse noise and dense noise components respectively, as shown as ‘‘Component1’’, ‘‘Component2’’ and ‘‘Component3’’ in Fig. 9. To further verify the above discussion, we draw the ROC curves of different values for K on five real image sequences, as shown in Fig. 11. It can be obviously observed that the performance is best when $K = 3$. Specifically, for Sequence 2 to 5, the configuration of $K = 3$ has considerably higher probabilities of detection (P_d) than other values. It is worth noting that in Fig. 11 (a) the ROC curve with $K = 4$ is better than that with $K = 3$ when $F_a < 0.5$. This is

because the separated target components with $K=4$ have cleaner backgrounds than those with $K=3$ for most of images. However, because the separated target components of some images lose the real small targets when $K = 4$, the P_d of the ROC curve can not reach 1. But this does not happen for $K = 3$. In following experiments, we always set $K = 3$.

3.3. Performance evaluation

In this section, we first evaluate the effect of the MRF constraint in the proposed model and then subjectively and objectively evaluate the performance of small target enhancement of the pro-

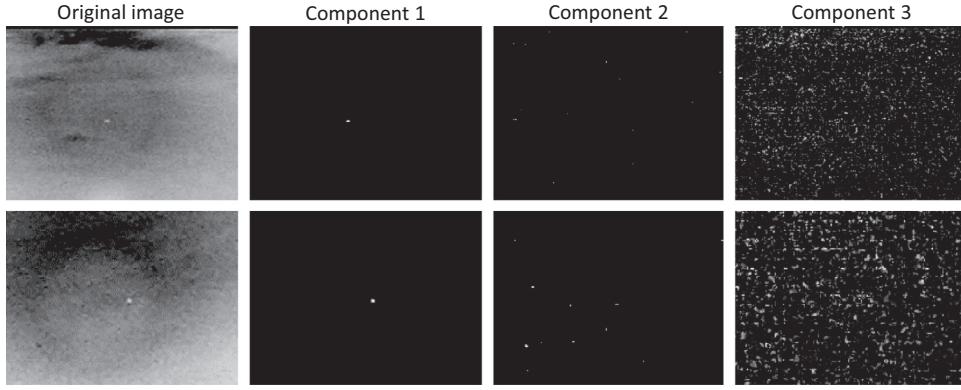


Fig. 9. The results of our method with three Gaussian components.

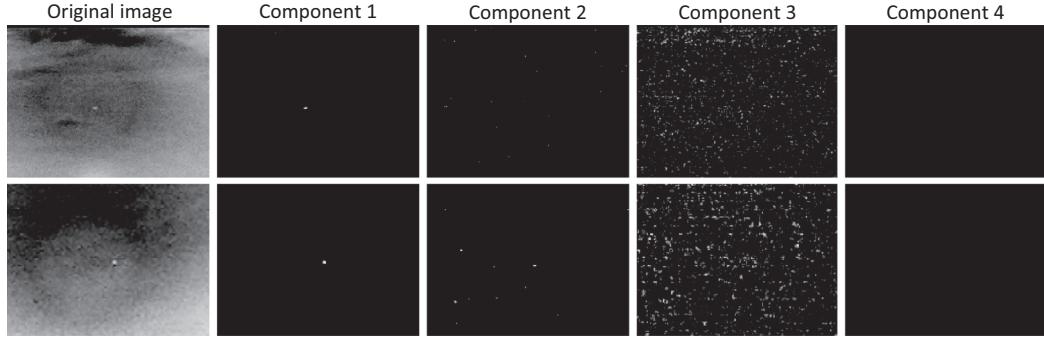


Fig. 10. The results of our method with four Gaussian components.

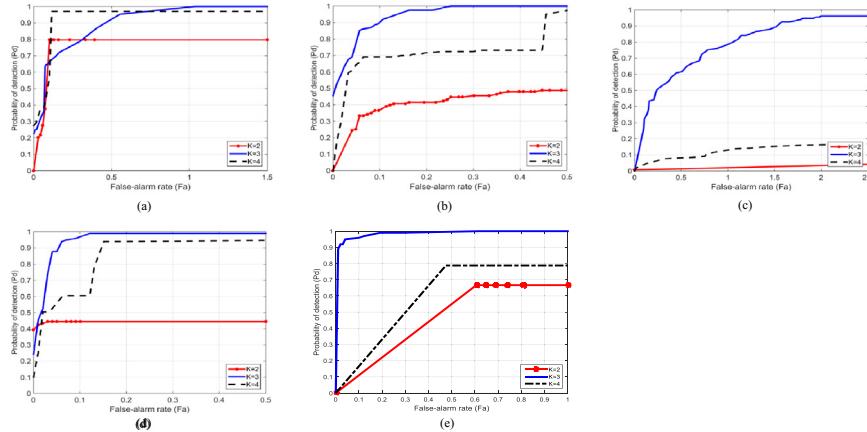


Fig. 11. The ROC curves of different values for the parameter K on the five real image sequences. (a) Result of Sequence 1, (b) result of Sequence 2, (c) result of Sequence 3, (d) result of Sequence 4, (e) result of Sequence 5.

posed method in comparison with baseline methods. Finally, we use five real infrared image sequences to evaluate the detection performance.

