Overlapping Group Sparsity TV with Lp-Norm Constraint for Infrared Dim Small Target Detection

Yebin Chen, Jingjing Gui, and Wei Xue, *Member, IAENG*

***Abstract*—Detecting dim small targets in infrared image is a critical aspect of infrared search and tracking systems. However, the conventional detection algorithms are often limited by the impact of complex background and noise. Therefore, we propose an infrared dim small target detection model that combines overlapping group sparse total variation and L_p-norm constraints. The introduction of overlapping group sparse total variation regularization effectively suppresses noise in the image and enhances the detection capability for small targets. Additionally, by incorporating the L_p-norm constraint, the sparsity of the targets is enhanced, while the low-rank component constraint is adaptively adjusted. During the detection phase, the infrared image is initially into block-based matrices, followed by optimization through the alternating direction method of multipliers. Tests confirm that this detect effectively infrared small targets, suppressing background noise in challenging environments, and delivers a boost in detection performance.**

***Index Terms*—Infrared dim small target detection, nuclear norm, overlapping group sparse total variation, L_p-norm.**

I. Introduction

INFRARED dim small target detection is a critical component in search and tracking systems [1, 2]. Detecting infrared targets amid complex backgrounds remains an ongoing technical hurdle due to their diminutive scale, minimal contrast, and weak signal relative to noise. Compounding the difficulty, these targets often lack clear features, combined with background clutter interference, makes their detection a persistent challenge.

To address this challenge, extensive research has been devoted to developing effective methods for identifying dim and small infrared targets in various scenarios. Among various detection techniques, suitable filtering methods can enhance the ability to detect small targets [3]; in particular, sparse low-rank matrix recovery methods are noteworthy. Gao [4] introduced the infrared patch image (IPI) model, employing nuclear along with L_1-norms as optimal convex approximations for rank functions and sparse features. Dai [5] incorporating a locally adaptive kernel [6], developed the weighted infrared patch image (WIPI) model. Dai [7] proposed the non-negative infrared patch image (NIPPS) model [8]. In [9], Wang employed the total variation framework [10] to isolate pronounced edges in infrared imagery, a method termed TV-PCP. Zhang [11] used the L_p-norm to describe the target and introduce a non-convex optimization (NOLC) model using L_p-norm constraints. Rawat [12] introduced a novel model utilizing Bi-nuclear norm minimization (NNM). Despite the considerable success of IPI model-based methods in recent research, they still face challenges in effectively suppressing background interference in complex environments.

This paper proposes a method for detecting infrared dim small targets that incorporates overlapping group sparse total variation (OGSTV) regularization alongside L_p-norm constraints. Combining non-convex L_p-norm optimization with the structured sparsity of OGSTV, the model achieves background suppression and noise reduction for infrared small target detection.

This paper makes the following primary contributions. (1) A new model for detecting dim small targets in infrared images combines OGSTV and L_p-norm constraints, providing strong background suppression and effective target extraction capabilities. (2) Non-convex optimization is employed to overcome the limitations of traditional convex optimization methods in sparse representation. (3) The model is effectively resolved through the integration of the majorization-minimization (MM) algorithm and the alternating direction method of multipliers (ADMM) optimization framework, ensuring computational efficiency.

II. Method

We begin by introduce two fundamental concepts and then detail the proposed detection model along with its optimization algorithm.

A. Basic Concepts

1) L_p-norm: Let B has a singular value decomposition expressed as $B = U * S * V^T$, S is a diagonal matrix of singular values. The definition of the one norms of B is as Eq. (1), where $\|B\|_p$ denotes L_p-norm.

$$\|B\|_p = \left(\sum_{h=1}^{m} \sum_{f=1}^{n} |a_{hf}|^p \right)^{1/p}, 0 < p < \infty \qquad (1)$$

Here, a_{hf} denotes the pixel values in the hth row and fth column of matrix B.

The sparsity of the solution in the context of L_p-norm is directly tied to the value of p. Mathematically, the behavior of L_p-norm function with different p values determines the distribution of nonzero elements in the solution. As discussed in [11], when $p > 1$, the norm tends to yield a solution with a more uniform distribution of non-zero values, resulting in a less sparse outcome. Conversely, when $p \leq 1$, the function

Manuscript received January 8, 2025; revised July 17, 2025. This work was supported in part by the Natural Science Foundation of the Anhui Higher Education Institutions of China (2022AH050310, 2023AH040149) and the Anhui Provincial Natural Science Foundation (2208085MF168).

Yebin Chen is a professor of the School of Computer Science and Technology, Anhui University of Technology, Maanshan 243032, China (e-mail: cyb7102@163.com).

Jingjing Gui is a postgraduate student at the School of Computer Science and Technology, Anhui University of Technology, Maanshan 243032, China (e-mail: gjingjing0523@163.com).

Wei Xue is an associate professor of the School of Computer Science and Technology, Anhui University of Technology, Maanshan 243032, China (Corresponding author; e-mail: xuewei@ahut.edu.cn).

promotes values to focus on a small number of elements, resulting to a sparse solution where two elements are zero and the intersection point is located at the coordinate axes. Hence, sparsity is achieved when $p \leq 1$. The L_1-norm is a special case in this paper when L_p=1. Relative to the L_1-norm, adopting an L_p-norm regularization approach markedly enhances the capability of the algorithm to recover sparse signals [13, 14, 15, 16].

2) Overlapping group sparsity total variation: For a matrix $a \in \mathbb{R}^{n\times n}$, we define a $C \times C$ point group at the location (h, f) of a

$$a_{hf,C} = \begin{bmatrix} a_{h-m_1,f-m_1} & \cdots & a_{h-m_1,f+m_2} \\ a_{h-m_1+1,f-m_1} & \cdots & a_{h-m_1+1,f+m_2} \\ \vdots & \ddots & \vdots \\ a_{h+m_2,f-m_1} & \cdots & a_{h+m_2,f+m_2} \end{bmatrix} \in \mathbb{R}^{C\times C} \tag{2}$$

Here, $m_1 = \lfloor \frac{C-1}{2} \rfloor$, $m_2 = \lfloor \frac{C}{2} \rfloor$. $\lfloor e \rfloor$ denotes the greatest integer not exceeding e. The matrix $\tilde{a}_{h,f,C}$ is centered at (h, f), and $a_{h,f,C}$ is obtained by vectorizing $\tilde{a}_{h,f,C}$, i.e., $a_{h,f,C} = \tilde{a}_{h,f,C}(:)$. Then the overlapping group sparsity total variation regularizer can be expressed as follows:

$$OT(a) = \sum_{h,f} ||a(h, f)_C||_2 \tag{3}$$

Following the approach in [17], We define the regularization functional $OT(a)$ as follows:

$$OT(a) = OT(D^{(1)}a) + OT(D^{(2)}a) \tag{4}$$

where $OT(D^{(1)}a)$ and $OT(D^{(2)}a)$ denote the finite difference matrices for horizontal and vertical directions, respectively.

B. Proposed Model

Traditional infrared image model is expressed as (5):

$$f(x, y) = f_B(x, y) + f_T(x, y) + f_N(x, y) \tag{5}$$

where $f(x, y)$, $f_B(x, y)$, $f_T(x, y)$, $f_N(x, y)$ and (x, y) represent the original image gray value, the background region gray value, the target region gray value, the noise region gray value, and the pixel location, respectively. The target image can be effectively extracted from the infrared data by characterizing the background, target, and noise with different features, thus enabling accurate detection. Based on this, IPI model[4] is formulated as shown in (6):

$$D = B + T + N \tag{6}$$

Here, D, B, T, and N denote the original infrared patch image, background patch image, target patch image and noise patch image, respectively. The IPI model assumes that local infrared background patches are highly correlated, modeling them as a low-rank matrix, while small targets are treated as sparse components. The methods for constructing and reconstructing the IPI model refer to [4]. The detection pipeline for infrared small targets using the IPI model, as introduced in [4], is formulated as follows:

$$\min_{B,T} \|B\|_* + \lambda\|T\|_1, \quad \text{s.t.} \|D - B - T\|_F \leq \delta \tag{7}$$

where$\|\cdot\|_*$ denotes the nuclear norm, $\|\cdot\|_1$ denotes the L_1 norm, $\|\cdot\|_F$ is the Frobenius norm, λ serves as a weighting parameter.

As mentioned earlier, when the value of $p < 1$, the L_p-norm is more effective than the L_1-norm in enforcing sparsity. It imposes a stronger penalties on non-zero elements during optimization, driving more coefficients toward zero and yielding a sparser solution. This is particularly advantageous for processing images with significant noise or complex backgrounds, as the L_p-norm better isolates small targets.

In the IPI model, the low-rank assumption can lead to significant blurring in background estimation when applied to non-smooth and uneven backgrounds. This blurring effect causes the loss of critical background details, such as edges and corners, which in turn compromises detection accuracy. To address this issue, the TV-PCP model incorporates a total variation (TV) regularization term. However, traditional TV regularization, when used in complex backgrounds or with weak targets, may lead to over-smoothing or target blurring, thereby reducing detection accuracy. To further improve image processing, we introduce the OGSTV method. By imposing overlapping group sparsity constraints on the image gradient, OGSTV effectively preserves image details and enhances the separation of targets.

