






Infrared Dim and Small Target Detection Based on Greedy Bilateral Factorization in Image Sequences

Dongdong Pang , Tao Shan , *Member, IEEE*, Wei Li , *Senior Member, IEEE*, Pengge Ma, Shengheng Liu , *Member, IEEE*, and Ran Tao , *Senior Member, IEEE*

Abstract—Fast and stable detection of dim and small infrared (IR) targets in complex backgrounds has important practical significance for IR search and tracking system. The existing small IR target detection methods usually fail or cause a high probability of false alarm in the highly heterogeneous and complex backgrounds. Continuous motion of a target relative to the background is important information regarding detection. In this article, a low-rank and sparse decomposition method based on greedy bilateral factorization is proposed for IR dim and small target detection. First, by analyzing the complex structure information of IR image sequences, the target is regarded as an independent sparse motion structure and an efficient optimization algorithm is designed. Second, the greedy bilateral factorization strategy is adopted to approximate the low-rank part of the algorithm, which significantly accelerates the efficiency of the algorithm. Extensive experiments demonstrate that the proposed method has better detection performance than the existing methods. The proposed method can still detect targets quickly and stably especially in complex scenes with weak signal-to-noise ratio.

Index Terms—Greedy bilateral factorization, image sequences, infrared (IR) dim and small targets detection, low-rank and sparse decomposition (LSD).

I. INTRODUCTION

INFRARED search and tracking system is a passive detection system and is widely used in video target monitoring and other fields. Compared with radar and visible light systems, infrared (IR) system has the advantages of strong antijamming ability and good concealment. Because the target is far from the imaging sensor, the target usually appears as a point target with several pixels, resulting in the lack of fine structural features such as texture and shape. Moreover, the background of target imaging is complex, including cloud clutter, building and strong

sunlight clutter, which leads to the very low signal-to-noise ratio (SNR) of target and brings great challenges to target detection. Fig. 1 shows four representative IR small target images.

Many methods have been proposed to detect dim and small IR targets. Traditional detection methods mainly include two categories: single-frame detection methods and sequential detection methods. Typical single-frame detection methods include top-hat filtering [1], max-mean/max-median filtering [2], 2-D least mean square adaptive filtering [3], etc. Most of these methods are based on morphological filtering to detect targets by suppressing the background. These methods show good detection performance in a simple and uniform background. Then, based on human visual attention mechanism, many IR small target detection methods are proposed, such as local contrast measure (LCM). LCM proposes a local contrast descriptor to measure the difference between the current position and its neighborhood to enhance the target while suppressing the background, so as to improve the image SNR [4]. On this basis, many improved LCM (ILCM) methods are proposed [5], such as multiscale relative LCM [6], high-boost-based multiscale LCM [7], multiscale patch-based contrast measure [8], and Gaussian scale-space enhanced LCM [9]. From the perspective of local image segmentation, an effective IR small target detection algorithm inspired by random walks (RW) is presented combined with local contrast characteristics and global uniqueness of small targets [10]. Qin *et al.* proposed a local contrast descriptor based on a facet kernel filtering and RWs (FKRW) algorithm for clutter suppression and target enhancement [11]. Huang *et al.* proposed an IR small target detection method based on density peak search and maximum gray area growth, where the target and background were segmented by selecting seed growth points [12]. These methods achieve better detection performance in complex scenes. Aiming at the chaotic cloudy/sea-sky background, Deng *et al.* proposed a multiscale fuzzy metric method [13] and a weighted local difference measurement method [14]. Based on the derivative information, the derivative entropy-based contrast measure method [15] and a derivative dissimilarity measure method [16] are proposed for small target detection. In addition, single-frame-based small target detection methods include multiscale gray difference weighted image entropy [17], multiscale gray and variance difference measurement [18], entropy-based window selection [19], etc.

Compared with single-frame detection methods, sequential detection methods can get better detection performance in low SNR. This is due to the fact that sequential detection methods

Manuscript received March 29, 2020; revised May 20, 2020; accepted May 26, 2020. Date of publication June 1, 2020; date of current version June 29, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant U1833203, Grant 61671060, and Grant 61421001, and in part by the Key Scientific and Technological Projects in Henan Province (192102210106). (*Corresponding author: Tao Shan.*)

Dongdong Pang, Tao Shan, Wei Li, and Ran Tao are with the Beijing Key Laboratory of Fractional Signals and Systems School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: 3120185397@bit.edu.cn; shantao@bit.edu.cn; liwei089@ieec.org; rantao@bit.edu.cn).

Pengge Ma is with the School of Intelligent Engineering, Zhengzhou University of Aeronautics, Zhengzhou 450000, China (e-mail: mapenge@163.com).

Shengheng Liu is with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: s.liu@seu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2020.2998822

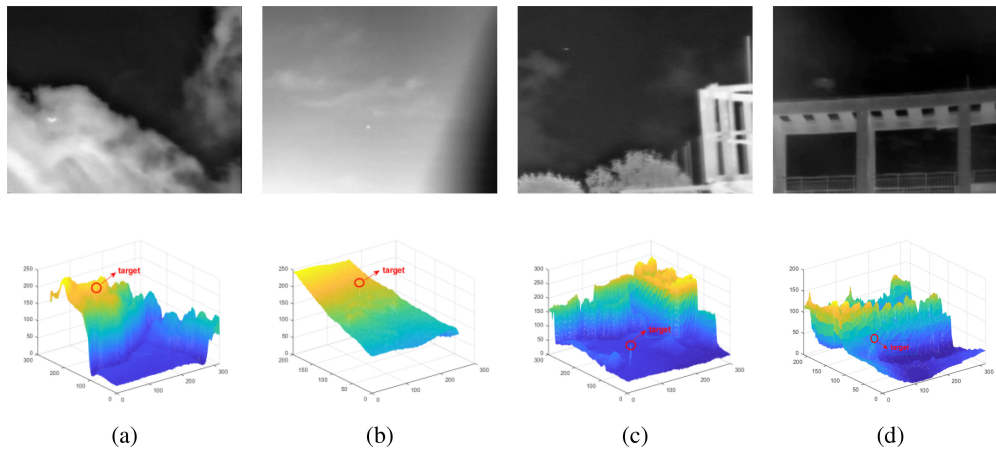


Fig. 1. Representative targets and corresponding 3-D maps in different backgrounds. Targets are marked by red circles on a 3-D map. (a) Heavy cloudy clusters. (b) Strong sunlight radiation. (c) Building with radiation. (d) Dark sky with cloud clutter and buildings.

have more target and background information available, such as motion information. Therefore, the detection algorithm based on multiframe has a better development prospects. Typical sequential detection methods include 3-D matched filtering [20], temporal profiles [21], hypothesis testing [22], maximum-likelihood estimation, and dynamic programming [23], [24]. However, these methods are not suitable for the detection of small target. Recently, many new sequential detection methods have been proposed. A new method based on in-frame and interframe information is proposed [25]. A small IR target detection method based on spatial-temporal enhancement using the quaternion discrete cosine transform (QDCT) is proposed by fusing the kurtosis feature, two-directional feature maps, and motion feature of image sequences [26]. Similarly, temporal and spatial information of sequential images is used to extract objects of interest in [27] and [28]. In addition, many methods based on sparse representation have been proposed for IR small target detection [29]–[31]. However, the current sequential detection methods do not achieve better detection performance.

By investigating the abovementioned literature, it is found that the single-frame detection methods only have good detection performance in specific scenes. However, sequential detection methods can utilize more target and background information. In order to solve the problem of IR small target detection, especially in the case of very low SNR, a new low-rank and sparse decomposition (LSD) method based on greedy bilateral factorization is proposed. In this article, the target is regarded as an independent sparse motion structure. An LSD model is established based on the time-domain motion information of the target. The theory guarantees that the target can be separated from the background and noise. The main contributions of this article are summarized as follows.

- 1) By analyzing the incoherent structure of IR image sequences, we separate the small IR target as an independent sparse motion structure from the background and noise. It overcomes the problem that the existing methods fail to detect the target in a highly heterogeneous scene. It is worth mentioning that the proposed method only takes

about 0.8 s to process 30 frames with a resolution of 540×398 . Moreover, the frame rate of the IR sensor is 30 frames per second. It is of great significance to the real-time performance of the algorithm.

- 2) The low-rank approximation based on greedy bilateral factorization is adopted in the background of the model, which significantly improves the detection efficiency of the algorithm. Also, the rank estimation of the low rank part is adaptive. Therefore, the proposed method can achieve fast and robust detection while using the inter-frame motion information.
- 3) In this article, the proposed method is verified in 8 real datasets. Compared with the existing datasets, the background of the datasets in this article is more complex, and the image SNR is very weak, which makes the proposed method more competitive than the existing methods.

The rest of this article is organized as follows. The related work of IR small target detection based on LSD is summarized in Section II. The detection framework of the proposed method, algorithm flow, and theoretical analysis are given in Section III. The results of the experiment in different scenes are shown, and the performance of different methods is analyzed with a series of evaluation indexes in Section IV. Finally, the work of this article is summarized and the future work is prospected in Section V.

II. RELATED WORK

With the rise of matrix decomposition theory, a number of IR small target detection methods based on LSD have emerged. Since the method in this article is also based on LSD, we briefly review the existing LSD-based IR small target detection methods in this section.