3.3.1. The effect of MRF

According to the characterises of infrared small targets, we adopt MRF to guide the MoG noise modeling. To verify the effectiveness of the MRF, we also test the version of the proposed method without MRF constraint, namely MoG. The results with/without MRFs for the same images are shown in Figs. 12 and 13, respectively. We can obviously observe that the MRF constraint can effectively help clean the background clutter in the small target image, as shown as "Component1" in Fig. 12. In contrast, without MRF constraint the small target image from original MoG noise model [29] has obvious noise residual, as shown as "Component1"

in Fig. 13. These noise residual could cause a high false alarm rate in the following detection procedure. Our method with MRF constraint can effectively avoid this problem.

3.3.2. Comparisons to baseline methods

We compare the abilities of background suppression and target detection of our method and baseline methods. For both *IPI model* and our method, we set the patch size as 50×50 , and the step size as 10. And the filter sizes of methods *TopHat*, *MaxMean*, and *MaxMedian* are set as 9×9 which is the usual configuration [3,24].

Fig. 14 shows results of different methods without segmentation on five real infrared images with heavy background clutters. Here, these original images are from five infrared image sequences, respectively. We can see that both the *IPI model* and our method have cleaner backgrounds than other four baseline methods, and

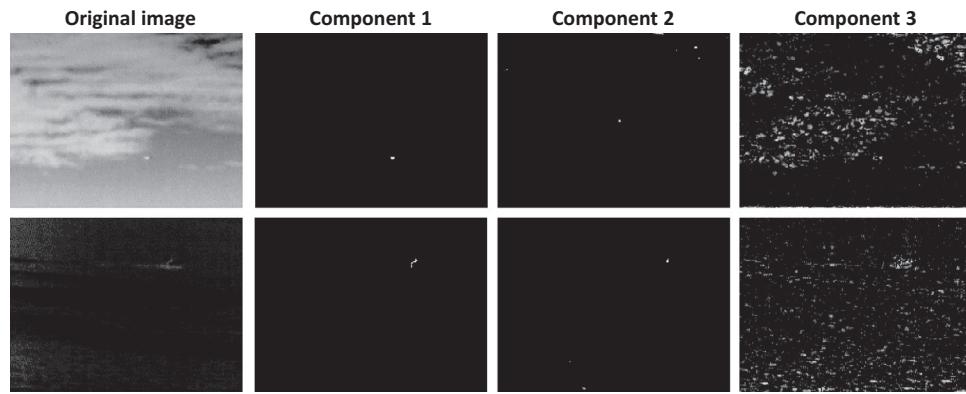


Fig. 12. The results of the proposed method (with MRF constraint).

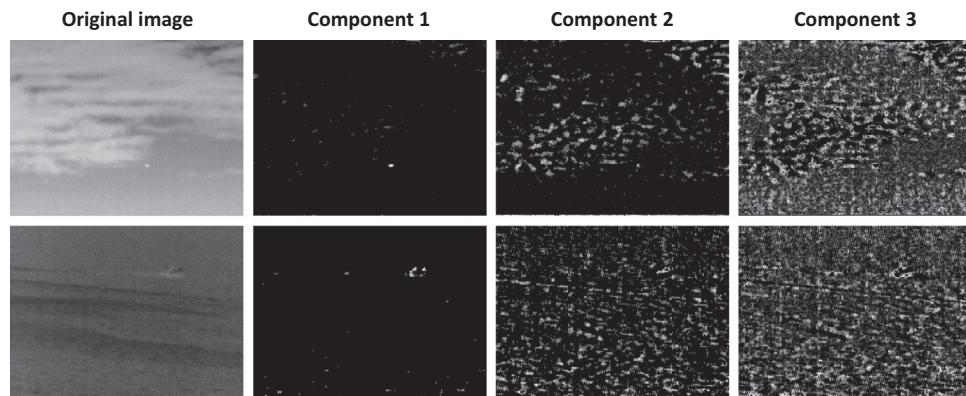


Fig. 13. The results of MoG (without MRF constraint).