Therefore, to enhance target detection, we propose we introduce a dim infrared small target detection model that combines OGSTV and L_p-norm constraints. The resulting OTLC objective function is presented in Eq. (8).

$$\begin{aligned} &\min_{B,T} \|B\|_* + \lambda_1\|T\|_p + \lambda_2 OT(T), \\ &s.t. \|D - B - T\|_F \leq \delta \end{aligned} \tag{8}$$

Here,$\mathrm{OT}(\cdot)$ denotes the OGSTV regularization term, $\|\cdot\|_p$ represents the L_p-norm, and λ_1 and λ_2 serves as weighting parameters. The infrared target image is obtained by solving Eq.(8). The target information is obtained by simple post-processing of the target image. The overview of the proposed method is presented in Fig.1.

C. Optimization Algorithm

The objective in (8) is reformulated as:

$$\begin{aligned} &\min ||B||_* + \lambda_1||T||_p + \lambda_2(OT(D^{(1)}T) + OT(D^{(2)}T)), \\ &\quad s.t.\ Z_1 = D^{(1)}T, Z_2 = D^{(2)}T, \|D - B - T\|_F \leq \delta \end{aligned} \tag{9}$$

The Lagrangian function of Eq. (9) is given by (10).

$$\begin{aligned} L_A =& ||B||_* + \lambda_1||T||_p + \lambda_2(OT(Z_1) + OT(Z_2)) \\ &+ < \mu_1, D^{(1)}T - Z_1 > + \frac{\beta}{2}\|D^{(1)}T - Z_1\|_F^2 \\ &+ < \mu_2, D^{(2)}T - Z_2 > + \frac{\beta}{2}\|D^{(2)}T - Z_2\|_F^2 \\ &+ < Y, D - B - T > + \frac{\beta}{2}\|D - B - T\|_F^2 \end{aligned} \tag{10}$$

1) For the B subproblem:

$$\begin{aligned} B^{k+1} &= \arg\min_B L_A \\ &= \arg\min_B ||B||_* + \frac{\beta}{2}||B - (D + \beta^{-1}Y^k - T^k)||_F^2 \end{aligned} \tag{11}$$

Soft thresholding [18] can be used to solved the convex optimization problem Eq. (11).

$$B^{k+1} = US_{\beta^{-1}}(\Sigma)V^T \tag{12}$$

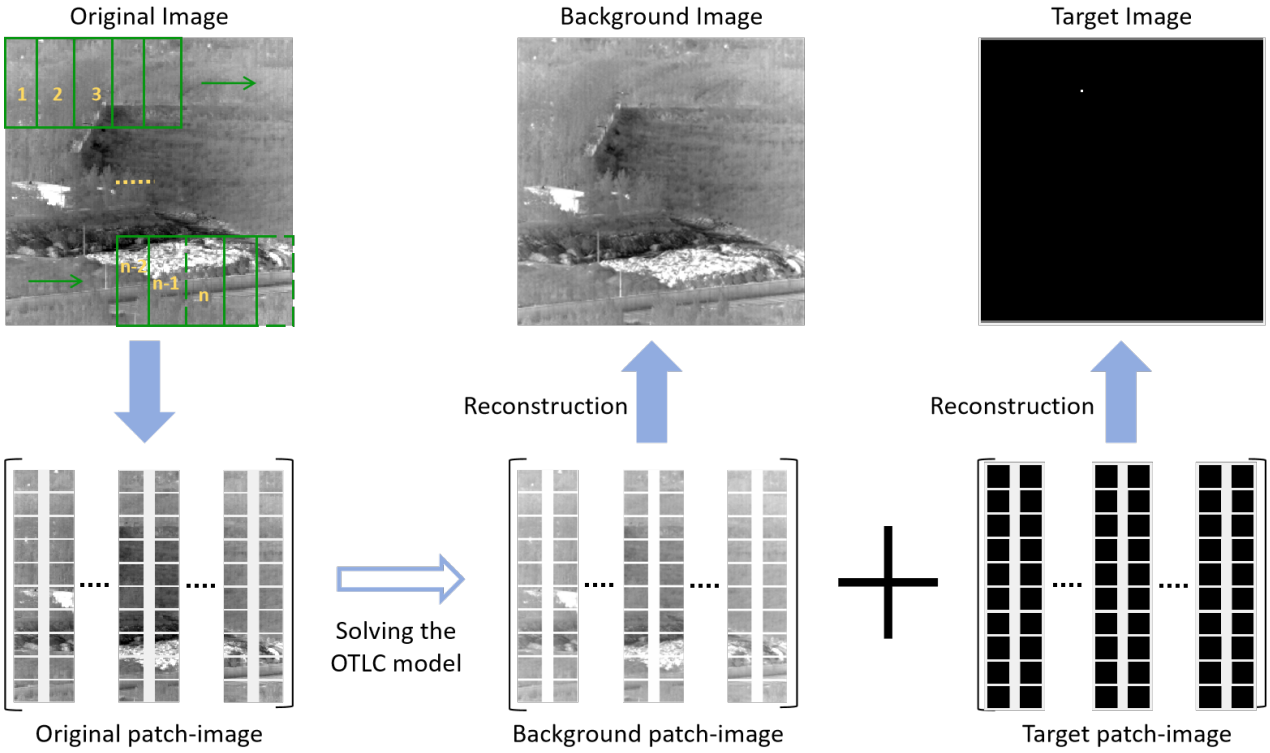


Fig. 1: Detection flow of OTLC model.

Given the singular value decomposition of Y as $Y = U\Sigma V^T$, the estimator $\widehat{x}$ is expressed as $\widehat{x} = US_\lambda(\Sigma)V^T$, and $S_\lambda(\Sigma)$ signifies the soft-thresholding operator implemented on the diagonal matrix Σ. Specifically, $S_\lambda[\bullet]$ is defined as:

$$S_\epsilon[X] = \begin{cases} \text{sign}(X)(|X| - \epsilon) & \text{if} |X| > \epsilon \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

2) For the T subproblem:

$$\begin{aligned} T^{k+1} &= \arg\min_T L_A \\ &= \arg\min_T \lambda\|T\|_p + \frac{\beta}{2}\|T - (D + \beta^{-1}Y^k - B^{k+1})\|_F^2 \end{aligned} \tag{14}$$

Due to the linear independence of the matrix elements, the problem can be solved on a per-pixel basis [19]. The optimization goal for each pixel is:

$$u^* = \min_u \frac{1}{2}(u-a)^2 + \lambda|u|^p \tag{15}$$

We define the optimization function for each pixel as $f(u)$.

$$f(u) = \frac{1}{2}(u-a)^2 + \lambda|u|^p \tag{16}$$

We take the first derivative with respect to $f(u)$.

$$f'(u) = u - a + \lambda p|u|^{p-1} \cdot sign|u| \tag{17}$$

We take the second derivative with respect to $f(u)$.

$$f''(u) = 1 + \lambda p(p-1) \cdot |u|^{p-2} \tag{18}$$

By examining the first, and second derivative of $f(u)$, we can get the minimum point of $f(u)$, which is either 0 or u_1. We define $k = \lambda p(1-p)^{\frac{1}{2-p}}$ and $k_1 = k + \lambda p|k|^{p-1}$. The solution to Eq. (15) can be written as follows:

$$u^* = \begin{cases} 0 & a \le k_1 \\ \arg\min_{u\in\{0,u_1\}} f(u) & a \ge k_1 \end{cases} \tag{19}$$

The solution u_1 of $f'(u) = 0$ within the interval $k < u < a$ can be obtained using the Newton iteration method, and the iterative formula is presented in Eq. (20).

$$U_{n+1} = U_n - \frac{f'(u)}{f''(u)} \tag{20}$$

To solve the matrix formulation of (15), we introduce the operator $Q_\lambda[\bullet]$.

$$Q_{\lambda,p}[M] = \arg\min_U \lambda\|U\|_p + \frac{1}{2}\|U - M\|_F^2 \tag{21}$$

An optimal solution can be obtained through Eq. (19). Thus, problem (14) is solved through the definition of $Q_\lambda[\bullet]$.

$$T^{k+1} = Q_{\beta^{-1}\lambda_1,p}\left[D + \beta^{-1}Y^k - B^{k+1}\right] \tag{22}$$

3) For the Z_1 and Z_2 subproblem:

$$\begin{aligned} Z_1^{k+1} &= \arg\min_{Z_1} L_A \\ &= \arg\min \frac{\beta}{2}||Z_1 - (D^{(1)}T^{k+1} + \beta^{-1}\mu_1^k)||_F^2 \\ &+ \lambda_2 OT(Z_1) \end{aligned} \tag{23}$$

$$\begin{aligned} Z_2^{k+1} &= \arg\min_{Z_2} L_A \\ &= \arg\min \frac{\beta}{2}\|Z_2 - (D^{(2)}T^{k+1} + \beta^{-1}\mu_2^k)\|_F^2 \\ &+ \lambda_2 OT(Z_2) \end{aligned} \tag{24}$$

Referring to [20], we can use MM algorithm to solve the problem (22) and (23). Refer to [17] for the detailed process of solving related problems. This can be formalized in Lemma 1 as follows:

Lemma 1: Let us consider the following minimization problem: $\min_\nu R(\nu) = \frac{\beta}{2}||\nu - \nu_0||_F^2 + \lambda\Phi(\nu)$, where β

and λ are positive parameters and the functional $\Phi(v) = \sum_{h,f} ||a(h,f)_C||_2$. Applying the MM algorithm to minimize $R(\nu)$, the estimate is updated at each iteration by

$$\nu^{k+1} = (I + \frac{\lambda}{\beta}\Lambda(\nu^k)^T\Lambda(\nu^k))^{-1}\nu_0 \tag{25}$$

The specific process of solving problem (10) combined with ADMM algorithm is shown in Algorithm 1.