The most representative one is the method based on an IR patch-image (IPI) model, which uses the local nonautocorrelation of the image to transform the detection of IR small targets into the low-rank and sparse optimization problem [32]. Based on an IPI model, Wang *et al.* think that different radiation sources in the IR image background come from different subspaces,

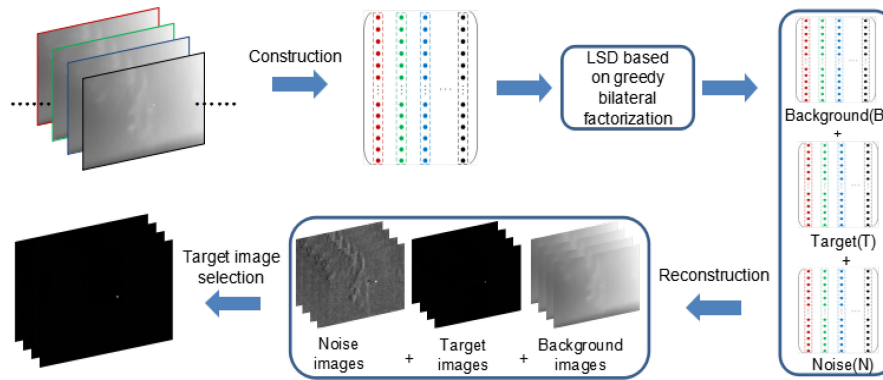


Fig. 2. Framework of the proposed method.

and then map the background to multiple subspaces. A method of IR dim and small target detection based on stable multiple subspace learning (SMSL) is proposed, where the background model of multi-subspace is constructed to realize the detection of IR dim and small target in complex scenes [33]. There are several methods to suppress background or enhance target by constraining a sparse term. Zhang *et al.* introduced structured $\ell_{2,1}$ norm to eliminate background strong residual and proposed a novel method based on nonconvex rank approximation minimization joint $\ell_{2,1}$ norm (NRAM) [34]. Zhang *et al.* used ℓ_p norm to strengthen the sparse term constraint. A new IR patch-tensor (IPT) model is proposed, which uses the prior information of image structure [35]. Based on the IPT model, a novel nonconvex low-rank constraint method named partial sum of tensor nuclear norm (PSTNN) joint weighted ℓ_1 norm is proposed [36]. The time information of image sequences is combined with the theory of low-rank matrix decomposition, and the IR small target is regarded as a special sparse noise component of complex background noise, which is modeled by Gaussian mixture and Markov random domain [37]. This method has achieved good results in complex environment, but it has a high computational complexity.

Matrix decomposition theory injects new vitality into IR small target detection, and has been widely used in the field of IR small target detection. However, most of the existing methods based on LSD are single-frame detection methods, which have high probability of false alarm under weak SNR. Also, most of the optimization algorithms have high computational complexity, which seriously restrict the real-time performance of the algorithm. Therefore, it is still a great challenge to detect IR small targets quickly and stably in complex environment.

III. PROPOSED METHOD

In this section, we first analyze the low-rank and sparse characteristics of IR image sequences in time domain. On this basis, we build an IR small target detection model and give the detection framework of the proposed method as shown in Fig. 2. Then, we describe the optimization algorithm based on greedy bilateral decomposition in detail and give the implementation

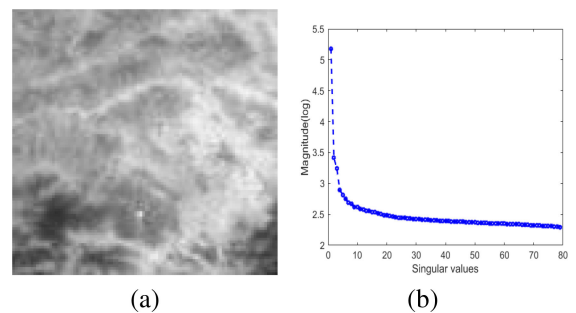


Fig. 3. (a) IR image scene. (b) Matrix singular value distribution.

steps of the algorithm. Finally, the computational complexity of the algorithm is analyzed.

A. Model Construction and Detection Framework

Compared with a single frame image, image sequences have the following two characteristics. First, most backgrounds change slowly from frame to frame and are therefore correlated in the time domain, whereas targets typically change quickly relative to backgrounds. As shown in Fig. 3, the distribution of singular values of the data observation matrix containing 80 frames is drawn, and it can be seen that its singular value converges rapidly. Therefore, it is considered that the background is low rank in time domain and the target is sparse. Second, compared with noise, the motion of small target is continuous in the space–time domain, and the distribution of noise is random in the space–time domain.

Most of the existing IR small target detection methods are based on single frame detection. Generally speaking, the target, noise, and background of IR image are independent of each other, which can be regarded as the superposition of three components. Here, the same IR image sequences model is considered as follows:

$$f_D = f_B + f_T + f_N \quad (1)$$

where f_B and f_T represent background image and target image respectively, and f_N is a noise image. Among them, f_B satisfies the low-rank property, f_T satisfies the sparse property, and f_N is usually random noise.

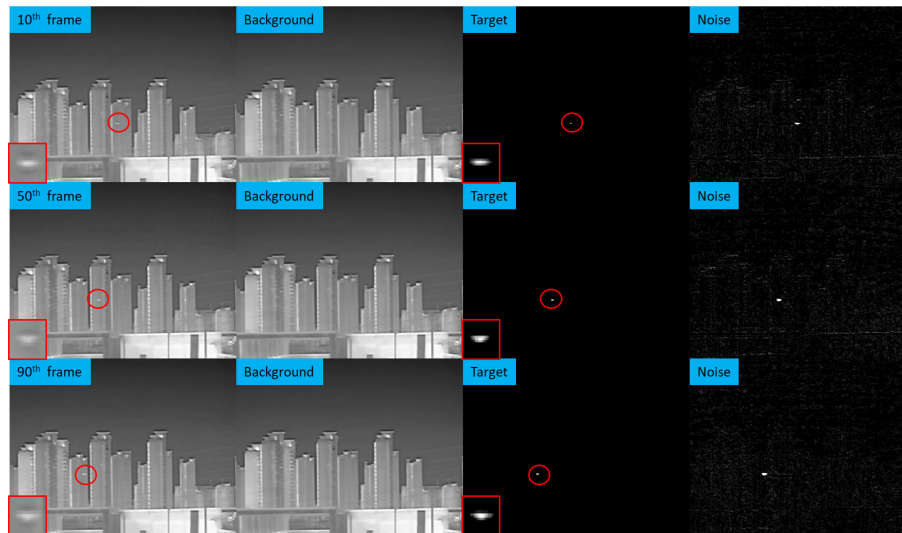


Fig. 4. Three-frame example (including target, background, and noise.). The target is marked with a red circle and a close-up is shown in the left-bottom part of each figure.

In order to make full use of the temporal information of image sequences to achieve better detection performance, an LSD model is established. Fig. 2 shows the framework of the proposed IR small target detection method. First, the pixel value of each frame in the image sequences is vectorized into a column vector in the spatiotemporal observation matrix. Then, LSD is implemented based on greedy bilateral factorization. Finally, the IR small target is extracted by reconstructing the spatiotemporal full-size images matrix. Three frames are extracted from a sequence scene, and the detection results of each frame obtained by the proposed detection framework are shown in Fig. 4. The decomposition includes a low-rank background, a sparse component containing IR moving targets, and a dense noise part.

B. Optimization Algorithm

Here, the following mathematical model of IR image sequences is established:

$$D = B + T + N \quad (2)$$

where D is the observation data matrix of IR image sequences, B is the background matrix, T is the target matrix, and N is the noise matrix.

For background matrix B , based on the abovementioned analysis, it is considered as a low-rank matrix and its rank satisfies

$$\text{rank}(B) \leq r \quad (3)$$

where r is the rank of the matrix B that is a constant, and the size of r is related to the correlation of the background. Generally, the smaller r is, the larger the background correlation will be, and the larger r is, the smaller the background correlation will be.

Accordingly, for the target matrix T , it is considered to be a sparse matrix, which satisfies the following requirements:

$$\|T\|_0 \leq j \quad (4)$$

where $\|\cdot\|_0$ is the ℓ_0 norm and j represents the number of nonzero entries, and $j \ll m \times n$ ($m \times n$ is the size of T), which means that most of the entries of the matrix T are zero.

In order to separate target T from background B and noise N , the problem of small target detection is transformed into the problem of low-rank and sparse matrix decomposition. In practice, the ℓ_1 norm is typically used to approximate the ℓ_0 norm and simplify the optimization problem. The following update rules are formulated:

$$\begin{aligned} \min_{B,T} \quad & \|D - B - T\|_F^2 + \lambda \|T\|_1 \\ \text{s.t.} \quad & \text{rank}(B) \leq r \end{aligned} \quad (5)$$

where $\|\cdot\|_1$ and $\|\cdot\|_F$ are the ℓ_1 norm and the Frobenius norm, respectively, and λ is a regularization parameter.

The low-rank parts of the existing LSD-based model are approximated by the first few principal components [38]. It is not difficult to find that the update of the low-rank part will call a complete singular value decomposition (SVD) per iterate. Thus, they suffer from a very low calculation efficiency. Especially for sequential detection methods, it is difficult to guarantee the real-time performance of the algorithm. Inspired by Zhou and Tao[39], the greedy bilateral (GreB) paradigm is chosen to model the low-rank part as bilateral factorization and update the left and right factors in a mutually adaptive and greedy incremental manner. The following update rules can be obtained by replacing B with its bilateral factorization $B = XY$:

$$\begin{aligned} \min_{X,Y,T} \quad & \|D - XY - T\|_F^2 + \lambda \|T\|_1 \\ \text{s.t.} \quad & \text{rank}(X) = \text{rank}(Y) \leq r \end{aligned} \quad (6)$$

where X is a tall matrix and Y is a broad matrix. It starts from X and Y , respectively, containing a very few columns and rows, and optimizes them alternately. It can be found that the object value in (6) is not determined by a single X or Y , but by the product XY . Different (X, Y) pairs can produce the same XY . Alternately

optimizing X , Y , and T in (4) immediately yields the following updating rules:

$$\begin{cases} X_k = (D - T_{k-1})Y_{k-1}^T(Y_{k-1}Y_{k-1}^T)^\dagger \\ Y_k = (X_k^T X_k)^\dagger X_k^T (D - T_{k-1}) \\ T_k = \mathcal{S}_\lambda(D - X_k Y_k) \end{cases} \quad (7)$$

where subscript k is the variable in the k th iterate and \dagger is the Moore–Penrose pseudoinverse, and \mathcal{S}_λ is an elementwise soft thresholding operator with threshold λ as follows:

$$\mathcal{S}_\lambda D = \max\{|D| - \lambda, 0\} \cdot \text{sgn}(D). \quad (8)$$

By observing (7), the following equation can be obtained:

$$X_k Y_k = X_k (X_k^T X_k)^\dagger X_k^T (D - T_{k-1}) = \mathcal{P}_{X_k}(D - T_{k-1}) \quad (9)$$

where \mathcal{P} is an orthogonal projection operator. In other words, $X_k Y_k$ is equal to the orthogonal projection of $(D - T_{k-1})$ on the column space of X_k . However, the column space of X_k can be indicated by any orthogonal basis for the columns of $(D - T_{k-1})Y_{k-1}^T$. Here, we can calculate it as Q by QR decomposing

$$(D - T_{k-1})Y_{k-1}^T = QR. \quad (10)$$

Thus, product $X_k Y_k$ can be calculated as

$$X_k Y_k = \mathcal{P}_Q(D - T_{k-1}) = QQ^T(D - T_{k-1}). \quad (11)$$

Therefore, the update process of the k th iteration in (7) is given by QR decomposition

$$\begin{cases} X_k = Q, \text{QR}((D - T_{k-1})Y_{k-1}^T) = QR \\ Y_k = Q^T(D - T_{k-1}) \\ T_k = \mathcal{S}_\lambda(D - X_k Y_k) \end{cases} \quad (12)$$

where the updates in (12) are iterated for K times or until the objective function converges. Then, add Δr extra rows into Y as the new directions for decreasing the object value. For $Y \in \mathbb{R}^{r_0 \times n}$ with a small integer r_0 , augment the rank of Y to $r_1 = r_0 + \Delta r$, where Δr is the rank step size. It is theoretically proved that the proposed method has the same approximate accuracy as SVD in [40]. The convergence is proved in [39], [41], and [42].