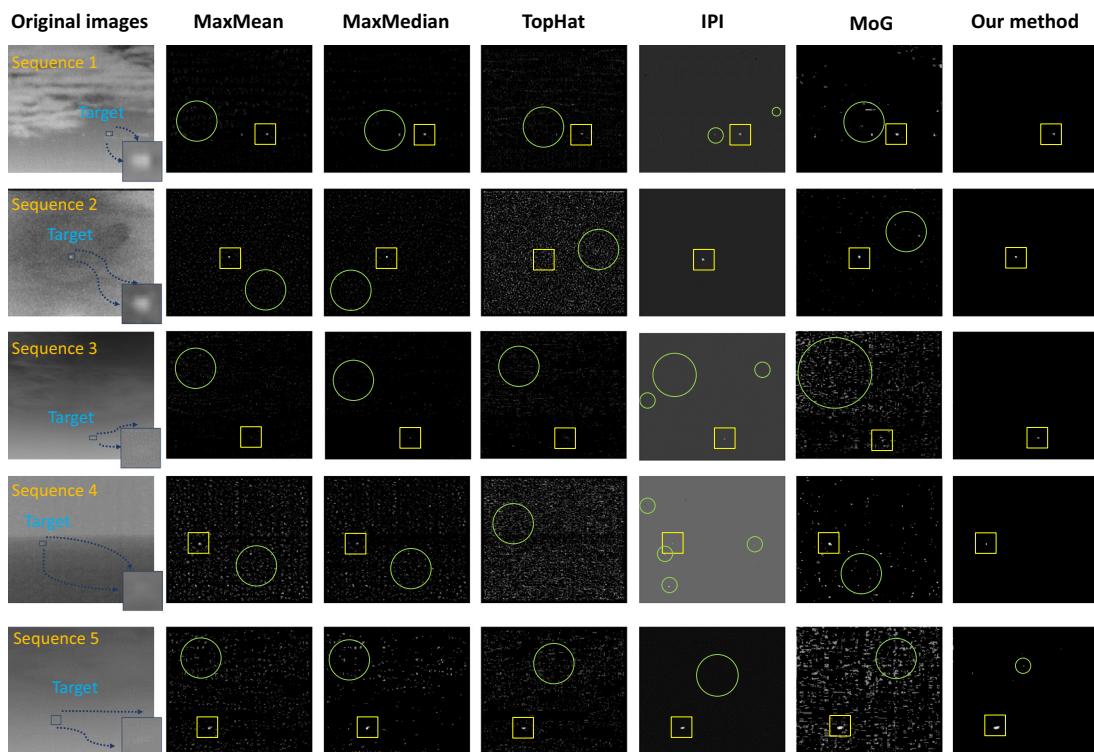


Fig. 14. The representative results of different methods on five real infrared image sequences. Images in the first column are the representative frame of five real infrared image sequences. The yellow rectangles denote the targets and the green circles are representative examples of noise. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1

The evaluation results of SCRG and BSF of different methods for images in the first column in Fig. 14.

Methods	image of Seq. 1		image of Seq. 2		image of Seq. 3		image of Seq. 4		image of Seq. 5	
	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF	SCRG	BSF
Max-mean	3.44	7.07	4.68	1.91	6.38	2.92	1.98	1.17	2.52	1.17
Max-median	2.17	9.83	5.94	2.06	3.76	5.55	1.73	1.51	2.76	1.4
Top-hat	2.27	5.96	3.04	0.75	7.72	2.48	0.71	0.87	1.91	1.21
IPI	7.31	8.03	15.38	5.88	10.12	27.86	8.96	12.37	3.81	6.98
MoG	6.21	9.61	9.58	2.95	2.83	0.32	5.07	1.69	3.44	0.52
Ours	19.41	26.79	31.65	9.76	14.95	18.76	17.05	20.88	6.6	3.55

Table 2

The evaluation results of average CG values of different methods for all image sequences.