Algorithm 1 Solving algorithm to OTLC model

Input: D, λ_1, λ_2, p
Output: T and B

Initialization parameters: $B^0 = D$, $T^0 = 0$, $Z^0 = 0$, $Y^0 = 0$, $\mu_1^0 = 0$, $\mu_2^0 = 0$, $\beta^0 = \frac{1}{5*\text{std}(D)}$
while not converged **do**
  Update B^{k+1} by solving $B^{k+1} = \arg\min_B L_A$
  $B^{k+1} = US_{\beta^{-1}}(\Sigma)V^T$
  Update T^{k+1} by solving $T^{k+1} = \arg\min_T L_A$
  $T^{k+1} = Q_{\beta^{-1}\lambda_1,p}\left[D + \beta^{-1}Y^k - B^{k+1}\right]$
  Compute Z_1^{k+1} and Z_2^{k+1} according to Eq.(25)
  Update Y^{k+1} and β^{k+1}
  $Y^{k+1} = Y^k + \beta^k(D - B^{k+1} - T^{k+1})$,
$\beta^{k+1} = 1.5 * \beta^k$
  Update μ_1^{k+1} and μ_2^{k+1}
  $\mu_1^{k+1} = \mu_1^k + \beta^k(D^{(1)}T^{k+1} - Z_1^{k+1})$,
$\mu_2^{k+1} = \mu_2^k + \beta^k(D^{(2)}T^{k+1} - Z_2^{k+1})$
  Judge whether it has converged
  stopCriterion $= \frac{\|D-B^{k+1}-T^{k+1}\|_F}{\|D\|_F}$
  if stopCriterion $< 10^{-7}$ **then**
    converged and stop iteration
  end if
end while

To validate the proposed model, experiments are carried out with different parameter settings and compared with other methods.

III. Results and Discussion

The experimental configuration is described in this section, along with an analysis of parameter sensitivity in relation to detection performance. A series of experiments are performed to demonstrate the robustness and accuracy of the OTLC method, especially under cluttered backgrounds and noise.

A. Experimental Setting

Patch size is a critical parameter that significantly affects detection performance, as its optimization can enhance target sparsity. However, smaller patch sizes often lead to higher computational costs. In our experiments, patch sizes of 30, 40, and 50 are evaluated. Fig. 2 presents the 3D visualizations associated with these patch sizes. The results indicate that a patch size of 30 provides an optimal balance between detection accuracy and computational efficiency.

Similarly, we further adjust the step size to improve performance. With the patch size fixed at 30 × 30, step sizes of 5, 10, and 15 are investigated. The 3D visualizations corresponding to these step sizes are shown in Fig. 3, which indicate that a step size of 10 yields the optimal result.

In the proposed algorithm, the parameters are configured as follows: a patch size of 30 × 30, a sliding step of 10, $L = 1$, and $p = 0.3$. The regularization parameter λ_1 is defined as $L/\sqrt{\max(\text{size}(D))}$, while λ_2 is empirically determined and generally set around 0.01. In our experiments, we set $\lambda_2 = 0.04$. To evaluate detection and tracking performance, six real-world infrared sequences featuring dim and small aircraft targets in diverse complex backgrounds are selected as datasets.

B. Quantitative Evaluation

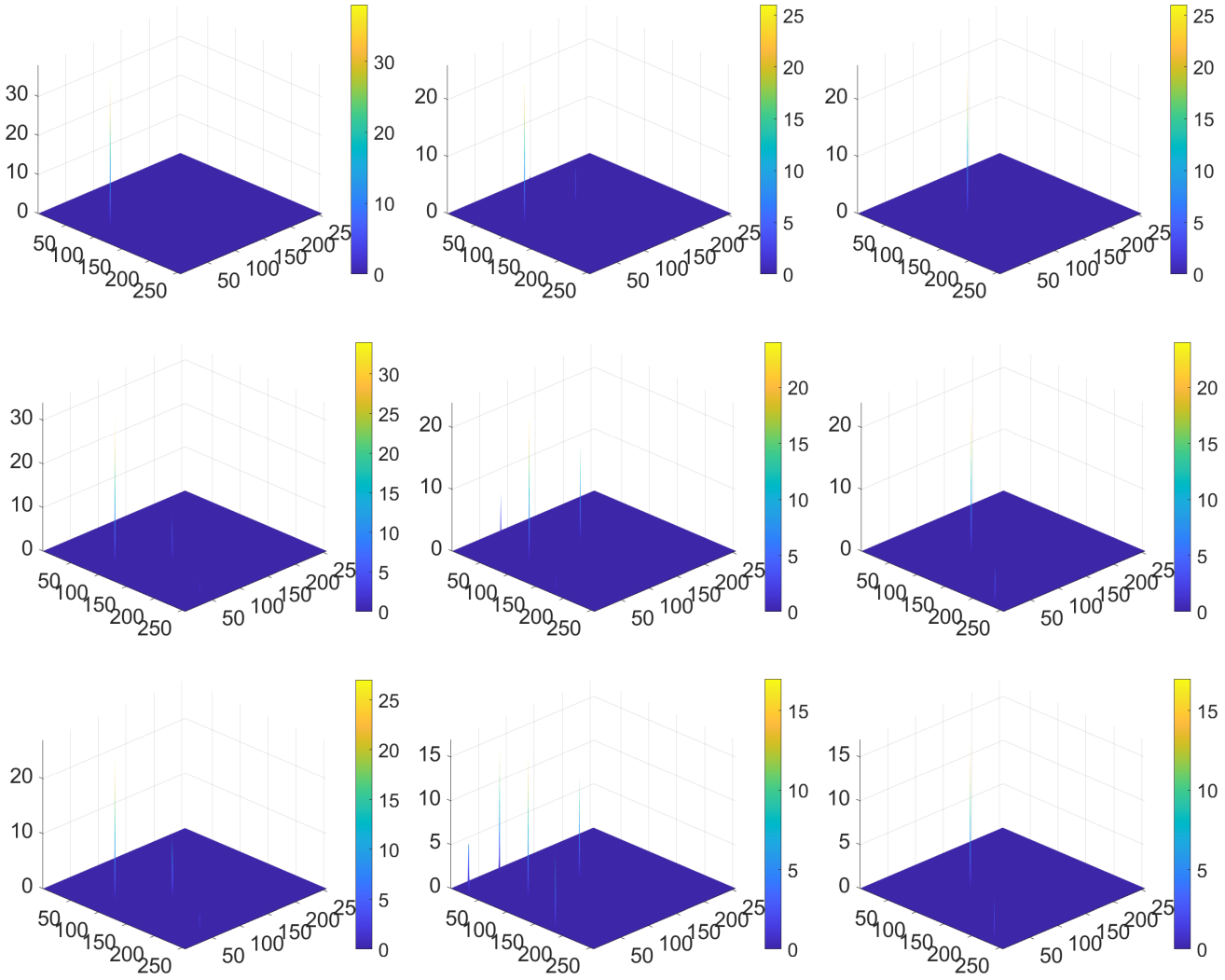
To demonstrate the proposed model outperforms nine baseline approaches under noisy and complex conditions, we perform comparative experiments. These include the Maxmedian method [21], the local contrast measure (LCM) method [22], the Tophat model [23], the IPI model, the TV-PCP model, the NIPPS model, the NOLC model, the nonconvex tensor low-rank approximation model (ASTTV-NTLA) [24], and the 4-D tensor ring (4-D TR) model [25]. The parameter configurations for all compared algorithms are summarized in Table I.

Fig. 4 shows the detection results of six sequences of single images across ten different methods. The experimental outcomes suggest that the OTLC approach achieves the effective detection performance in infrared small target scenarios involving noise and scene complexity. In Seq1 to Seq6, OTLC effectively identifies the target while significantly reducing background interference. Even under substantial noise, OTLC maintains accurate target identification.

Fig. 5–8 presents a 3D display comparing various infrared small target detection algorithms across different sequences. In some sequences, the Maxmedian and Tophat methods fail to effectively extract the target signal, resulting in signal blurring or confusion with the background. Although the LCM, IPI, and TV-PCP methods suppress background signals to some extent in certain sequences, residual background signals remain prominent. While the NIPPS, NOLC, and ASTTV-NTLA methods demonstrate a strong ability to suppress noise, background fluctuations persist in multiple sequences, indicating that noise removal is incomplete. In Seq2, the 4-D TR method exhibits a slight advantage over OTLC; nonetheless, OTLC outperforms it when considering overall detection performance. In Seq1 to Seq6, OTLC successfully detects weak targets and exhibits superior background noise suppression capabilities.

To highlight the robustness and effectiveness of the OTLC method, we introduce zero-mean Gaussian noise into Seq6 and compare the performance of OTLC with the ASTTV-NTLA and 4-D TR methods. Fig. 9 shows the results for noise variances of 0.001. The OTLC method consistently demonstrates stable detection performance even at increased noise levels and effectively suppressing background noise.