Here, we apply the strategy in [39] to select the added Δr rows as the top Δr right singular vectors of the following partial derivative:

$$\frac{\partial \|D - XY - T\|_F^2}{\partial Y} = D - XY - T \quad (13)$$

where the selected Δr rows maximize the magnitude of (13). Thus, the decomposition error can be reduced as quickly as possible.

The detailed algorithm steps are listed in the following table. Referring to the table, we can see that the algorithm contains two layers of loops. The inner *for loop* is to find a (X, Y) pair, so that it is the same as the (X_k, Y_k) product in (7), and the outer *while loop* is to find the loop after the fixed rank.

The proposed algorithm has the following two advantages. First, through the adaptive updation of X and Y , the optimization of the algorithm does not require the cyclic iteration of SVD. Second, the proposed algorithm repeatedly increases the rank until a sufficiently small decomposition error is achieved. So the

Algorithm 1: Proposed Target Detection Method Based on Greedy Bilateral Factorization.

Input: IR image sequences observation matrix D ; rank step size Δr ; power K ; tolerance τ

Output: low-rank matrix XY and sparse matrix T

- 1: Initialize $Y \in \mathbb{R}^{r_0 \times n}$ and T
 - 2: **while** residual error $\leq \tau$ **do**
 - 3: **for** $k = 1$ to K **do**
 - 4: Update X, Y and T by alternating minimization rules, calculate (12) and find a pair of (X, Y) that have the same product as (X_k, Y_k) in (7);
 - 5: **end for**
 - 6: Calculate the top Δr right singular vectors v of (13);
 - 7: Augment the rank of Y to $r_1 = r_0 + \Delta r$;
 - 8: Set $Y := [Y; v]$;
 - 9: **end while**
-

rank of the low-rank component is adaptively estimated and does not relies on initial rank value. Based on the theory, compared with the single-frame detection methods, the proposed method can use the interframe motion information to achieve fast and robust detection.

C. Computational Complexity

In this part, we briefly analyze the complexity of the proposed algorithm. For $X \in \mathbb{R}^{m \times r_i}$ and $Y \in \mathbb{R}^{r_i \times n}$, (12) can be performed in $3mnr_i + mr_i$ flops. Under the GreB paradigm, the time complexity of completing the whole matrix is $\mathcal{O}(\max\{\|\Omega\|_0 r^2, (m+n)r^3\})$, where Ω is the sampling set, $m \times n$ is the size of the observation matrix, and r is the rank of the matrix. In the iteration, the convergence speed of the algorithm can be accelerated by selecting the optimization direction with formula (13). However, for robust principal component analysis methods with SVD, the computational complexity is $\mathcal{O}(mnr^2)$.

IV. EXPERIMENTS AND ANALYSIS

Here, we introduce several commonly used IR dim and small target detection and evaluation indexes. In order to verify the effectiveness of the algorithm, we have carried out experiments on IR images in different scenes, and compared the proposed algorithm with the current algorithm with the best performance. In the experiment, the parameters of all algorithms are adjusted to achieve the best detection results.

A. Evaluation Metrics

In order to further quantify the advantages of the proposed algorithm, the performance of all methods is evaluated by using a receiver operating characteristic (ROC) curve. The ROC curve reflects the relationship between the probability of detection (P_d) and probability of false alarm (P_f). P_d and P_f are defined as follows:

$$P_d = \frac{n_t}{N_t} \quad (14)$$

$$P_f = \frac{n_f}{N} \quad (15)$$

TABLE I
DETAILS OF EIGHT IR TARGET DATASETS

	Frame Number	Image Resolution	Background Description	Average SNR Value
Scene 1	80	256 × 320	Sky scene with heavy cloudy clusters	2.2555
Scene 2	200	200 × 256	Complicated background with sunny sky	1.9972
Scene 3	200	318 × 210	Gloomy sky scene with cloudy clusters	2.1903
Scene 4	40	300 × 300	Gloomy sky scene with building	3.1495
Scene 5	200	330 × 230	Gloomy sky scene with building and cloud	3.2508
Scene 6	400	258 × 200	Sky scene with cloudy clusters and building	2.9202
Scene 7	100	540 × 398	Building scene with strong radiation bright spot suspected target	2.8465
Scene 8	200	480 × 358	Building scene with tower crane and tower crane hanger suspected target	4.6886

TABLE II
CALCULATION TIME OF DIFFERENT FRAMES (S)

Frame Number	10	20	30	40
Scene a	0.1135	0.2104	0.3027	0.3944

where n_t , N_t , n_f , and N represent the number of detected true pixels, ground-truth target pixels, false alarm pixels, and the total number of pixels in the image sequences, respectively.

The SNR is used to describe the difficulty of detecting IR small targets. Usually, the higher the SNR, the easier the target will be detected. The SNR is defined as follows:

$$\text{SNR} = (I_{\max} - I_{\text{mean}})/\sigma \quad (16)$$

where I_{\max} is the maximum gray value of the image, I_{mean} is the mean value of the image, and σ is the standard deviation. The SNR gain (SNRG) is defined as follows:

$$\text{SNRG} = 20 \times \log_{10}(\text{SNR}_{\text{out}}/\text{SNR}_{\text{in}}) \quad (17)$$

where SNR_{in} and SNR_{out} represent the SNRs of the input image and the output image, respectively. Generally, the higher the SNRG is, the better the target enhancement ability of the corresponding algorithm will be.

Another commonly used evaluation index is background suppression factor (BSF). BSF is used to describe the ability of background suppression. Generally, the higher the BSF is, the stronger the ability of background suppression will be. BSF is defined as follows:

$$\text{BSF} = \frac{C_{\text{in}}}{C_{\text{out}}} \quad (18)$$

where C_{in} and C_{out} represent the standard deviations of clutter in the input and output image, respectively.

B. Experimental Setup

1) *Datasets*: The effectiveness and robustness of the proposed method are verified by the image sequences of eight real scenes. Table I gives a detailed description of the target and background of all datasets. The target is a small UAV. The background of Scenes 1–3 is relatively simple. In Scene 4, the local background of the target is more uniform, and there are brighter buildings. Scenes 5–8 are highly heterogeneous scenes, where the target is mixed in the cloud layer and buildings, and there are suspected targets, such as point cloud and tower crane pendant. The average SNR values are given in Table I (data sets).

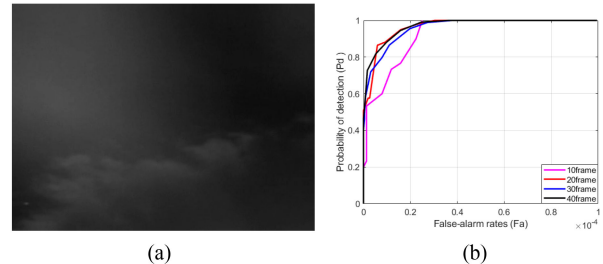


Fig. 5. ROC with different frames on the image sequences. (a) Representative frame of image sequences. (b) Corresponding ROC curve of the proposed method.

2) *Number of Frames*: Compared with single-frame detection method, sequential detection methods require higher performance, such as hardware storage. We choose one of the scenes to experiment in groups from 10 to 40 frames. Fig. 5 shows the ROC curves of each group of experiments. It can be seen that when P_f is low, P_d increases with the increase in the number of frames. Table II shows the time cost of each group of experiments. It can be seen that the time cost increases with the increase in the number of frames. The time cost of the proposed method for processing 30 frames is only about 0.3 s. The sampling rate of the IR sensor is 30 frames per second. Considering these factors, 30 frames were selected for each group of experiments. All the experiments are conducted on a computer with 8 GB of main memory and Intel Core i7-2600K CPU with 3.40 GHz. The code was implemented in MATLAB R2018b.

3) *Baseline Methods*: In order to reflect the performance of the proposed method, the classical methods and the current methods with better performance are used as the baseline methods. Classical methods include top-hat [1], max-mean/max-median [2], and LCM [4]. Methods based on LSD include IPI [32], NRAM [34], SMSL [33], and PSTNN [36]. These methods have achieved good detection performance in specific scene. However, these methods often fail to detect the target in complex environment. In addition, IPI and NRAM also have high computational complexity. QDCT [26] is a sequential detection method. FKRW [11] based on image segmentation is one of the best methods at present. The parameters of various methods are adjusted to achieve the best detection results. Table III shows the parameter settings for all methods.