Method	Seq. 1	Seq. 2	Seq. 3	Seq. 4	Seq. 5
Max-mean	1.68	2.97	22.25	18.48	14.35
Max-median	1.42	2.98	15.49	15.66	9.31
Top-hat	0.97	1.46	17.63	9.34	11.92
IPI	1.26	2.24	4.66	8.83	12.14
MoG	3.37	3.81	20.8	20.94	20.14
Ours	2.31	4.28	35.2	24.11	29.05

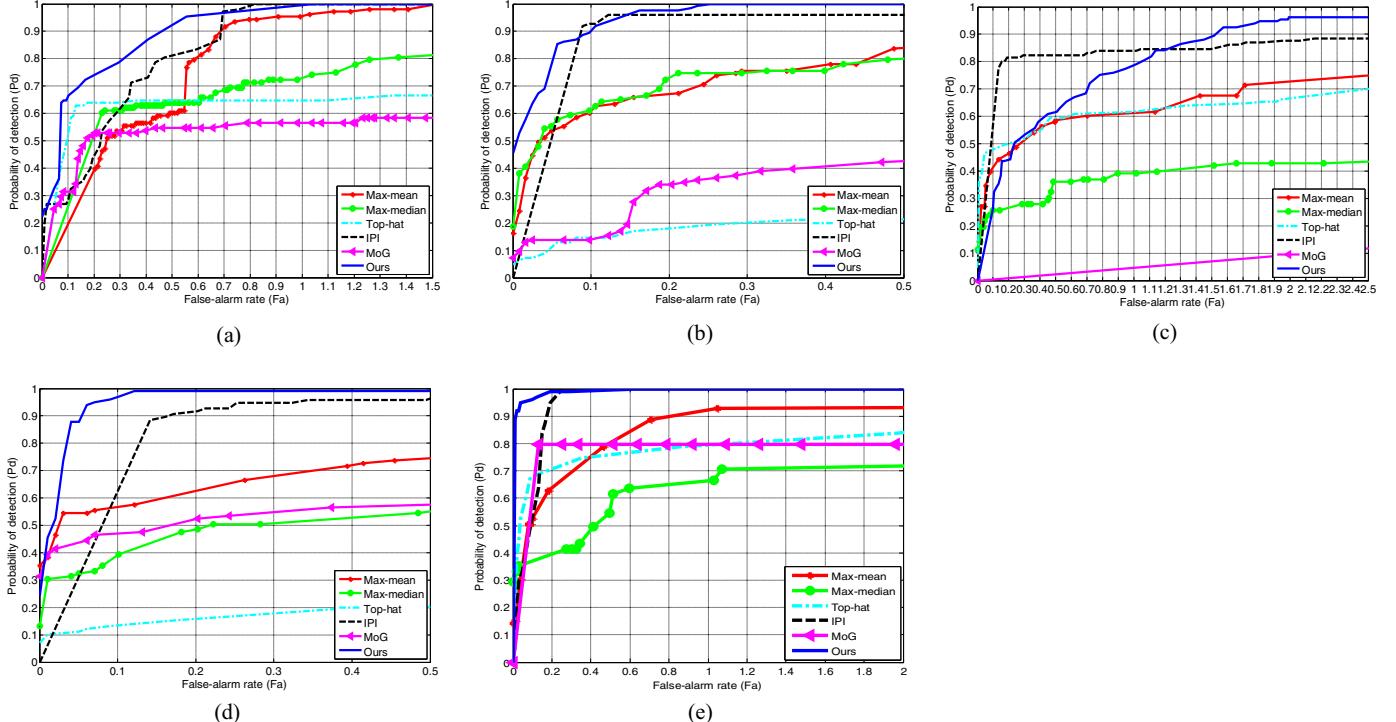


Fig. 15. The ROC curves of six methods for five real image sequences. (a) Result of Sequence 1. (b) Result of Sequence 2. (c) Result of Sequence 3. (d) Result of Sequence 4. (e) Result of Sequence 5.

TopHat, *MaxMean*, *MaxMedian* and *MoG* have relatively dense noise residual. Comparing our method to *IPI model*, we can observe that *IPI model* still has some sparse noise residual while our method has better results. To objectively evaluate the performance of background suppression of our method, the SCRG and BSF are computed based on Fig. 14 and the results of different methods are listed in Table 1. We can obviously see that the proposed method can get higher performance than baseline methods for both two metrics. *Tophat* has the smallest values for its failure in suppressing background clutters and reserving the target information. Since *MaxMean* and *MaxMedian* can retain parts of the target information, it is a little higher than *TopHat*, but obviously lower than both the *IPI model* and the proposed approach for two metrics. the *IPI model* obtains better BSF performance in images of Seq. 3 and Seq.

5. This is because the background intensity of the results is higher in the *IPI model*, making the standard deviation smaller. *MoG* has unbalanced performance in different images, for its instability. This further means that the MRF constraint is very important for small target noise modeling.