Similarly, to further evaluate the robustness of OTLC under different noise conditions, we introduce Poisson noise into Seq5 and compare its performance with the ASTTV-NTLA and 4-D TR methods. Fig. 10 presents the detection results under Poisson noise conditions. Unlike Gaussian noise, which maintains a constant variance across all pixel intensities, Poisson noise exhibits signal-dependent



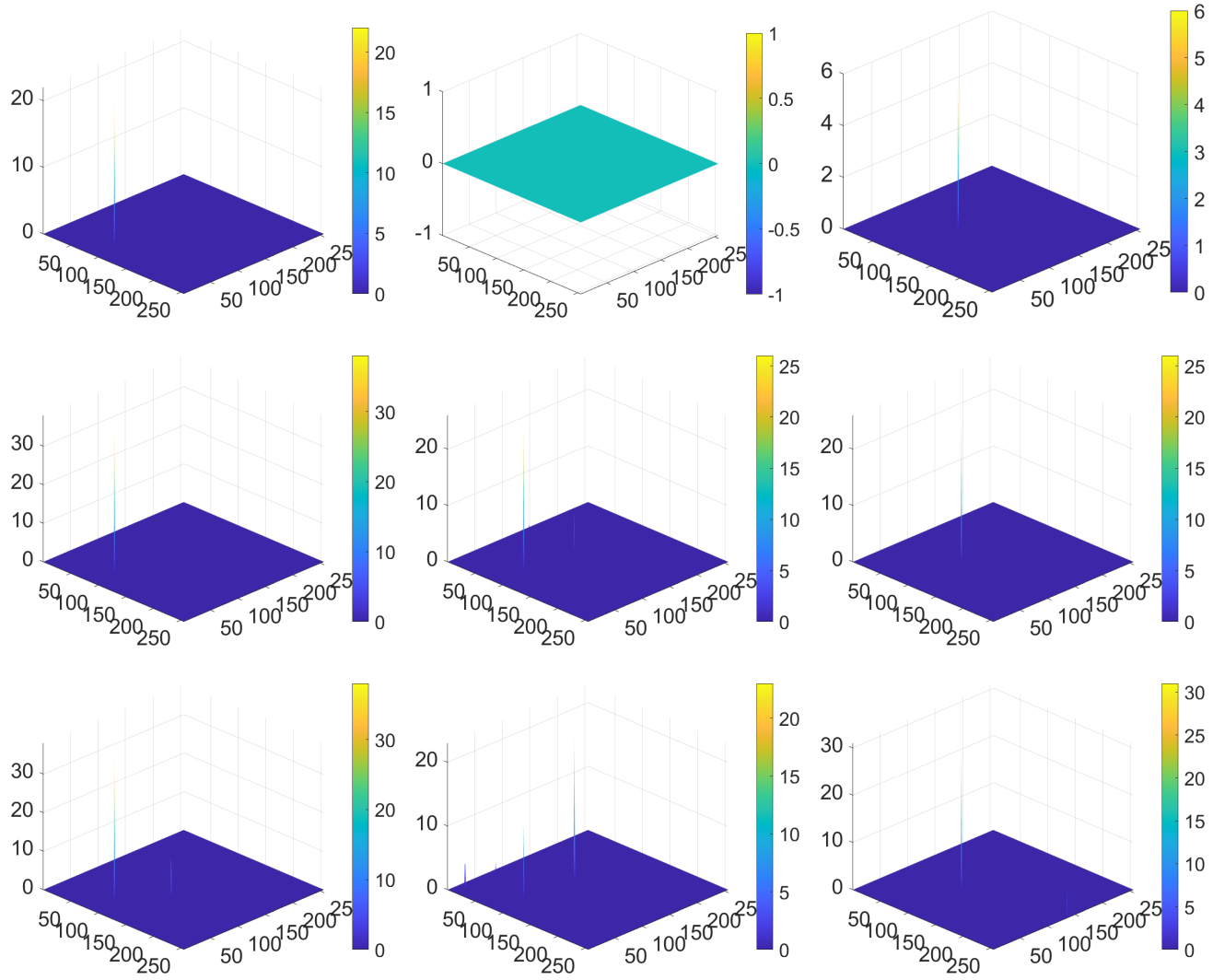
(a) Seq1 (b) Seq2 (c) Seq3

Fig. 2: 3D displays of Seq1 to Seq3 with patch sizes of 30×30, 40×40, and 50×50, shown in the first to third rows.

TABLE I: Baseline methods and primary parameter settings

Methods	**Primary parameter settings**
Maxmedian	Support size: 5×5
Tophat	Structure size: 5×5
LCM	Window radius: 1, 2, 3, 4
IPI	Patch size: 50×50; Sliding step: 10;$\lambda = L/\sqrt{\max(m,n)}$; $L = 1$
TV-PCP	$\lambda_1 = 0.005$;$\lambda_2 = \frac{L}{\sqrt{\min(m,n)}}$; $L = 1$; $\beta = 0.025$; $\gamma = 1.5$
NIPPS	Patch size: 50×50; Sliding step: 10; $\lambda = L/\sqrt{\min(m,n)}$; $L = 2$;
NOLC	Patch size: 30×30; Sliding step: 10; $\lambda = L/\sqrt{\max(m,n)}$; $L = 1$; $p = 0.4$
ASTTV-NTLA	L=3; H=6;$\lambda_{tv} = 0.005$;$\lambda_s = \frac{H}{\sqrt{\max(m,n)*L}}$; $\lambda_3 = 100$
4-D TR	Patch size: $N_1 \times N_2 = 70 \times 70$; Temporal size: $N_3 = 15$; $L = 2$; $\lambda_1 = \sum_{i=1}^{l} \frac{L}{\sqrt{\max\left(\prod_{i=n}^{n+l-1} N_i, \prod_{i=n+l}^{n-1} N_i\right)}}$

characteristics, whereby higher intensity regions experience stronger noise fluctuations while lower intensity regions



(a) Seq1 (b) Seq2 (c) Seq3

Fig. 3: 3D displays of Seq1 to Seq3 with sliding steps of 5×5, 10×10, and 15×15, shown in the first to third rows.

remain relatively unaffected. In contrast, the OTLC method consistently shows stable detection performance and effectively suppresses noise. These experiments further highlight the superior detection capabilities of the OTLC method.

C. Quantitative Evaluation

The effectiveness is further verified through quantitative experiments using two evaluation metrics: the signal-to-noise ratio gain (SNRG) and the background suppression factor (BSF).

SNRG is a metric used to evaluate the improvement in the signal-to-noise ratios (SNR) after applying a processing algorithm, particularly in signal detection tasks such as infrared small target detection. Higher SNRG values imply stronger noise suppression and better target signal retention, which contribute to enhanced detection results. SNRG is defined as

$$SNRG = (SNR_{out}/SNR_{in}) \tag{26}$$

where SNR_{in} and SNR_{out} respectively denote the SNR of the original infrared image and the processed image.

The BSF serves as an important criterion for measuring the effectiveness of background interference suppression in target detection algorithms. It evaluates the efficacy of background suppression by contrasting the signal intensities of the background and target areas. A higher BSF value is associated with stronger background signal suppression by the algorithm. BSF is defined as

$$BSF = C_{in}/C_{out} \tag{27}$$

where C_{in} and C_{out} denote the background standard deviation in the input and output images, respectively.

Based on the above definitions of SNRG and BSF, we carried out experiments on various sequences, the corresponding results are detailed in Table II. Experimental results reveal that the OTLC method consistently performs excellently in both the SNRG and the BSF across different sequences. Across all six sequences, the OTLC algorithm yields the highest mean BSF values. Although OTLC falls slightly behind 4-D TR in terms of average SNRG for Seq 2, it maintains leading performance across the rest of the sequences. This indicates that OTLC effectively enhances the SNR while exhibiting excellent background suppression