C. Comparison of Detection Performance

The proposed algorithm is verified in eight real sequence scenes. Figs. 6–9 show the detection results of different methods

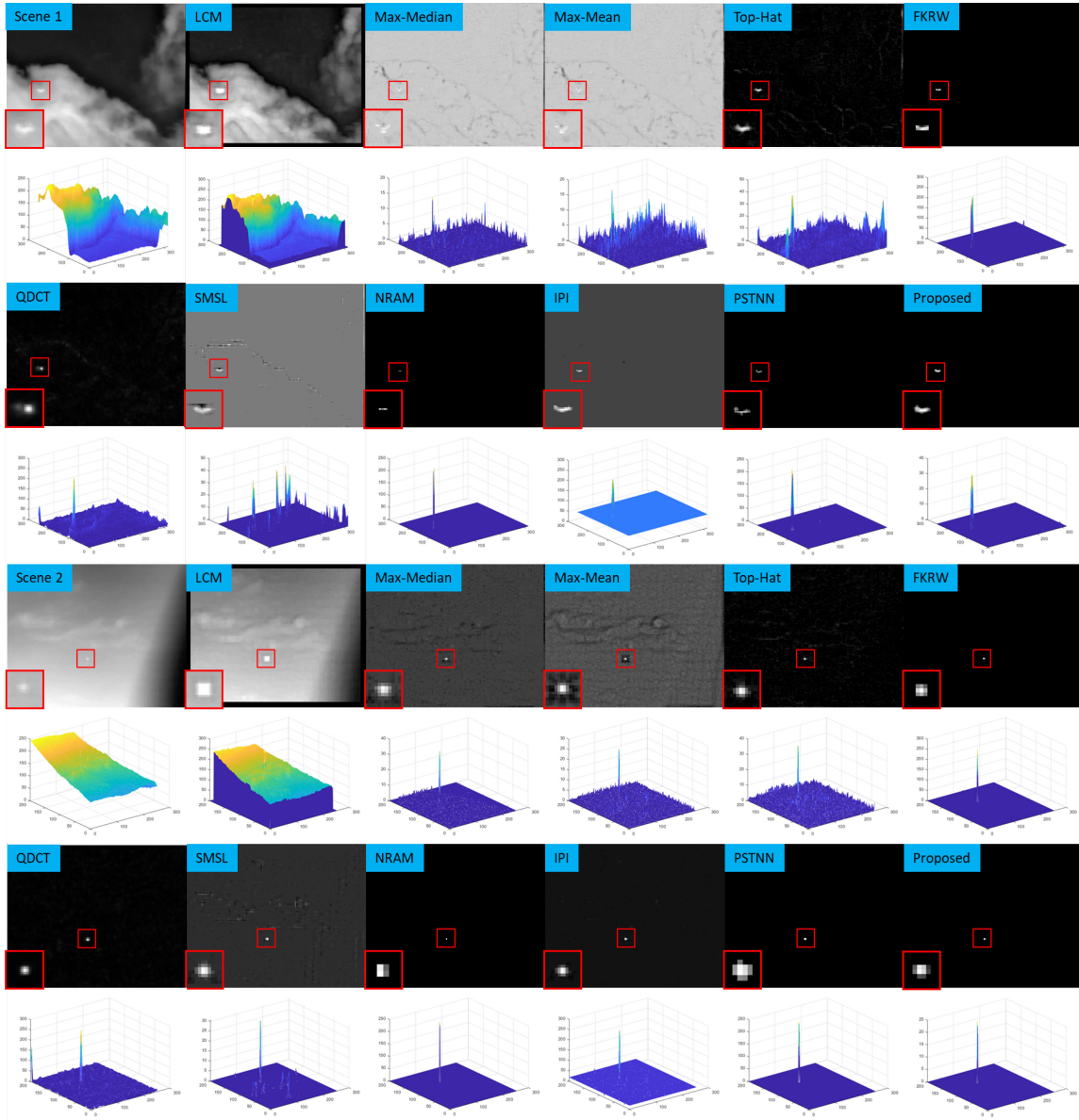


Fig. 6. Detection results and 3-D gray distribution map obtained by different methods on Scene 1 and Scene 2. The real target on the ground is displayed as a red rectangle with a close-up version in the left-bottom part of each figure.

TABLE III
PARAMETER SETTINGS OF DIFFERENT METHODS

No.	Methods	Acronyms	Parameter settings
1	Local Contrast Measure [4]	LCM	Local window size: $N=1,3,5,7$
2	Max-mean Filter [2]	Max-mean	Local window size: 5×5
3	Max-median Filter [2]	Max-median	Local window size: 5×5
4	Top-hat Filter [1]	Top-Hat	Local window size: 5×5
5	Facet Kernel and Random Walks [11]	FKRW	$K=4, p=6, \tau=200, \text{windows size: } 11 \times 11$
6	Quaternion Discrete Cosine Transform [26]	QDCT	Gaussian Filter: $\sigma=1.5, \tau=1, \text{Patch size: } 7 \times 7$
7	Stable Multiple Subspace Learning [33]	SMSL	Patch size: $30 \times 30, \text{sliding step: } 10, \lambda=3/\sqrt{\min(m, n)}, \mu^0=0.5^*s_4, \bar{\mu}=0.05^*s_5, \varepsilon=10^{-7}$
8	Non-convex Rank Approximation Minimization [34]	NRAM	Patch size: $50 \times 50, \text{sliding step: } 10, \lambda=1/\sqrt{\min(m, n)}, \mu^0=3\sqrt{\min(m, n)}, \gamma=0.002, C=\sqrt{\min(m, n)}/2.5, \varepsilon=10^{-7}$
9	IR Patch-Image [32]	IPI	Patch size: $50 \times 50, \text{sliding step: } 10, \lambda=1/\sqrt{\min(m, n)}, \varepsilon=10^{-7}$
10	Partial Sum of Tensor Nuclear Norm [36]	PSTNN	Patch size: $40 \times 40, \text{sliding step: } 40, \lambda=0.6/\sqrt{\min(n_1, n_2)} * n_3, \varepsilon=10^{-7}$
11	Proposed		$\text{rank}=2, \Delta r=1, \tau=10^{-3}, K=3$

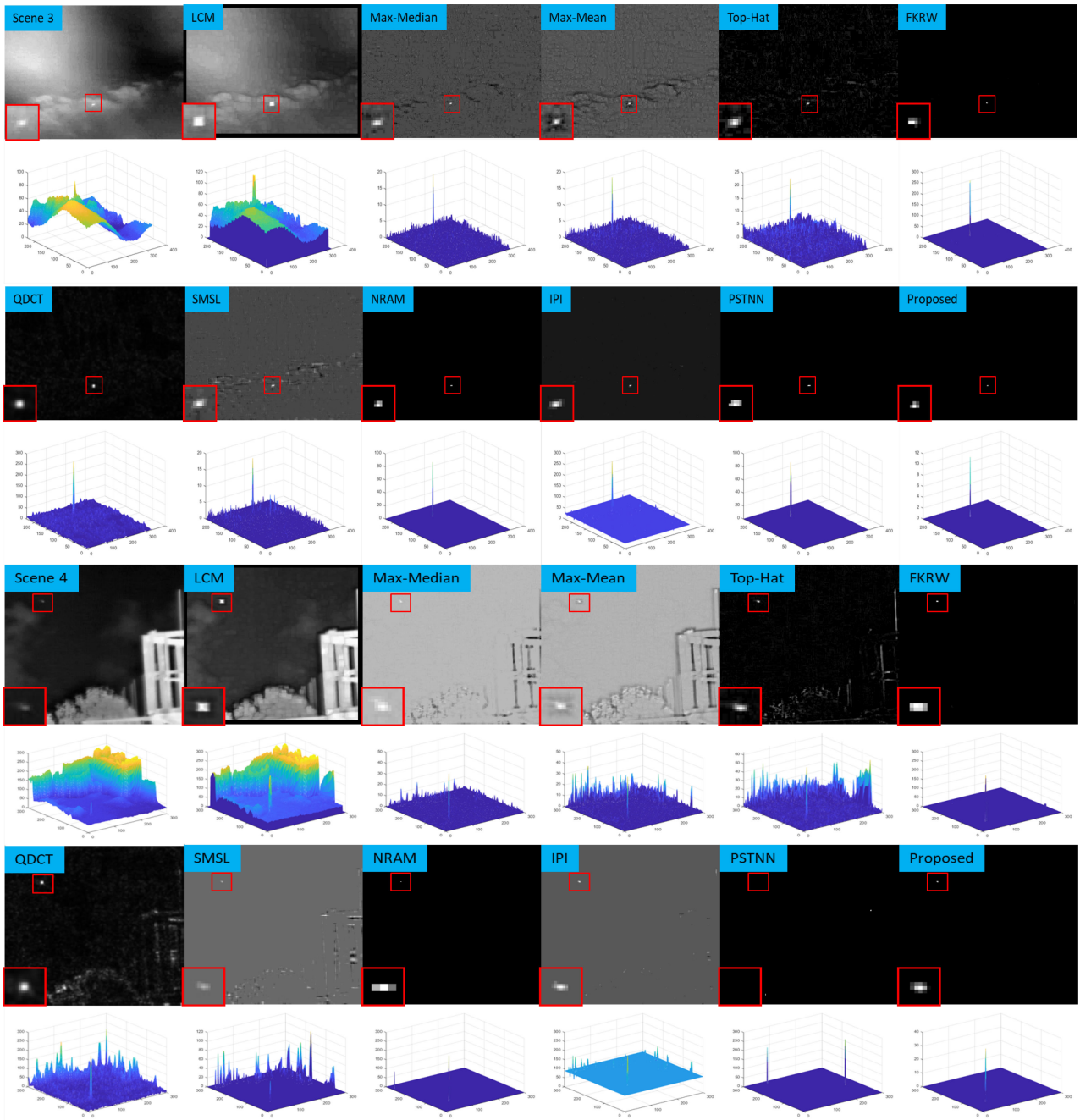


Fig. 7. Detection results and 3-D gray distribution map obtained by different methods on Scene 3 and Scene 4. The real target on the ground is displayed as a red rectangle with a close-up version in the left-bottom part of each figure.

and the corresponding 3-D display. The target is marked by a red rectangle, and each target is given a close-up in the lower left corner. The 3-D distribution of image can show the performance of detection method intuitively. From the 3-D distribution of the image, it can be seen that the detection results of the proposed method only retain the target, and there is almost no background noise. For Scene 1–3, the background is relatively single and the target has maximum brightness. The baseline methods such as SMSL, NRAM, PSTNN, and IPI have higher P_d . Compared with Scenes 1–3, the background building in Scene 4 is brighter. FKRW, NRAM, and the proposed method are still valid. It

can be seen from the 3-D distribution that FKRW and NRAM still retain a small amount of background clutter. The proposed method only retains the target. Although other baseline methods can also detect the target, their weak ability to suppress the background results in a high P_f . Scenes 5–8 have more complex backgrounds, which greatly increases the difficulty of target detection. Most baseline methods have a high false alarm rate or fail to detect the target. For Scene 8, the local background of the target is relatively clean, and tower crane hangers form interference targets. It can be found that in the detection results of QDCT, the real target has higher brightness than the