As discussed in Section 3.1, the case that standard deviation may be close to zero prevents us from calculating the average values of above used metrics for all images. Instead, we calculate average contrast gain of all images for each sequence and the results are listed in Table 2. It is obviously observed that the proposed method totally has the best results among all baseline methods. According to the definition of CG in Eq. (28), the CG is the gray value difference of background and small targets. The results of Table 2 means that our method has good performance on enlarg-

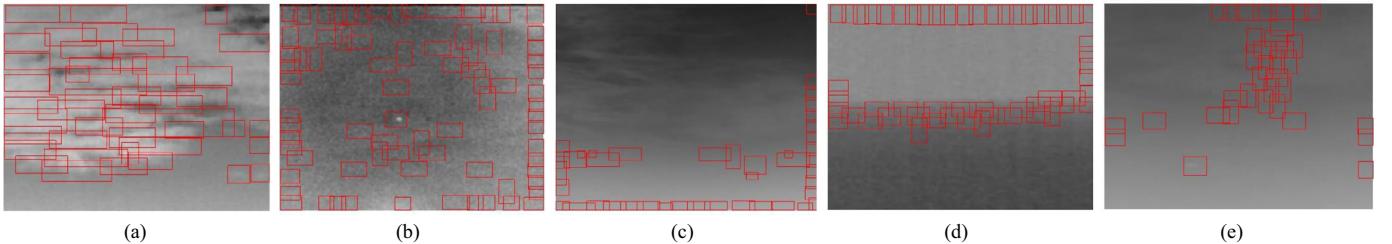


Fig. 16. The representative detection results of the faster-RCNN method [35] on five real infrared images. (a) The result from the sequence 1. (b) The result from the sequence 2. (c) The result from the sequence 3. (d) The result from the sequence 4. (e) The result from the sequence 5.

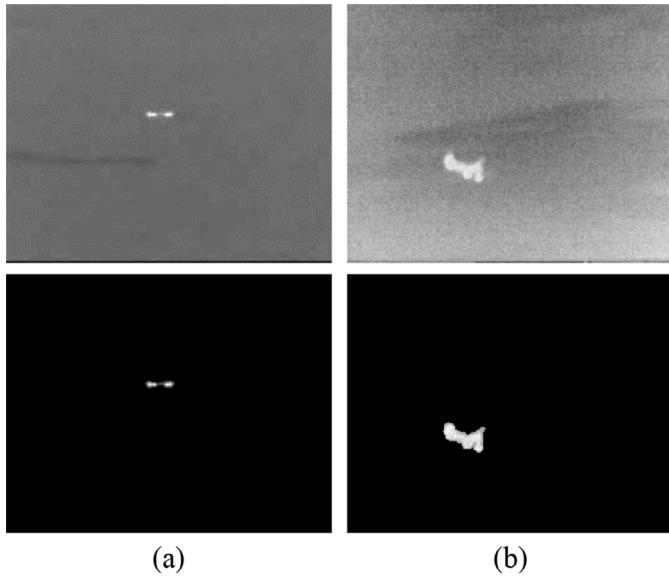


Fig. 17. The representative detection results of the proposed method on real infrared images with different target sizes. (a) The original image and result with a moderate target size. (b) The original image and result with a little large target size.

ing this gray value difference. This will greatly help the following threshold segmentation task for extracting small targets.

Fig. 15 shows the ROC curves of the methods for five image sequences, respectively. It can be seen that the proposed method has better detection performance than baseline methods. For Sequence 1, 2, 4 and 5, our method has higher probabilities of detection (P_d) and relatively lower false alarm rates (F_a). For Sequence 3, the *IPI model* has a little higher performance than our method when $F_a \leq 1.2$. However, our method reaches around 0.97 (97%) in a faster speed than the *IPI model* when $F_a > 1.2$. **Fig. 15** also shows that the *MoG* without MRF constraint has relatively poor detection performance.

Currently, modern detection algorithms such as Convolutional Neural Networks show astonishing results in visual object detection. For comparison, we test the state-of-the-art object detection method, namely the powerful faster-RCNN method [35]. We resize the anchor size to 16^2 , 32^2 and 64^2 pixels to adapt to the scales of infrared small targets. We randomly split our dataset to 1.5:1 for training set and test set, and use the ZF model for training. The representative test results are shown in **Fig. 16**. It can be observed that there are a number of false detections. The reason could be that the small targets are too small, and thus the feature of small targets is not obvious, with heavy background clutter. This characteristic of small targets makes the powerful ability of CNN weak on feature learning.

As discussed above, the proposed algorithm mainly handles the case of small-dim targets, similar to noise. However, it is valid for a certain range of small target size. In infrared small target applications, a common problem is that small target sizes in the image could become large during approach, e.g., in missile systems.