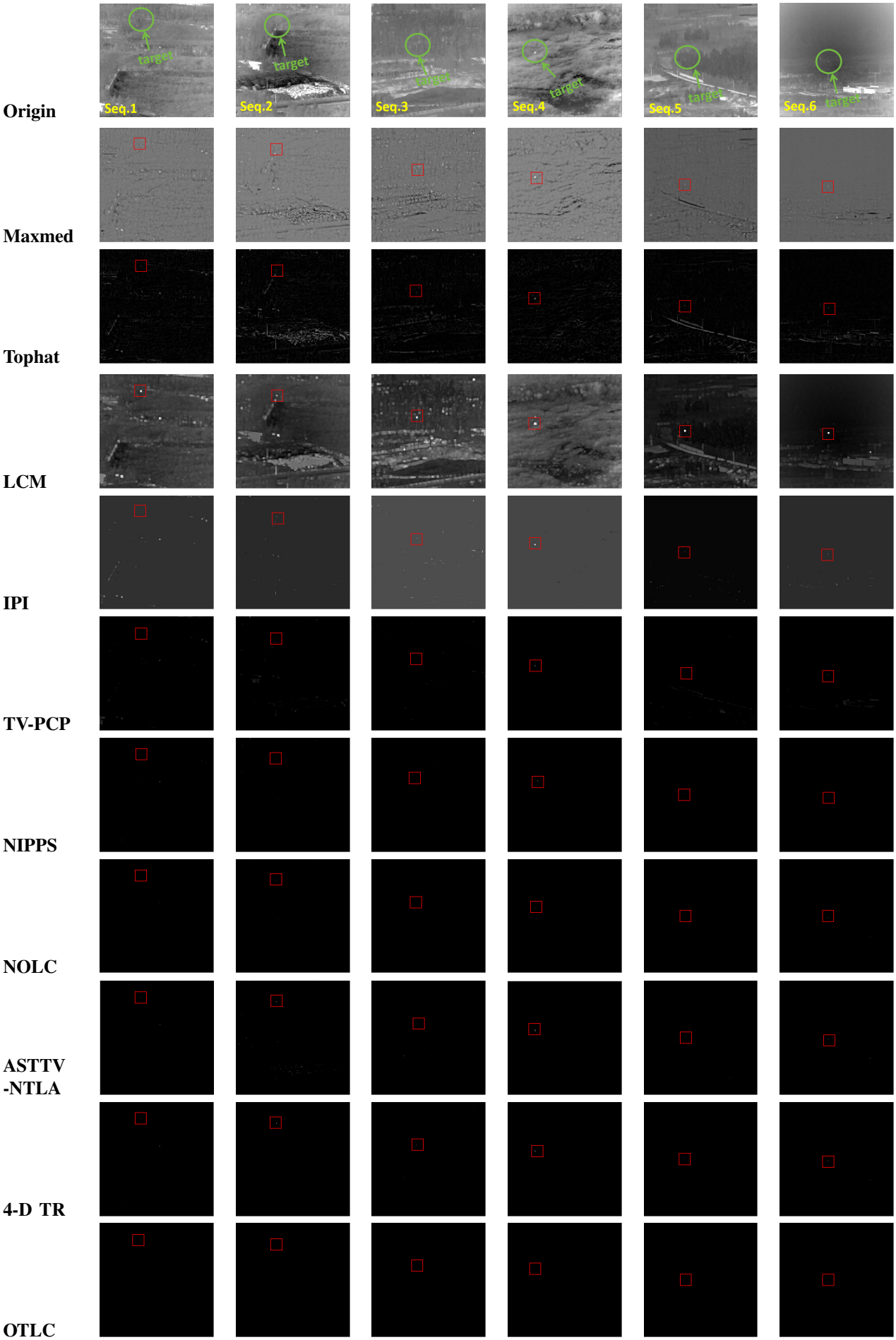
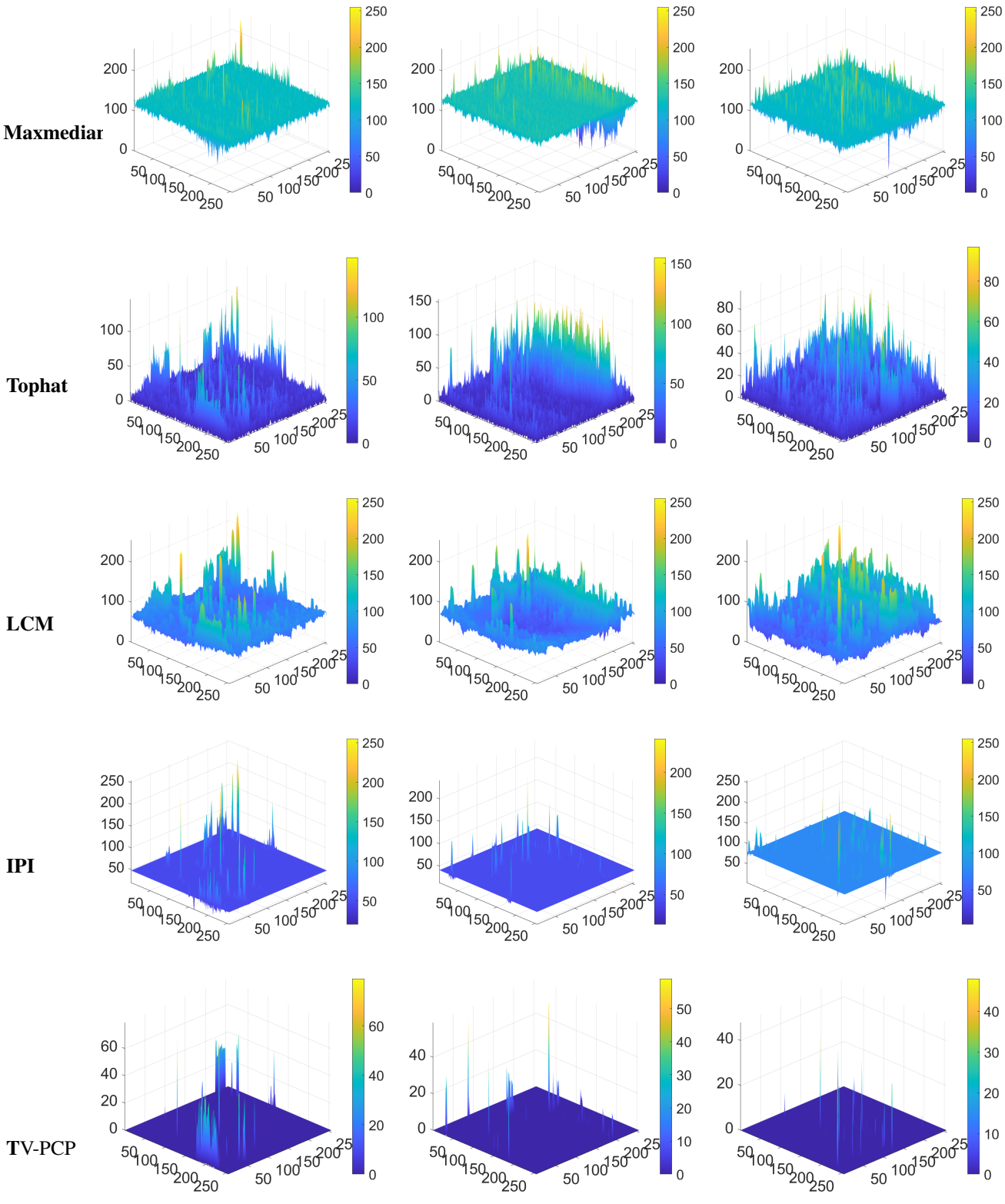


Fig. 4: Detection results of the Seq1 to Seq6 for multiple methods.



(a) Seq1 (b) Seq2 (c) Seq3

Fig. 5: 3D display of the Seq1 to Seq3 for multiple methods: Maxmedian, Tophat, LCM, IPI, and TV-PCP.

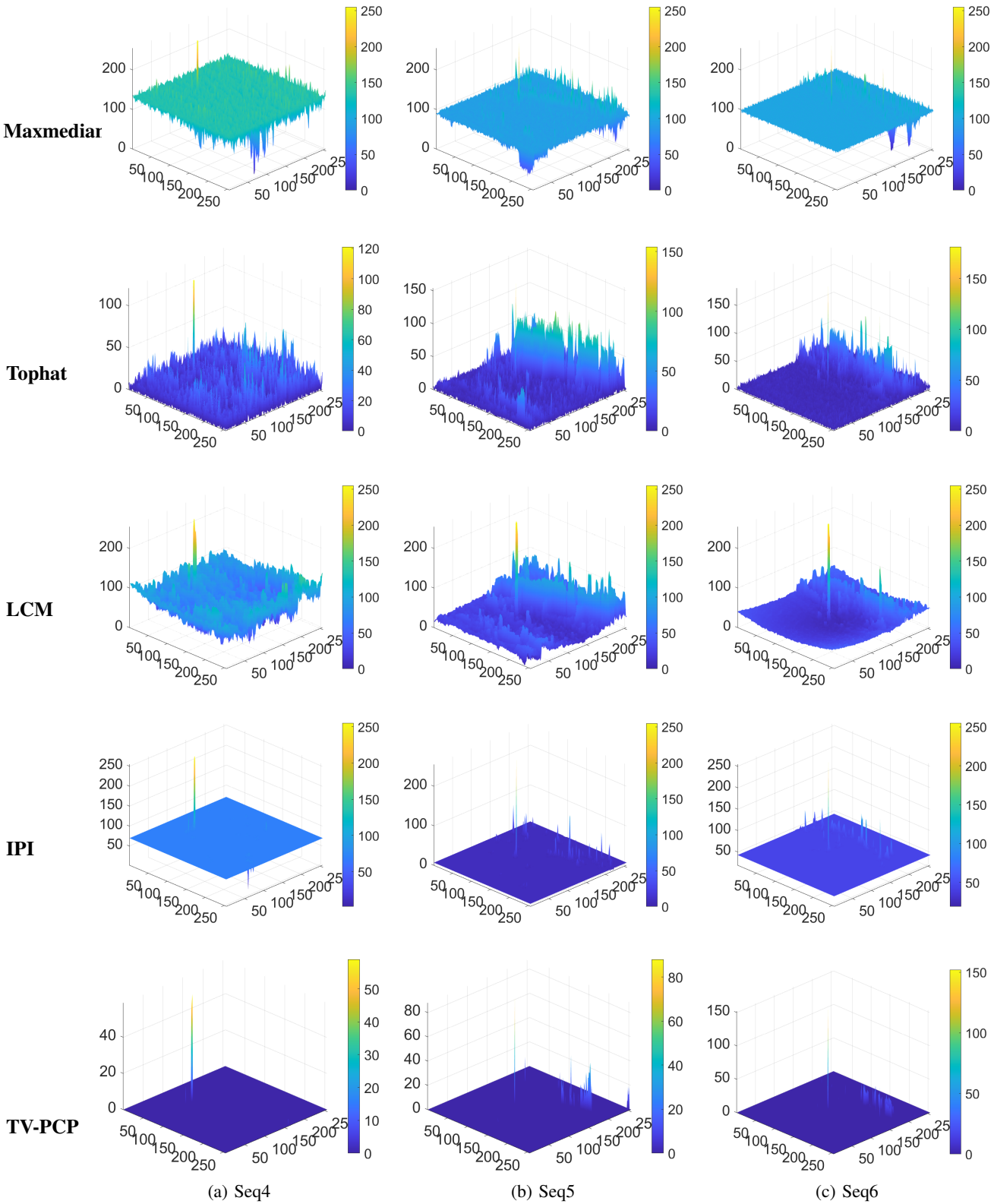
capabilities.