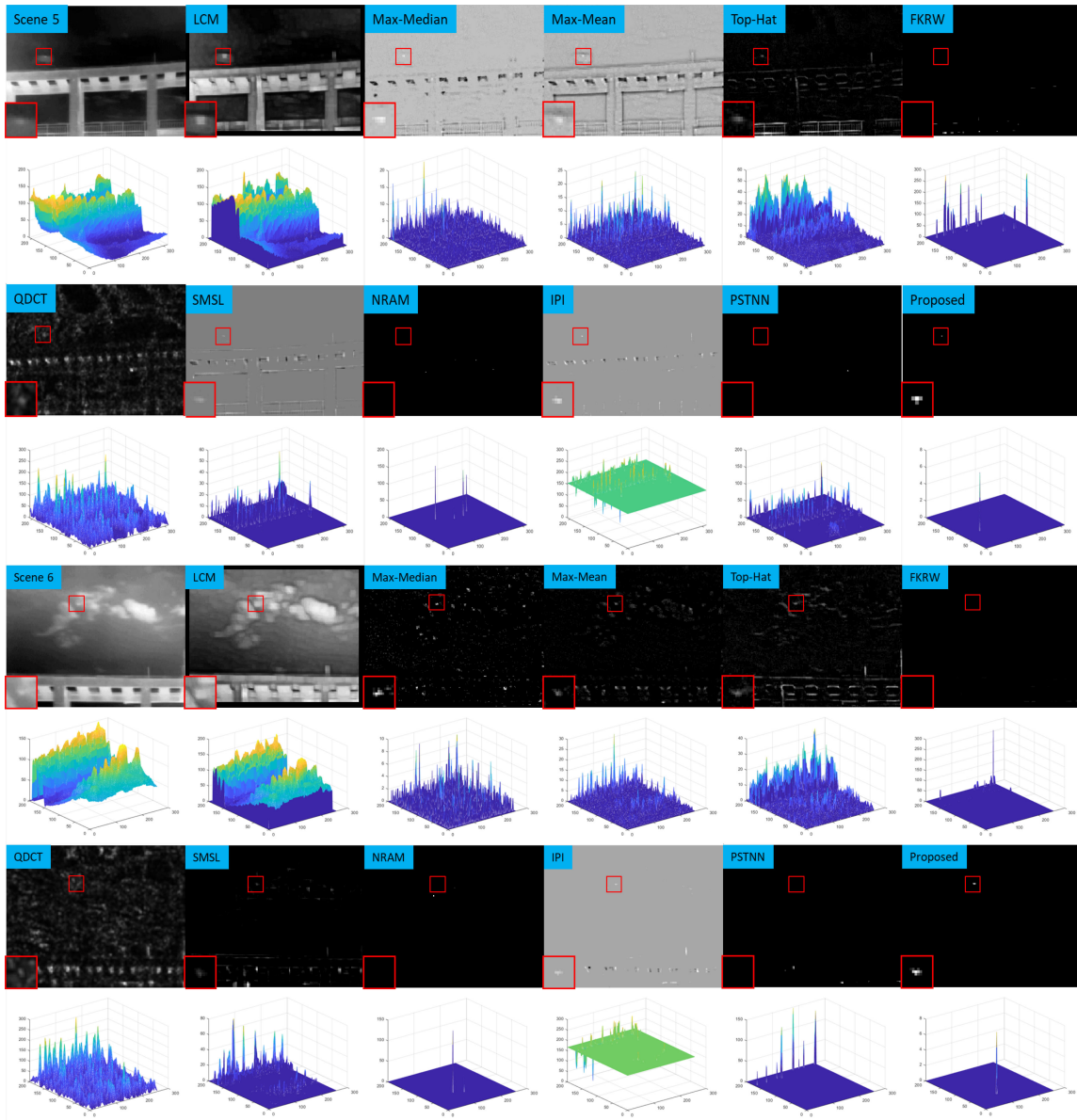


Fig. 8. Detection results and 3-D gray distribution map obtained by different methods on Scene 5 and Scene 6. The real target on the ground is displayed as a red rectangle with a close-up version in the left-bottom part of each figure.

suspected target. QDCT can suppress background clutter by using multiframe target motion information. FKRW can also detect targets. FKRW thinks that the background clutter tends to spread in a certain direction in a local small area, and uses the consistency of the direction of the background clutter to suppress the background clutter [11]. It can be seen from the close-up that the target detected by the proposed method is smoother than that those detected by FKRW and NRAM. It can be seen from the 3-D distribution that there are still a few edge background clutter in FKRW, and the proposed method has no background clutter. Our method suppresses the background while using timing information. For Scene 7, when the target and the building in the background overlap, other baseline methods will cause higher P_f , and the proposed method is still effective.

SNRG and BSF of all methods in different scenes are shown in Tables IV and V, respectively. The — mark indicates that the value is negative. Black bold type indicates the maximum value, and the second best method is marked with a horizontal line. It can be seen that the recently proposed baseline methods of FKRW, NRAM, PSTNN, and the proposed methods have higher SNRG and BSF values than those by traditional methods. This indicates that these methods have strong target enhancement capabilities and background suppression capabilities. In Scenes 1–3 and Scene 7 and 8, NRAM has the maximum SNRG value and the maximum BSF value, and the method proposed in this article is the second best. In other scenes, the proposed method has the largest values of SNRG and BSF. Although NRAM and PSTNN also achieve considerable SNRG and BSF, these methods fail to detect target in complex scenes. The proposed

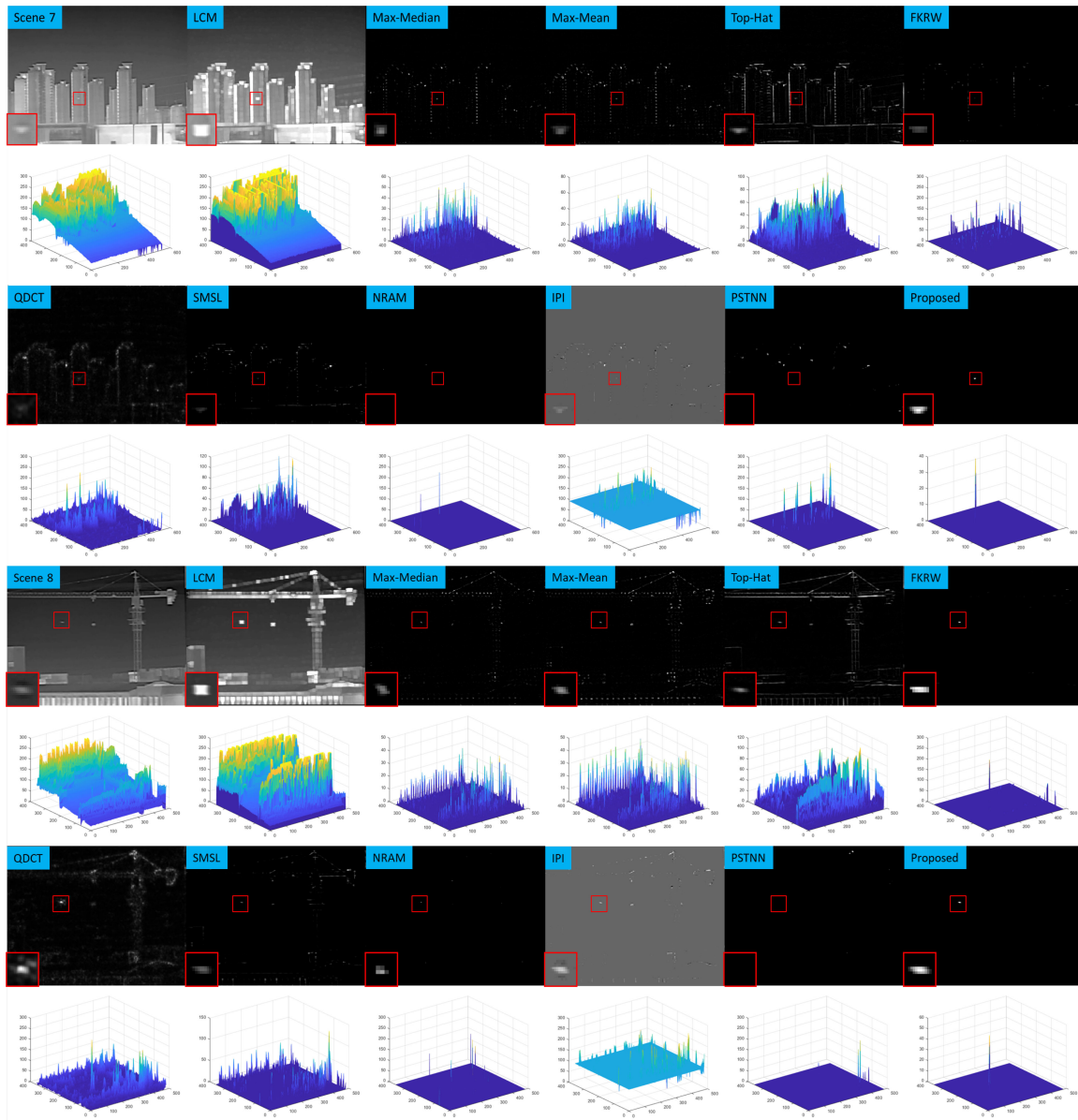


Fig. 9. Detection results and 3-D gray distribution map obtained by different methods on Scene 7 and Scene 8. The real target on the ground is displayed as a red rectangle with a close-up version in the left-bottom part of each figure.

method has high SNRG and BSF and high P_d in complex scenes. This is due to the fact that the proposed method uses the correlation of the background in the time domain to suppress the background.