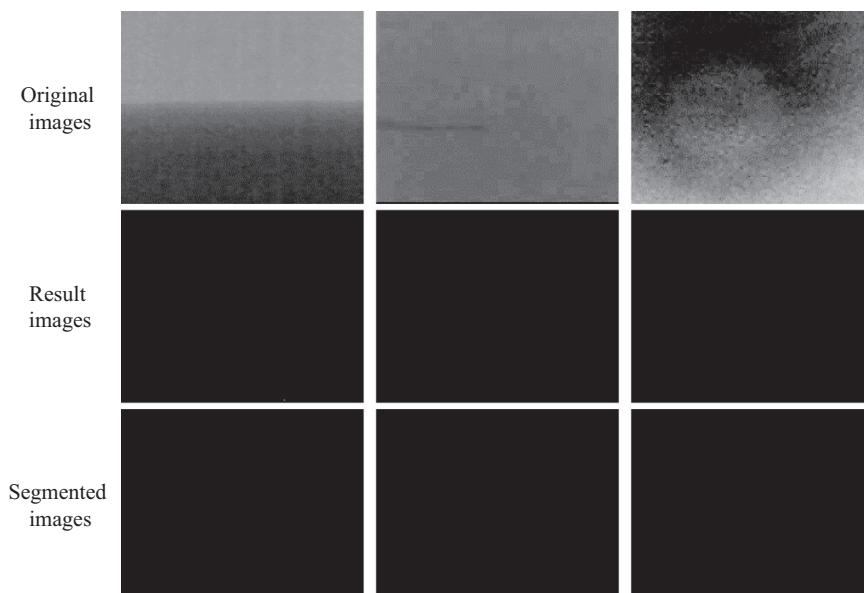


Fig. 18. The detection results of the proposed method on different background images without any target.

Fig. 17 shows the results of this situation. It can be seen that even for infrared targets with a little large size, the proposed method still can get good results. Generally, the small target detection algorithm is applied to the early-warning stage. In this stage, the interested targets are just of a small size. Following this stage is a tracking stage. During the tracking stage, we can switch an appropriate the detection algorithm if we find the size of the target becomes quite large.

Furthermore, to evaluate the performance on the case of no target, we test our method on different background images without targets. As can be seen in **Fig. 18**, the test results are clean without false alarms. Thus, the proposed method is also robust in infrared images without any target.

For utilizing temporal information, the proposed method has a delay of several frames. For the case of three consecutive images for the patch image construction, there is a delay of two frames. If the frame rate of one infrared sensor is 30 frames per second, there is just a delay of around 0.067s, assuming that the proposed method can real-timely run on an extremely optimal configuration of the software and hardware. If this delay can not meet the requirement, we can reduce the number of consecutive images for the patch image construction. Especially, our method supports only one image for the patch image construction without temporal information. As discussed in **Section 3.2.1**, our method can still achieve good performance even just using one frame.

Averagely, for one frame with the size of 256×200 , the real running time of the proposed method is around 93 seconds with Matlab implementation, running on a PC with I7 CPU, 64GB of memory. Actually, the real running time depends on many factors, including the iteration times of VB, the number of frames for the patch image, the patch size and so on. For practical applications, these parameters can be further optimized to reach a good trade-off between the performance and running time.

4. Conclusion

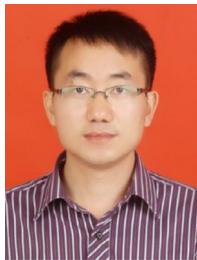
A novel method based on MRF guided MoG noise model is presented in this paper, to address the small-dim target detection problem with heavy background noise. Spatio-temporal patch images are constructed using local cubic patches. And then the background, small target and noise components are formulated into a unified model under the Bayesian framework. After that, variational Bayesian is adopted to solve the proposed model and the experiments show that the very small and dim targets can be effectively separated from complex background noise by our methods.

Acknowledgment

This work is supported by the National Natural Science Foundation of China (No. 61571071, 61373114, 61661166011, 61603292, 11690011, 61721002) and the National Grand Fundamental Research 973 Program of China under Grant No. 2013CB329404.