To evaluate the detection accuracy of the OTLC algorithm, we draw the receiver operating characteristic (ROC) curve with the true positive rate (TPR) and the false positive rate (FPR).

$$\mathrm{TPR} = \frac{\text{number of detected true targets}}{\text{total number of actual targets}} \tag{28}$$

$$\mathrm{FPR} = \frac{\text{number of detected false pixels}}{\text{total number of pixels in images}} \tag{29}$$

ROC curves are commonly employed to assess the performance of infrared small target detection algorithms. Its horizontal axis represents the TPR, indicating the fraction of actual targets correctly identified, while the vertical axis corresponds to the FPR, reflecting the fraction of background elements misclassified as targets. Consequently, an ROC

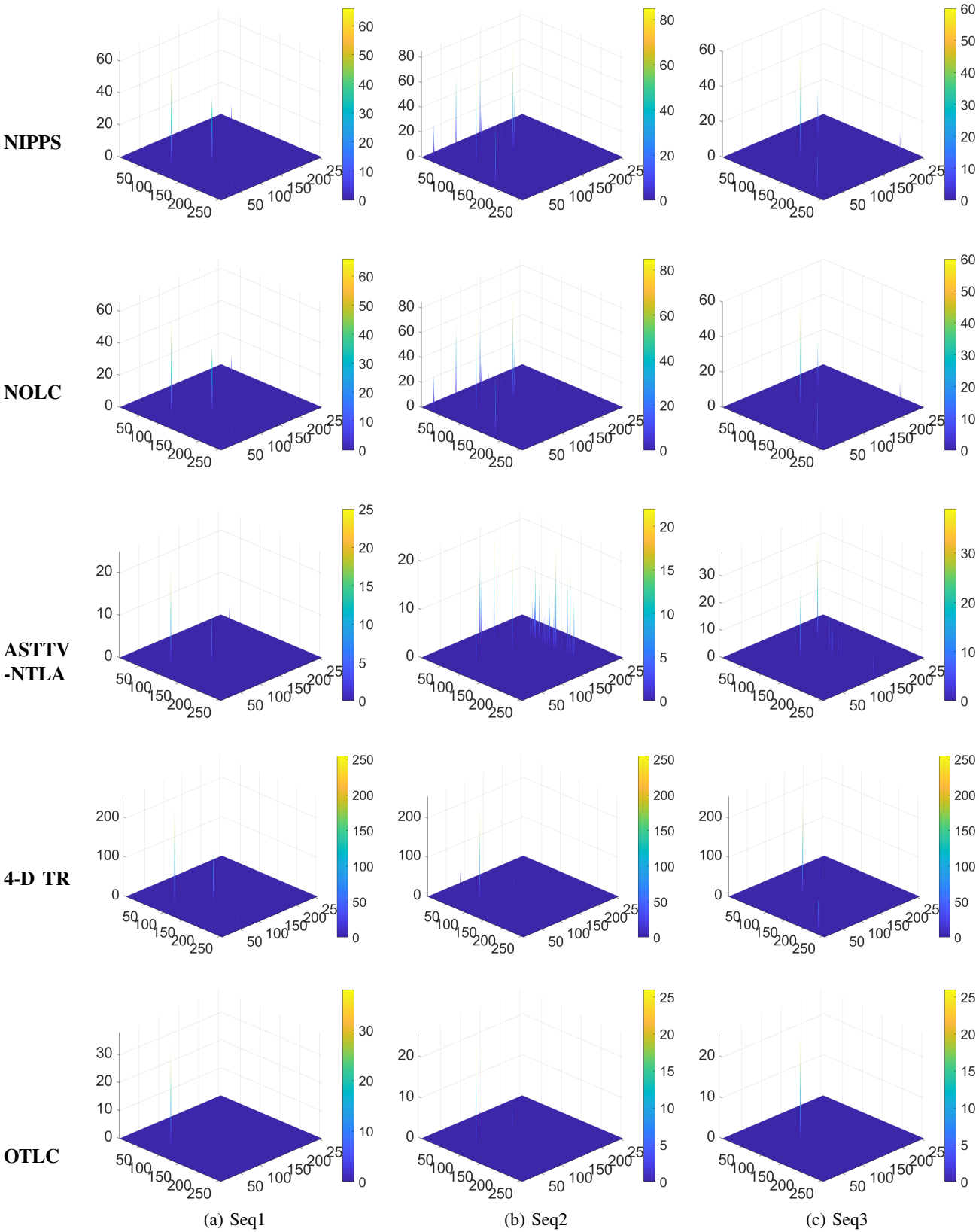


(a) Seq4 (b) Seq5 (c) Seq6

Fig. 6: 3D display of the Seq4 to Seq6 for multiple methods: Maxmedian, Tophat, LCM, IPI, and TV-PCP.

curve that near the top-left corner indicates superior detection capability. Additionally, the area under the ROC curve (AUC) is a key indicator: a larger AUC value generally signifies more effective target detection.

Fig. 11 displays the ROC curves for the six sequences using ten different methods. The ROC curves for the OTLC model cluster near the top-left corner, demonstrating that a higher TPR is achieved for a given FPR. Despite its slightly inferior AUC in sequence 1 compared to NIPPS, the OTLC method achieves superior AUC performance across sequences 2 to 6. Overall, the OTLC method consistently maintains a high detection rate across different sequences.



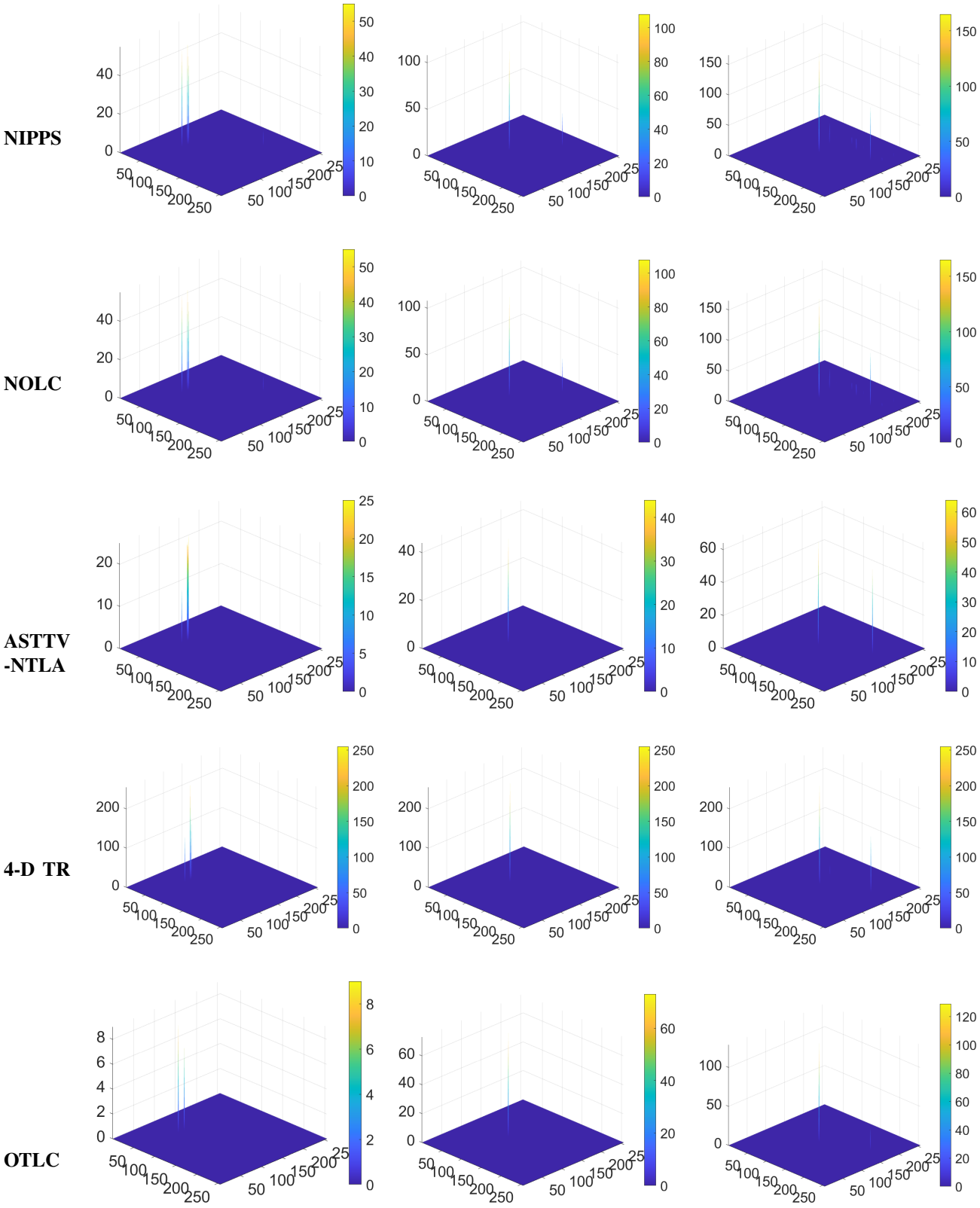
(a) Seq1 (b) Seq2 (c) Seq3

Fig. 7: 3D display of the Seq1 to Seq3 for multiple methods: NIPPS, NOLC, ASTTV-NTLA, 4D-TR, and OTLC.