Fig. 10 shows the ROC curves of different methods. In general, ROC curves in the upper left corner achieve better detection performance. In Scenes 1–4, the background is relatively simple. The baseline methods (such as IPI, NRAM, and PSTNN) still have a high P_d . The detection performance of the proposed method is not as good as that of IPI in Scene 1. The proposed method in other scenes has the highest P_d and the lowest P_f . However, the target is almost drowned in the background clutter in Scene 5. In Scene 6, the target is mixed into the cloud layer making it difficult to distinguish it from a suspicious target (such as a point cloud). In Scenes 7 and 8, some bright points with

large gray values in the background of the building will be mistaken for the target. These factors increase the difficulty of target detection, and single-frame detection methods often fail. However, our method still has higher P_d and lower P_f . The motion information of the target relative to the background is important information to separate the background and the target, and has been well applied in this article. We also calculated the area under curve (AUC) values for the different methods in Table VI. The AUC value is in the range of 0–1. Generally, the larger the AUC value is, the better the detection performance of the corresponding method will be. In Scene 1, IPI has the largest AUC value, and our method has the second-largest AUC value. In other scenes, the proposed method has the largest AUC value. For the AUC value of Scenes 5–8, it can be seen that the AUC value of the baseline method is significantly reduced,

TABLE IV
AVERAGE SNRG COMPARISON OF EIGHT REAL SCENES

	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5	Scene 6	Scene 7	Scene 8
LCM	0.5422	0.6418	3.0637	—	—	1.7610	—	—
Max-mean	23.3732	28.7375	23.7427	18.7178	16.8791	18.9916	21.6193	14.1895
Max-median	28.1977	32.3088	26.6737	26.6568	21.3538	18.2323	22.875	18.0342
Top-Hat	23.6409	26.2656	21.2774	15.8134	11.8479	13.9506	14.8745	10.5898
QDCT	26.4598	30.0192	26.6616	20.0066	13.4383	14.0775	20.4572	13.6057
FKRW	31.8615	35.1092	35.1889	32.5099	25.4191	31.6078	28.1477	27.4779
SMSL	26.043	32.0394	28.0161	21.0922	23.3699	19.8353	23.7241	20.6692
NRAM	36.1998	38.8672	36.8084	35.6703	33.2739	34.4124	40.5591	30.5525
IPI	29.5591	34.0517	34.6835	26.9477	18.4188	19.1971	25.1887	18.5475
PSTNN	28.1908	34.1955	34.5214	31.1643	20.5305	28.3407	25.0365	29.327
Proposed	32.1690	37.2854	40.1779	32.8293	33.8658	35.3668	35.4444	30.0353

TABLE V
AVERAGE BSF COMPARISON OF EIGHT REAL SCENES

	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5	Scene 6	Scene 7	Scene 8
LCM [4]	1.0045	0.9757	1.2724	0.9923	0.9873	1.0697	0.9833	0.9158
Max-mean [2]	10.5285	11.1632	8.8915	6.766	5.2988	5.0345	7.0080	3.7551
Max-median [2]	17.9915	16.677	12.4635	16.8251	8.8445	4.6343	8.0875	5.8318
Top-Hat [1]	10.7357	8.4923	6.8232	4.9222	3.0220	2.8808	3.2706	2.5064
QDCT [26]	14.5172	12.9514	12.7576	8.2248	3.8896	3.0938	6.3728	3.7882
FKRW [11]	25.2927	22.9133	32.4096	32.8665	14.2599	23.7078	15.0286	17.6812
SMSL [33]	12.9632	16.5419	14.6048	8.8521	11.1468	5.5127	8.9642	7.8956
NRAM [34]	45.1349	35.7154	38.9739	47.3806	34.9799	29.2106	63.1183	24.5325
IPI [32]	18.5387	21.8735	34.6484	25.1968	13.8484	13.3562	16.3613	10.2242
PSTNN [37]	16.6766	20.6397	29.9991	28.515	8.0320	14.6025	10.3477	21.3021
Proposed	26.3101	29.5695	55.1762	35.4860	37.1501	33.3278	34.5064	23.2706

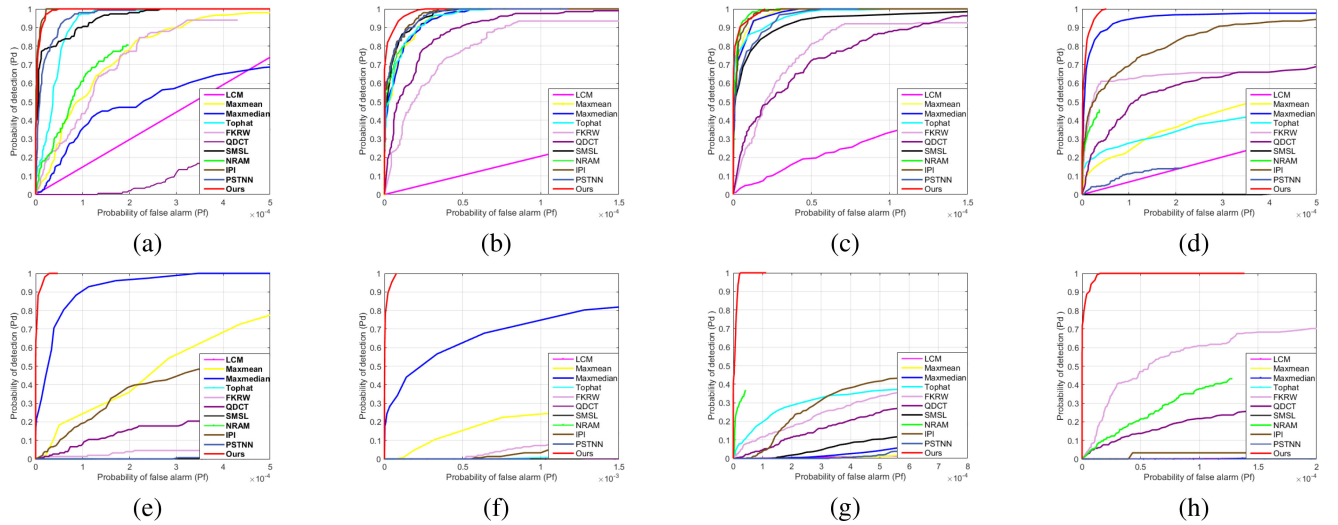


Fig. 10. (a)–(h) ROC curves of 11 methods in real sequence Scenes 1–8.

TABLE VI
AUC VALUES OF ALL METHODS IN EIGHT REAL SCEQUENCE SCENES ($\times 10^{-3}$)

	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5	Scene 6	Scene 7	Scene 8
LCM	24.6861	13.7885	20.3979	28.4788	20.3971	15.6825	67.0016	54.5429
Max-mean	0.5952	0.4587	0.5201	0.6661	0.5504	0.4782	0.9776	0.6351
Max-median	0.9532	0.6910	0.8103	0.9285	0.8563	0.7597	1.5708	1.1933
Top-Hat	0.1004	0.0977	0.1762	0.1034	0.1005	0.1085	0.1142	0.0130
QDCT ¹	0.6371	0.5468	0.6445	0.6257	0.6107	0.5227	1.0020	0.8824
FKRW	0.2212	0.1573	0.1457	0.2138	0.1510	0.2446	0.5964	0.242
SMSL	0.5832	0.7868	0.5711	0.7707	0.8508	0.7342	2.5293	1.9382
NRAM	3.0587	1.6591	2.6978	6.8566	3.0284	1.6416	44.2619	2.9419
IPI	21.2892	3.8154	16.1176	27.5263	13.4202	7.8853	267.656	117.2631
PSTNN	0.3614	0.0732	0.1218	0.4051	0.1371	0.1137	1.0114	0.7538
Proposed ²	0.3101	0.2055	0.2963	0.2353	0.2614	0.2136	0.7975	0.6452

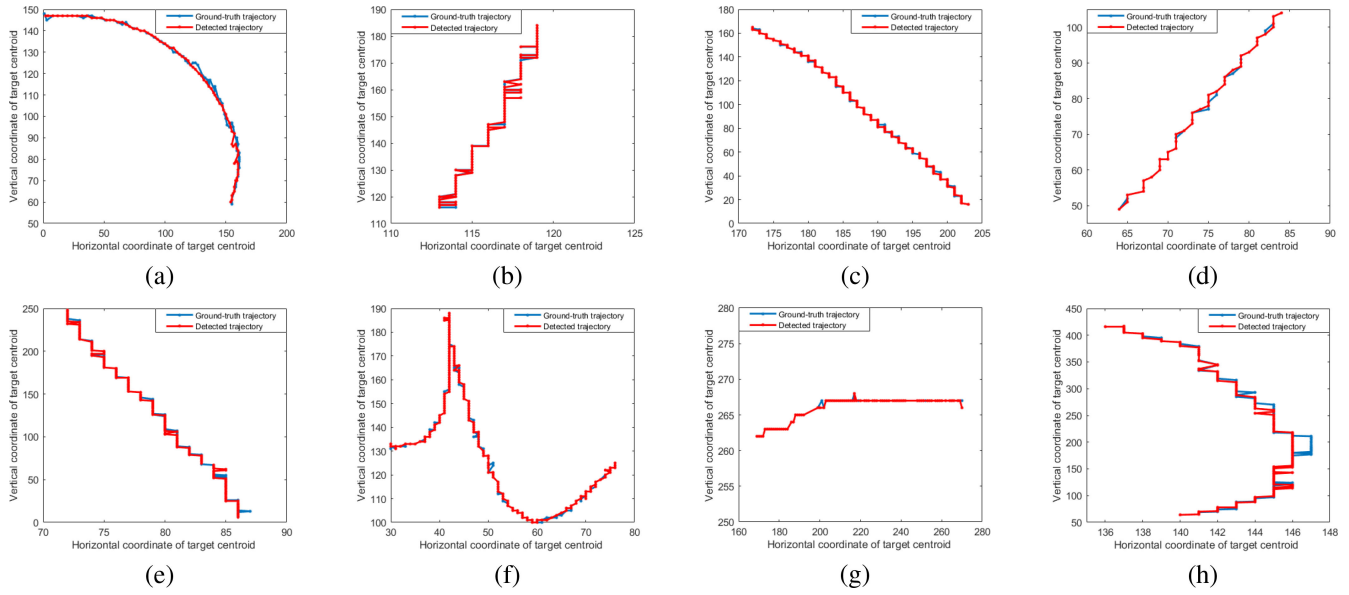


Fig. 11. Ground-truth and detected trajectories based on the proposed method in real sequence Scenes 1–8.