References

- [1] C. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, A.G. Hauptmann, Infrared patch-image model for small target detection in a single image, *Image Process. IEEE Trans.* 22 (12) (2013) 4996–5009.
- [2] Y. He, M. Li, J. Zhang, Q. An, Small infrared target detection based on low-rank and sparse representation, *Infrared Phys. Technol.* 68 (2015) 98–109.
- [3] C. Gao, J. Tian, P. Wang, Generalised-structure-tensor-based infrared small target detection, *Electron. Lett.* 44 (23) (2008) 1349–1351.
- [4] Y. Wei, X. You, H. Li, Multiscale patch-based contrast measure for small infrared target detection, *Pattern Recognit.* 58 (2016) 216–226.
- [5] Y. Li, S. Liang, B. Bai, D. Feng, Detecting and tracking dim small targets in infrared image sequences under complex backgrounds, *Multimed. Tools Appl.* 71 (3) (2014) 1179–1199.
- [6] C. Huang, Y. Li, R. Nevatia, Multiple target tracking by learning-based hierarchical association of detection responses, *Pattern Anal. Mach. Intell. IEEE Trans.* 35 (4) (2013) 898–910.
- [7] S. Kim, J. Lee, Scale invariant small target detection by optimizing signal-to-clutter ratio in heterogeneous background for infrared search and track, *Pattern Recognit.* 45 (1) (2012) 393–406.
- [8] H. Chen, Y. Bar-Shalom, K.R. Pattipati, T. Kirubarajan, Mdl approach for multiple low observable track initiation, *Aerospace Electron. Syst. IEEE Trans.* 39 (3) (2003) 862–882.
- [9] E. Grossi, M. Lops, Sequential along-track integration for early detection of moving targets, *Signal Process. IEEE Trans.* 56 (8) (2008) 3969–3982.
- [10] E. Grossi, M. Lops, L. Venturino, A novel dynamic programming algorithm for track-before-detect in radar systems, *Signal Process. IEEE Trans.* 61 (10) (2013) 2608–2619.
- [11] K.A. Melendez, J.W. Modestino, Spatiotemporal multiscan adaptive matched filtering, in: SPIE's 1995 International Symposium on Optical Science, Engineering, and Instrumentation, International Society for Optics and Photonics, 1995, pp. 51–65.
- [12] M. Li, T. Zhang, W. Yang, X. Sun, Moving weak point target detection and estimation with three-dimensional double directional filter in ir cluttered background, *Opt. Eng.* 44 (10) (2005). 107007–107007.
- [13] D. Liu, J. Zhang, W. Dong, Temporal profile based small moving target detection algorithm in infrared image sequences, *Int. J. Infrared. Millimeter Waves* 28 (5) (2007) 373–381.
- [14] T. Zhang, M. Li, Z. Zuo, W. Yang, X. Sun, Moving dim point target detection with three-dimensional wide-to-exact search directional filtering, *Pattern Recognit. Lett.* 28 (2) (2007) 246–253.
- [15] T.-W. Bae, Small target detection using 3-dimensional bilateral filter, *J. Korea Multimed. Soc.* 16 (6) (2013) 746–755.
- [16] X. Liu, Z. Zuo, A Dim Small Infrared Moving Target Detection Algorithm Based on Improved Three-dimensional Directional Filtering, in: Advances in Image and Graphics Technologies, Springer, 2013, pp. 102–108.
- [17] L. Deng, H. Zhu, C. Tao, Y. Wei, Infrared moving point target detection based on spatial-temporal local contrast filter, *Infrared Phys. Technol.* 76 (2016) 168–173.
- [18] J. Zhou, H. Lv, F. Zhou, Infrared small target enhancement by using sequential top-hat filters, *International Symposium on Optoelectronic Technology and Application 2014*, International Society for Optics and Photonics, 2014. 93011L–93011L.
- [19] S.D. Deshpande, H.E. Meng, R. Venkateswarlu, P. Chan, Max-mean and max-median filters for detection of small targets, in: SPIE's International Symposium on Optical Science, Engineering, and Instrumentation, International Society for Optics and Photonics, 1999, pp. 74–83.
- [20] Y. Zhao, H. Pan, C. Du, Y. Peng, Y. Zheng, Bilateral two-dimensional least mean square filter for infrared small target detection, *Infrared Phys. Technol.* 65 (2014) 17–23.
- [21] J. Cao, C. Gao, Y. Xiao, P. Li, M. Cai, Infrared small target detection based on the tensor model, in: *Image and Signal Processing (CISP), 2014 7th International Congress on, IEEE*, 2014, pp. 169–173.
- [22] Z. Chen, S. Luo, T. Xie, J. Liu, G. Wang, G. Lei, A novel infrared small target detection method based on bemd and local inverse entropy, *Infrared Phys. Technol.* 66 (2014) 114–124.
- [23] T.-W. Bae, F. Zhang, I.-S. Kweon, Edge directional 2d lms filter for infrared small target detection, *Infrared Phys. Technol.* 55 (1) (2012) 137–145.
- [24] C. Gao, T. Zhang, Q. Li, Small infrared target detection using sparse ring representation, *Aeros. Electron. Syst. Mag. IEEE* 27 (3) (2012) 21–30.
- [25] J.-H. Kim, J.-J. Park, S.-H. Ahn, D.G. Lee, D. Moon, S.-K. Kim, Small target detection using morphology and modified gaussian distance function, *Secur. Commun. Netw.* 9 (6) (2016) 555–560.
- [26] J. Wright, A. Ganesh, S. Rao, Y. Peng, Y. Ma, Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization, in: *Advances in Neural Information Processing Systems*, 2009, pp. 2080–2088.
- [27] X. Ding, L. He, L. Carin, Bayesian robust principal component analysis, *Image Process. IEEE Trans.* 20 (12) (2011) 3419–3430.
- [28] S.D. Babacan, M. Luessi, R. Molina, A.K. Katsaggelos, Sparse bayesian methods for low-rank matrix estimation, *Signal Process. IEEE Trans.* 60 (8) (2012) 3964–3977.
- [29] Q. Zhao, D. Meng, Z. Xu, W. Zuo, L. Zhang, Robust principal component analysis with complex noise, in: *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, 2014, pp. 55–63.
- [30] Z. Lin, A. Ganesh, J. Wright, L. Wu, M. Chen, Y. Ma, Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix, *Comput. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*, 2009, 61 (6) (2009).
- [31] C.M. Bishop, et al., *Pattern Recognition and Machine Learning*, 4, Springer New York, 2006.
- [32] D. Meng, F. Torre, Robust matrix factorization with unknown noise, in: *Proceedings of the IEEE International Conference on Computer Vision*, 2013, pp. 1337–1344.
- [33] X. Cao, Q. Zhao, D. Meng, Y. Chen, Z. Xu, Robust low-rank matrix factorization under general mixture noise distributions, *IEEE Trans. Image Process.* 25 (10) (2016) 4677–4690, doi:10.1109/TIP.2016.2593343.
- [34] V. Tom, T. Peli, M. Leung, J. Bondaryk, Morphology-based algorithm for point target detection in infrared backgrounds, in: SPIE, 1994, 1993, pp. 25–32.
- [35] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, in: *Advances in Neural Information Processing Systems*, 2015, pp. 91–99.