D. Sensitivity Analysis

To substantiate the robustness and stability of our approach, we examined the sensitivity of the model with respect to λ_2 and p, investigating the effects of different parameter settings on SNRG and BSF. We adjusted λ_2 and p, and recorded the corresponding SNRG and BSF values for each parameter configuration. The parameter λ_2 controls the degree of target sparsity and background suppression, while p determines the extent to which the L_p-norm constraint enforces sparsity.

During the experiments, all other parameters remained

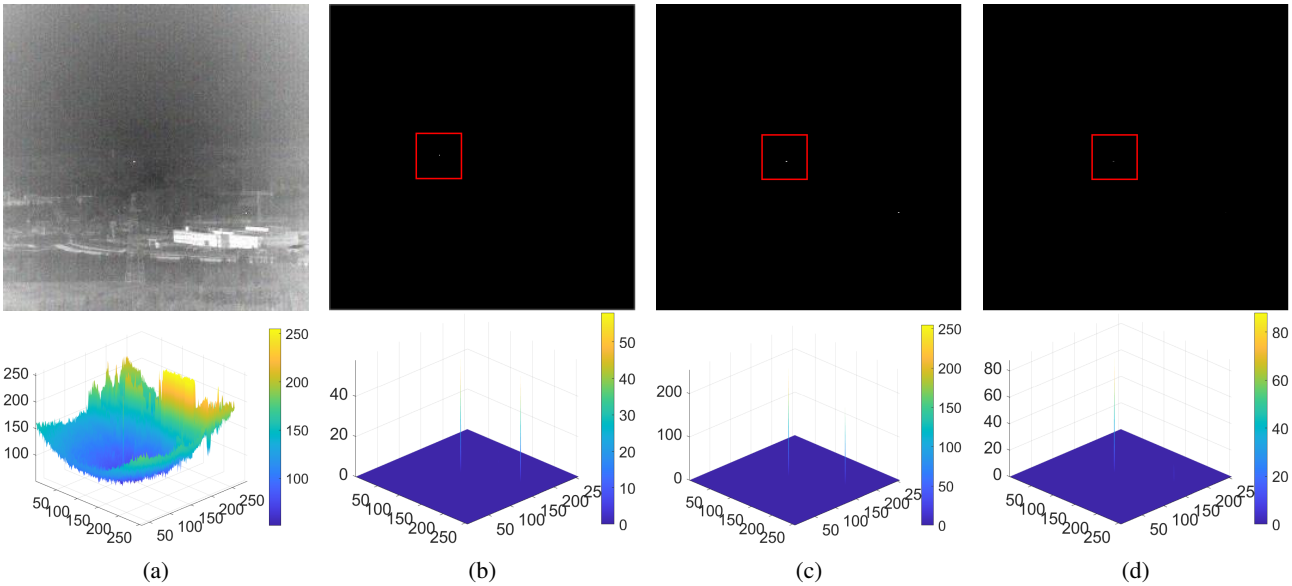


(a) Seq4 (b) Seq5 (c) Seq6

Fig. 8: 3D display of the Seq1 to Seq3 for multiple methods: NIPPS, NOLC, ASTTV-NTLA, 4D-TR, and OTLC.

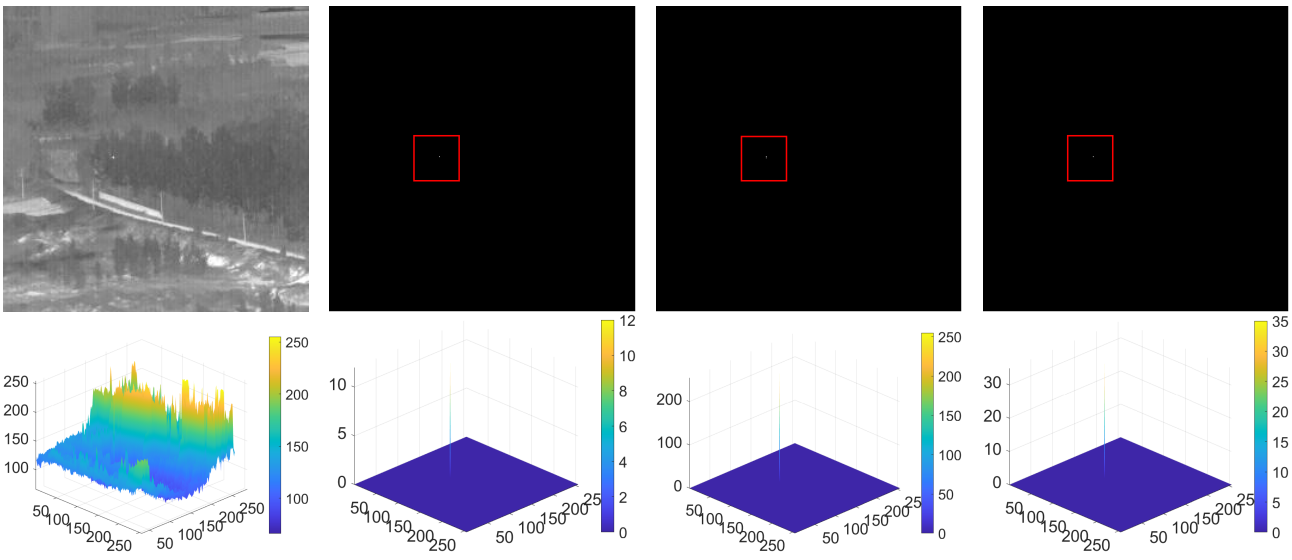
fixed while p was assigned values of 0.1, 0.2, 0.3, 0.4, and 0.5. Fig. 12 presents the variations in SNRG and BSF with different p values. The experimental results indicate that when p is set to 0.3, the algorithm achieves the best balance between detecting target and suppressing background across every test sequences.

Similarly, we set λ_2 to 0.01, 0.02, 0.03, and 0.04, respectively. Excessively large values of λ_2 may lead to the loss of target signals. Fig. 13 illustrates the variations in SNRG and BSF with different λ_2 values. It is observed that setting λ_2 to 0.04 yields optimal target detection and stronger suppression of background noise.



(a) (b) (c) (d)

Fig. 9: Comparison for Gaussian noise images in Seq6 (variance = 0.001): (a) noisy image; (b) ASTTV-NTLA result; (c) 4-D TR result; (d) OTLC result.



(a) (b) (c) (d)

Fig. 10: Comparison for Poisson noise images in Seq5: (a) original Poisson noise image; (b) ASTTV-NTLA result; (c) 4-D TR result; (d) OTLC result.

TABLE II: The average SNRG and BSF of six sequences.

Methods	Seq 1		Seq 2		Seq 3		Seq 4		Seq 5		Seq 6	
	SNRG	**BSF**	**SNRG**	**BSF**	**SNRG**	**BSF**	**SNRG**	**BSF**	**SNRG**	**BSF**	**SNRG**	**BSF**
Maxmedian	2.634	2.442	3.460	3.369	2.110	1.885	2.719	3.736	4.032	3.109	7.936	6.383
Tophat	2.937	2.659	3.279	2.862	2.076	2.890	4.072	4.502	2.981	2.564	6.633	4.804
LCM	1.812	1.205	2.791	1.837	0.793	1.193	1.623	2.303	2.409	1.358	4.300	2.406
IPI	7.470	4.610	26.149	16.732	9.719	6.969	21.583	14.888	24.240	12.473	31.615	19.052
TV-PCP	5.637	7.666	12.333	15.660	9.305	16.974	23.362	37.473	12.040	16.974	20.780	16.398
NIPPS	18.393	24.708	28.005	55.995	15.923	54.907	30.953	53.382	42.022	47.443	56.064	45.610
NOLC	30.093	58.176	39.542	59.520	33.189	70.551	37.323	86.602	44.830	52.952	53.163	41.152
ASTTV-NTLA	18.124	50.283	12.130	38.867	28.898	94.512	26.828	136.955	33.838	73.171	35.880	58.760
4-D TR	34.707	17.365	75.382	37.819	36.587	18.300	47.728	23.885	47.416	23.721	53.101	26.596
OTLC (ours)	39.634	133.075	74.642	367.278	41.149	201.858	47.881	678.928	47.412	82.853	62.619	61.997