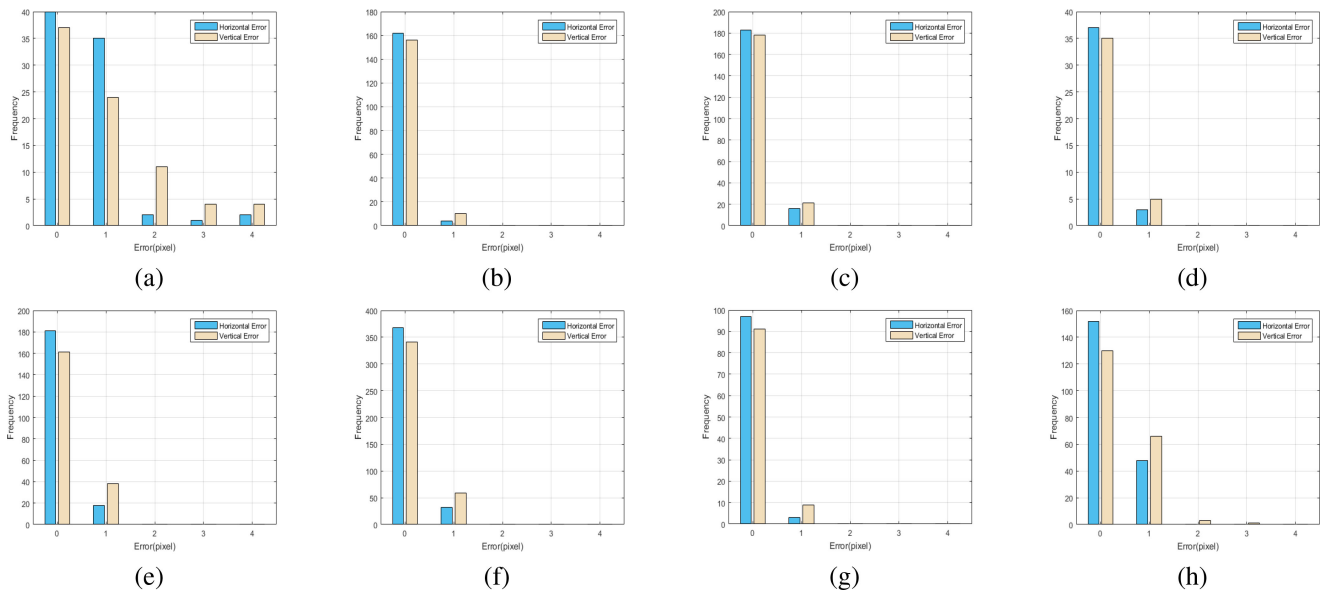


Fig. 12. Histograms of detected error distributions of real sequence scene 1–8.

whereas the AUC value of the proposed method is still close to 1. This indicates that the proposed method has better detection performance in complex scenes.

Considering that the target keeps moving in each image sequence, we choose the method presented in [13] to draw the ground-truth trajectory and detection trajectory curves in Fig. 11. In general, if the pixel error between centers of the ground-truth and the detected result is within 4 pixels, the detection is considered to be correct. Scene 2 contains 200 frames of images. From Fig. 11(b), it can be found that the vertical coordinate change range of the target in Scene 2 is 115–185, and the horizontal coordinate change range is only 113–118. This is because the target keeps hovering during the motion,

that is, the target’s motion speed is less than one pixel/frame. However, the proposed method still has high P_d and low P_f . Fig. 12 shows the corresponding histograms of error distributions. The horizontal coordinate is the pixel error and the vertical coordinate is the frequency of pixel error. It can be seen from Fig. 12 that the frequency of 4-pixel errors only appears in Scene 1. In other scenes, the horizontal pixel error and the vertical pixel error are less than 3 pixels, which indicates that the proposed method can detect the target stably. It can be seen that the detection trajectory is basically consistent with the ground-truth trajectory.

Table VII shows the calculation efficiency of different methods. All the experiments are conducted on a computer with 8 GB

TABLE VII
TIME COSTS COMPARISON OF ALL METHODS IN EIGHT REAL SCEQUENCE SCENES(S)

	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5	Scene 6	Scene 7	Scene 8
LCM	999.662	999.7579	996.88	966.8916	887.1712	898.9821	978.0812	982.3507
Max-mean	999.8581	999.9925	999.9968	993.6711	999.6702	996.3887	996.9701	997.6538
Max-median	923.4513	999.9923	999.9961	995.2515	999.9588	973.7911	973.1402	990.1002
Tophat	999.961	999.9927	999.995	999.2128	992.2031	991.9392	997.649	997.1258
QDCT	999.3715	999.9808	999.9585	998.0179	996.0453	963.4565	996.5837	998.2294
FKRW	969.6772	967.2972	962.7143	828.0099	528.6116	564.6276	894.1434	858.8568
SMSL	999.9745	999.9948	999.9881	993.7382	917.678	966.4786	996.9383	998.3261
NRAM	902.6114	899.4947	999.9981	727.7632	499.939	499.969	684.0679	716.1567
IPI	999.9968	999.9956	999.9974	999.6397	999.4278	897.1913	998.9054	999.5087
PSTNN	999.9808	999.9953	999.9936	570.7328	987.3153	499.5339	712.5133	499.8646
Proposed	999.9953	999.9983	999.9985	999.9941	999.9975	999.9926	999.9945	999.9924

¹Time costs of 2 frames.

²Time costs of 30 frames.

of main memory and Intel Core i7-2600K CPU with 3.40 GHz. The code was implemented in MATLAB R2018b. Among them, QDCT and the proposed method are sequential detection methods. QDCT selects at least two frames for experiments, and the proposed method selects 30 frames for experiments. Other methods are single-frame detection methods. Due to the use of multiscale, LCM takes a long time to process one frame. Other morphological methods, such as max-mean and max-median, take less time in a frame, and top-hat has the lowest time cost. However, these methods have lower P_d or higher P_f . SMSL, NRAM, PSTNN, and IPI are LSD-based methods. SMSL uses an optimization method based on acceleration approximate gradient, which results in lower time cost. However, this method does not obtain a high P_d and is not robust enough to noise. The time costs of NRAM and IPI are relatively large. The resolution size of Scene 7 is 540×398 , and the resolution size of Scene 8 is 480×358 . NRAM takes about 44 s and 3 s to process one frame, respectively, and IPI takes more than 100 s to process one frame. This is because the nuclear norm is used to approximate the low rank part, and each iteration requires SVD decomposition, which is time consuming. With the improvement of image resolution, the time costs of NRAM and IPI have increased significantly. In particular, IPI can hardly guarantee the real-time performance of the algorithm. Although the proposed method is a sequential detection method, it takes less time to process 30 frames. In addition, the proposed method uses SVD-free decomposition to ensure the real-time performance of the sequential detection method.

V. CONCLUSION

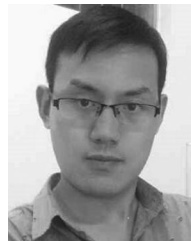
In this article, aiming at the problem of IR dim and small target detection, an LSD method based on greedy bilateral factorization is proposed. By using the interframe timing information of image sequences, the target is regarded as an independent sparse motion structure, which overcomes the problem of detection failure of existing methods in complex backgrounds. For most methods based on LSD, the optimization model is usually solved by SVD cycle iteration, which makes the algorithm less real time. In view of this, the proposed method uses GreB paradigm to model the low-rank part so as to avoid the cycle iteration of SVD. The proposed method can still use a small number

of frames and a small amount of time cost to achieve better detection performance. Extensive experiments demonstrate that the proposed method achieves better detection performance compared with the other baseline methods. Especially in Scenes 5–8, the proposed method has obvious advantages. However, the proposed method struggles to overcome backgrounds that change significantly from frame to frame. This is a subject of future work. In addition, we will consider whether the proposed algorithm is suitable for multitarget situations.

REFERENCES

- [1] J. Zhou, H. Lv, and F. Zhou, "Infrared small target enhancement by using sequential top-hat filters," in *Proc. Int. Symp. Optoelectron. Technol. Appl., Image Process. Pattern Recognit.*, 2014, vol. 9301, Art. no. 93011L.
- [2] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," in *Proc. SPIE*, 1999, vol. 3809, pp. 74–83.
- [3] H. Fan and C. Wen, "Two-dimensional adaptive filtering based on projection algorithm," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 832–838, Mar. 2004.
- [4] C. P. Chen, H. Li, Y. Wei, T. Xia, and Y. Y. Tang, "A local contrast method for small infrared target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 574–581, Jan. 2014.
- [5] J. Han, Y. Ma, B. Zhou, F. Fan, K. Liang, and Y. Fang, "A robust infrared small target detection algorithm based on human visual system," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 12, pp. 2168–2172, Dec. 2014.
- [6] J. Han, K. Liang, B. Zhou, X. Zhu, J. Zhao, and L. Zhao, "Infrared small target detection utilizing the multiscale relative local contrast measure," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 612–616, Apr. 2018.
- [7] Y. Shi, Y. Wei, H. Yao, D. Pan, and G. Xiao, "High-boost-based multiscale local contrast measure for infrared small target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 33–37, Jan. 2018.
- [8] Y. Wei, X. You, and H. Li, "Multiscale patch-based contrast measure for small infrared target detection," *Pattern Recognit.*, vol. 58, pp. 216–226, 2016.
- [9] X. Guan, Z. Peng, S. Huang, and Y. Chen, "Gaussian scale-space enhanced local contrast measure for small infrared target detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 327–331, Feb. 2020.
- [10] C. Xia, X. Li, and L. Zhao, "Infrared small target detection via modified random walks," *Remote Sens.*, vol. 10, no. 12, 2018, Art. no. 2004.
- [11] Y. Qin, L. Bruzzone, C. Gao, and B. Li, "Infrared small target detection based on facet kernel and random walker," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 7104–7118, Sep. 2019.
- [12] S. Huang, Z. Peng, Z. Wang, X. Wang, and M. Li, "Infrared small target detection by density peaks searching and maximum-gray region growing," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 12, pp. 1919–1923, Dec. 2019.
- [13] H. Deng, X. Sun, and X. Zhou, "A multiscale fuzzy metric for detecting small infrared targets against chaotic cloudy/sea-sky backgrounds," *IEEE Trans. Cybern.*, vol. 49, no. 5, pp. 1694–1707, May 2019.