Chenqiang Gao received the B.S. degree in computer science from China University of Geosciences, Wuhan, China, in 2004 and the Ph.D. degree in pattern recognition and intelligence systems from the Huazhong University of Science and Technology, Wuhan, China, in 2009. In August 2009, he joined School of Communications and Information Engineering at Chongqing University of Posts and Telecommunications (CQUPT), Chongqing, China. In September 2012, he joined the Informedia Group in School of Computer Science at Carnegie Mellon University (CMU), working on Multimedia Event Detection (MED) and Surveillance Event Detection (SED) as a visiting scholar. In April 2013, he became a Post Doctoral Fellow and continued work on MED and SED until March 2014 when he returned to CQUPT. He is currently a Professor at CQUPT. His research interests include image processing, infrared target detection, action recognition and event detection.



Lan Wang is currently studying for the M.S degree in information and telecommunication engineering from Chongqing Key Laboratory of Signal and Information Processing, Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include infrared target detection, infrared video based action and event detection.



Yongxing Xiao is currently studying for the M.S degree in information and telecommunication engineering from Chongqing Key Laboratory of Signal and Information Processing, Chongqing University of Posts and Telecommunications, Chongqing, China. Her research interests include computer vision and action recognition.



Qian Zhao received the B.Sc., Ph.D degrees in 2009 and 2015, respectively, from Xi'an Jiaotong University, Xi'an, China. From November 2013 to November 2014, he was visiting at Human Sensing Lab and Component Analysis Lab of The Robotics Institute, Carnegie Mellon University. He is currently a lecturer in School of Mathematics and Statistics, Xi'an Jiaotong University. His current research interests include Low-rank matrix/tensor analysis, Bayesian modeling, and Self-pace learning.



Deyu Meng received the B.Sc., M.Sc., and Ph.D degrees in 2001, 2004, and 2008, respectively, from Xi'an Jiaotong University, Xi'an, China. He is currently an associate professor with the institute for information and system sciences, faculty of science, Xi'an Jiaotong University. His current research interests include principal component analysis, nonlinear dimensionality reduction, feature extraction and selection, compressed sensing, and sparse machine learning methods.