- [14] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Small infrared target detection based on weighted local difference measure," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 4204–4214, Jul. 2016.
- [15] X. Bai and Y. Bi, "Derivative entropy-based contrast measure for infrared small-target detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 2452–2466, Apr. 2018.
- [16] X. Cao, C. Rong, and X. Bai, "Infrared small target detection based on derivative dissimilarity measure," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 8, pp. 3101–3116, Aug. 2019.
- [17] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Entropy-based window selection for detecting dim and small infrared targets," *Pattern Recognit.*, vol. 61, pp. 66–77, 2017.
- [18] H. Deng, X. Sun, M. Liu, C. Ye, and X. Zhou, "Infrared small-target detection using multiscale gray difference weighted image entropy," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 52, no. 1, pp. 60–72, Feb. 2016.
- [19] J. Gao, Y. Guo, Z. Lin, W. An, and J. Li, "Robust infrared small target detection using multiscale gray and variance difference measures," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 12, pp. 5039–5052, Dec. 2018.
- [20] I. S. Reed, R. M. Gagliardi, and L. B. Stotts, "Optical moving target detection with 3-D matched filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 327–336, Jul. 1988.
- [21] D. Liu, Z. Li, X. Wang, and J. Zhang, "Moving target detection by nonlinear adaptive filtering on temporal profiles in infrared image sequences," *Infrared Phys. Technol.*, vol. 73, pp. 41–48, 2015.
- [22] A. P. Tzannes and D. H. Brooks, "Detecting small moving objects using temporal hypothesis testing," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 2, pp. 570–586, Apr. 2002.
- [23] B. Balasingam, Y. Bar-Shalom, P. Willett, and K. Pattipati, "Maximum likelihood detection on images," in *Proc. 20th Int. Conf. Inf. Fusion*, 2017, pp. 1–8.
- [24] R. Succary, A. Cohen, P. Yaractzi, and S. R. Rotman, "Dynamic programming algorithm for point target detection: Practical parameters for DPA," in *Proc. SPIE*, 2001, vol. 4473, pp. 96–100.
- [25] M. Wan, G. Gu, E. Cao, X. Hu, W. Qian, and K. Ren, "In-frame and inter-frame information based infrared moving small target detection under complex cloud backgrounds," *Infrared Phys. Technol.*, vol. 76, pp. 455–467, 2016.
- [26] P. Zhang, X. Wang, X. Wang, C. Fei, and Z. Guo, "Infrared small target detection based on spatial-temporal enhancement using quaternion discrete cosine transform," *IEEE Access*, vol. 7, pp. 54712–54723, 2019.
- [27] Y. Li, Y. Zhang, J.-G. Yu, Y. Tan, J. Tian, and J. Ma, "A novel spatio-temporal saliency approach for robust dim moving target detection from airborne infrared image sequences," *Inf. Sci.*, vol. 369, pp. 548–563, 2016.
- [28] L. Dong, B. Wang, M. Zhao, and W. Xu, "Robust infrared maritime target detection based on visual attention and spatiotemporal filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 3037–3050, May 2017.
- [29] Z. Li *et al.*, "Dim moving target detection algorithm based on spatio-temporal classification sparse representation," *Infrared Phys. Technol.*, vol. 67, pp. 273–282, 2014.
- [30] H. Qin *et al.*, "Infrared small moving target detection using sparse representation-based image decomposition," *Infrared Phys. Technol.*, vol. 76, pp. 148–156, 2016.
- [31] Y. Lu, S. Huang, and W. Zhao, "Sparse representation based infrared small target detection via an online-learned double sparse background dictionary," *Infrared Phys. Technol.*, vol. 99, pp. 14–27, 2019.
- [32] C. Gao, D. Meng, Y. Yang, Y. Wang, X. Zhou, and A. G. Hauptmann, "Infrared patch-image model for small target detection in a single image," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4996–5009, Dec. 2013.
- [33] X. Wang, Z. Peng, D. Kong, and Y. He, "Infrared dim and small target detection based on stable multisubspace learning in heterogeneous scene," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 10, pp. 5481–5493, Oct. 2017.
- [34] L. Zhang, L. Peng, T. Zhang, S. Cao, and Z. Peng, "Infrared small target detection via non-convex rank approximation minimization joint l_2, l_1 norm," *Remote Sens.*, vol. 10, no. 11, 2018, Art. no. 1821.
- [35] Y. Dai and Y. Wu, "Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3752–3767, Aug. 2017.
- [36] L. Zhang and Z. Peng, "Infrared small target detection based on partial sum of the tensor nuclear norm," *Remote Sens.*, vol. 11, no. 4, 2019, Art. no. 382.
- [37] C. Gao, L. Wang, Y. Xiao, Q. Zhao, and D. Meng, "Infrared small-dim target detection based on Markov random field guided noise modeling," *Pattern Recognit.*, vol. 76, pp. 463–475, 2018.
- [38] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, 2011, Art. no. 11.
- [39] T. Zhou and D. Tao, "Greedy bilateral sketch, completion & smoothing," in *Proc. Int. Conf. Artif. Intell. Statist.*, Apr. 2013, pp. 650–658.
- [40] N. Halko, P.-G. Martinsson, and J. A. Tropp, "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions," *SIAM Rev.*, vol. 53, no. 2, pp. 217–288, 2011.
- [41] T. Zhou and D. Tao, "Unmixing incoherent structures of big data by randomized or greedy decomposition," 2013, *arXiv:1309.0302*.
- [42] S. Shalev-Shwartz, A. Gonen, and O. Shamir, "Large-scale convex minimization with a low-rank constraint," 2011, *arXiv:1106.1622*.



Dongdong Pang received the B.S. degree in communication engineering from Zhengzhou University of Aeronautics, Zhengzhou, China, in 2015, and the M.S. degree in communication engineering from Zhengzhou University, Zhengzhou, China, in 2018. He is currently working toward the Ph.D. degree in electronics and communications engineering with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China.

His research interests include image processing and infrared target detection.



Tao Shan (Member, IEEE) was born in 1969. He received the B.S. degree in electronic engineering from Xidian University, Xi'an, China, in 1991, and the Ph.D. degree in information and communications engineering from the Beijing Institute of Technology, Beijing, China, in 2004.

From 2014 to 2015, he was a Senior Visiting Scholar with the Center for Advanced Communications, Villanova University, Villanova, PA, USA. He is currently a Professor with the School of Information and Electronics, Beijing Institute of Technology. His

research interests include radar signal processing and time-frequency analysis for nonstationary signals.

Dr. Shan was the recipient of the first prize of Science and Technology Progress awarded by the Ministry of Education, in 2006 and 2007, respectively.



Wei Li (Senior Member, IEEE) received the B.E. degree in telecommunications engineering from Xidian University, Xi'an, China, in 2007, the M.S. degree in information science and technology from Sun Yat-Sen University, Guangzhou, China, in 2009, and the Ph.D. degree in electrical and computer engineering from Mississippi State University, Starkville, MS, USA, in 2012.

Subsequently, he spent one year as a Postdoctoral Researcher with the University of California, Davis, CA, USA. He is currently with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. His

research interests include hyperspectral image analysis, pattern recognition, and data compression.

Dr. Li is an active Reviewer for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE REMOTE SENSING LETTERS, and IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS). He is currently an Associate Editor for both the IEEE SIGNAL PROCESSING LETTERS and the IEEE JSTARS. He was the Guest Editor for Special Issue of *Journal of Real-Time Image Processing, Remote Sensing*, and IEEE JSTARS. He was the recipient of the 2015 Best Reviewer Award from IEEE Geoscience and Remote Sensing Society for his service for the IEEE JSTARS.



Pengge Ma received the B.S. degree in electronic engineering from Beihang University, Beijing, China, in 1996, the M.S. degree in electronic engineering from the School of Physical Science and Engineering, Zhengzhou University, Zhengzhou, China, in 2004, and the Ph.D. degree in communication and information system from the School of Information Engineering, Zhengzhou University, in 2012.

Since 1996, he has been with the Zhengzhou University of Aeronautics, Zhengzhou, China. His current research interests include signal processing,

photoelectric detection, target detection, and recognition.

Dr. Ma was the recipient of the first prize of Henan Natural Science Thesis and the second prize of Henan Natural Science Thesis, in 2013. He presided over and received the second prize of scientific and technological progress in Henan Province and the first prize of scientific and technological achievements of Henan Provincial Department of Education, in 2014. He was recognized as the Academic and Technical Leader of Henan Provincial Education Department, in 2015. He presided over and received the third prize of Scientific and Technological Progress in Henan Province, in 2017.



Shengheng Liu (Member, IEEE) received the B.Eng. and Ph.D. degrees in electronics engineering from the School of Information and Electronics, Beijing Institute of Technology, Beijing, China, in 2010 and 2017, respectively.

He is currently an Associate Professor with the School of Information Science and Engineering, Southeast University (SEU), Nanjing, China. Prior to joining SEU, he held a postdoctoral position with the Institute for Digital Communications, The University of Edinburgh, Edinburgh, U.K., from 2017 to 2018.

From 2015 to 2016, he was a Visiting Research Associate with the Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA, USA, under the support of the China Scholarship Council. His research interests mainly focus on intelligent sensing and wireless communications.

Dr. Liu was the recipient of the 2017 National Excellent Doctoral Dissertation Award from the China Institute of Communications.



Ran Tao (Senior Member, IEEE) was born in 1964. He received the B.S. degree from the Electronic Engineering Institute of PLA, Hefei, China, in 1985, and the M.S. and Ph.D. degrees from Harbin Institute of Technology, Harbin, China, in 1990 and 1993, respectively.

In 2001, he was a Senior Visiting Scholar with the University of Michigan, Ann Arbor, MI, USA. He is currently a Professor with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. He has authored or coauthored three

books and more than 100 peer-reviewed journal articles. His current research interests include fractional Fourier transform and its applications and theory and technology for radar and communication systems.

Dr. Tao has been the Chief Professor of the Creative Research Groups of the National Natural Science Foundation of China since 2014, and he was a Chief Professor of the Program for Changjiang Scholars and Innovative Research Team with a university during 2010 to 2012. He is currently the Vice-Chair of IEEE China Council. He is also the Vice-Chair of the International Union of Radio Science (URSI) China Council and a member of Wireless Communication and Signal Processing Commission of URSI. He was a recipient of the first prize of Science and Technology Progress, in 2006 and 2007, and the first prize of Natural Science, in 2013, both awarded by the Ministry of Education. He was the also recipient of the National Science Foundation of China for Distinguished Young Scholars, in 2006, and the Distinguished Professor of Changjiang Scholars Program, in 2009.