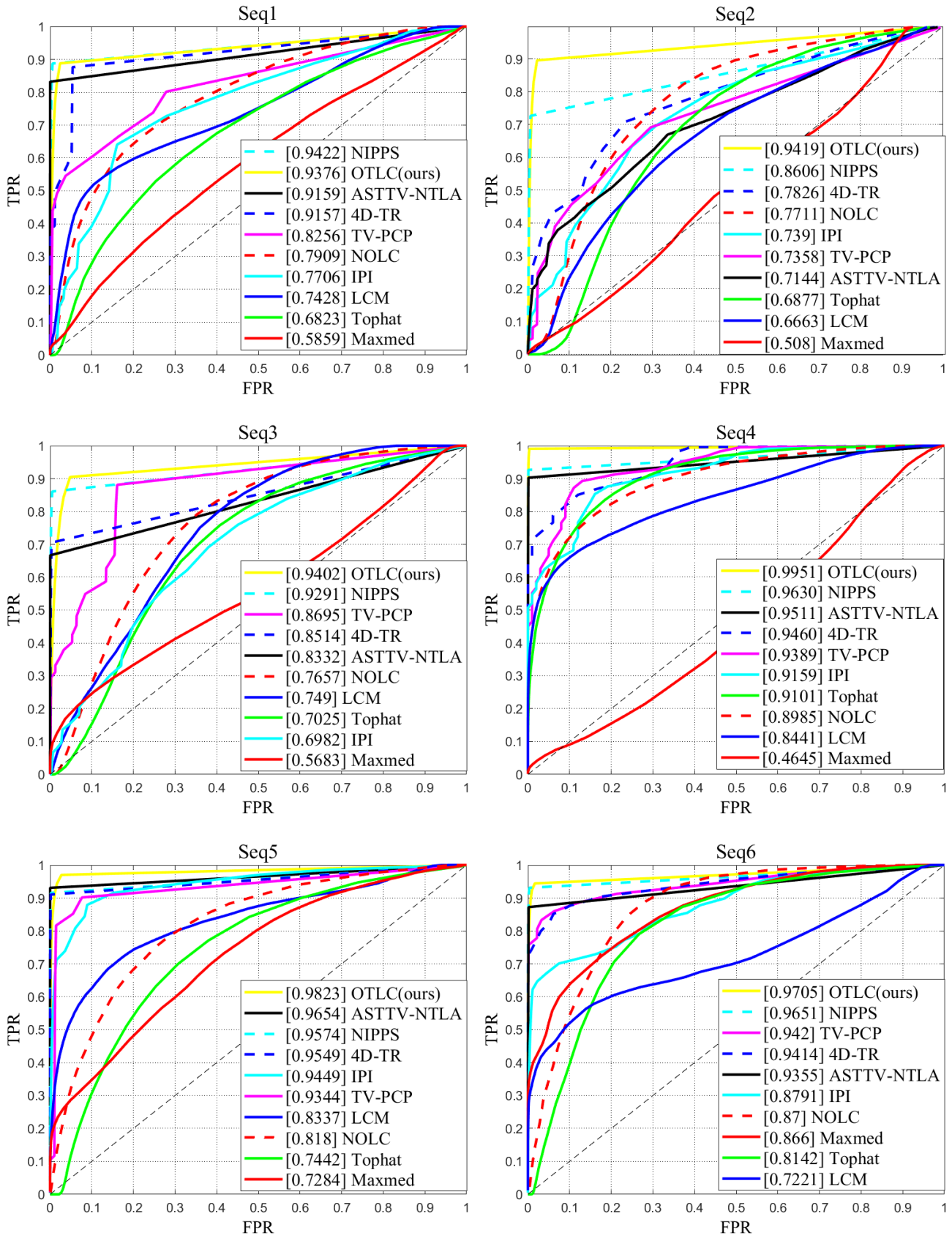


Fig. 11: ROC curves illustrating the detection results for six sequences.

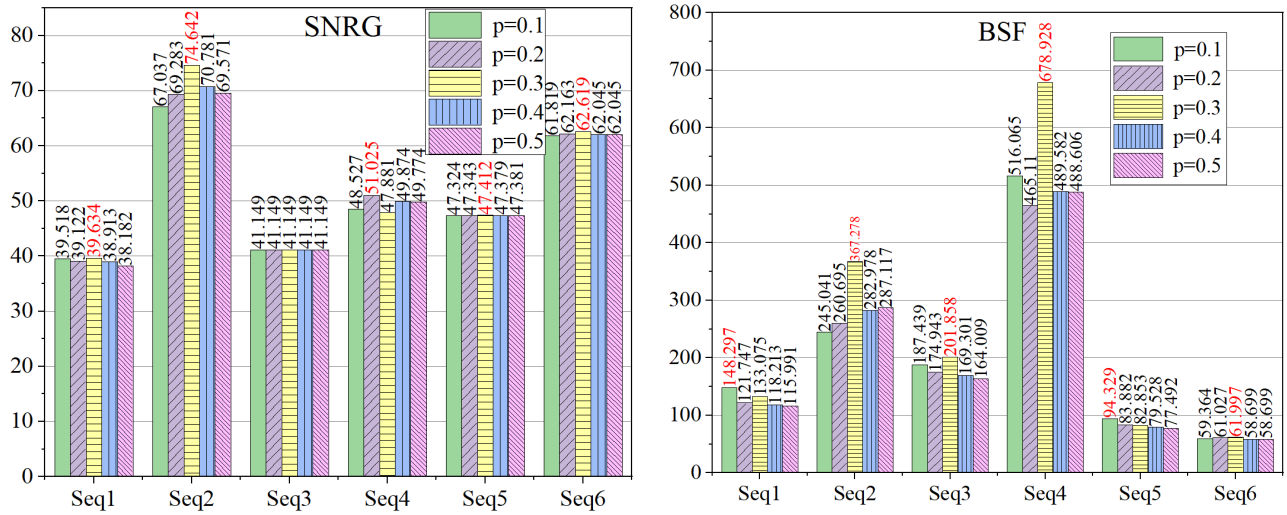


Fig. 12: SNRG and BSF values corresponding to different p values for six sequences.

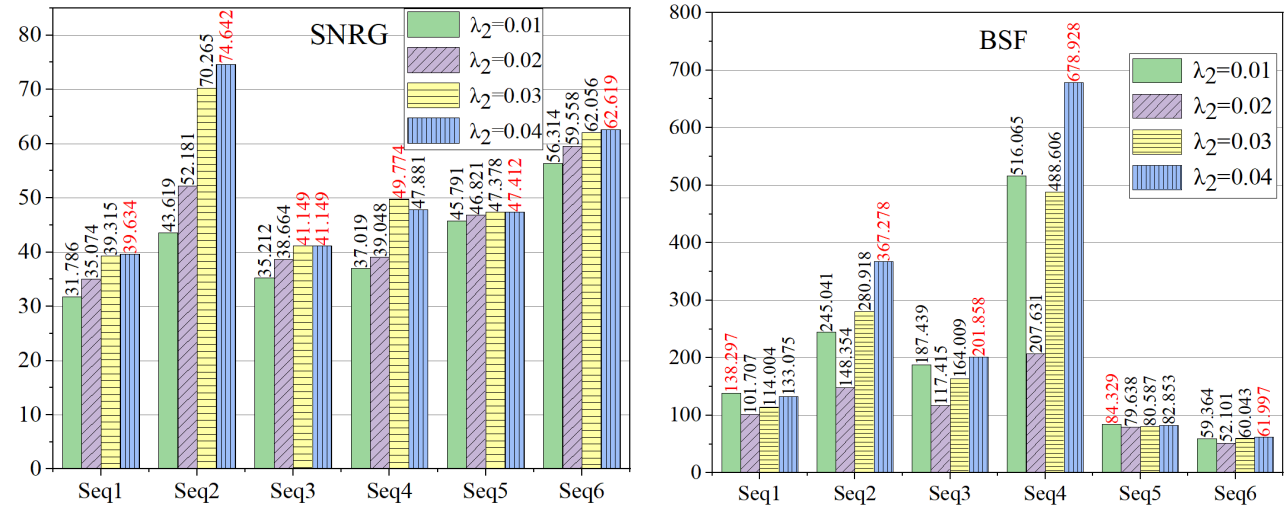


Fig. 13: SNRG and BSF values corresponding to different λ_2 values for six sequences.

IV. Conclusions

This paper introduces a new model for detecting infrared dim small targets, with its optimization solved through ADMM. The model combines OGSTV with L_p-norm regularization. Experiments confirm that the OTLC model obtains superior detection performance while effectively suppressing background noise.

References

[1] X. Shao, H. Fan, G. Lu, and J. Xu, "An improved infrared dim and small target detection algorithm based on the contrast mechanism of human visual system," *Infrared Physics & Technology*, vol. 55, no. 5, pp. 403–408, 2012.

[2] Y. Chen and Y. Xin, "An efficient infrared small target detection method based on visual contrast mechanism," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 7, pp. 962–966, 2016.

[3] H. Shahraki, S. Moradi, and S. Aalaei, "A noise-robust method for infrared small target detection," *Signal, Image and Video Processing*, vol. 17, pp. 2489–2497, 2023.

[4] 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 Transactions on Image Processing*, vol. 22, no. 12, pp. 4996–5009, 2013.

[5] Y. Dai, Y. Wu, and Y. Song, "Infrared small target and background separation via column-wise weighted robust principal component analysis," *Infrared Physics & Technology*, vol. 77, pp. 421–430, 2016.

[6] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 349–366, 2007.

[7] Y. Dai and Y. Wu, "Reweighted infrared patch-tensor model with both nonlocal and local priors for single-frame small target detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 8, pp. 3752–3767, 2017.

[8] Y. Dai, Y. Wu, Y. Song, and J. Guo, "Non-negative infrared patch-image model: Robust target-background separation via partial sum minimization of singular values," *Infrared Physics & Technology*, vol. 81, pp. 182–194, 2017.

[9] X. Wang, Z. Peng, D. Kong, P. Zhang, and Y. He, "Infrared dim target detection based on total variation regularization and principal component pursuit," *Image and Vision Computing*, vol. 63, pp. 1–9, 2017.

[10] D. Kong and Z. Peng, "Seismic random noise attenuation using shearlet and total generalized variation," *Journal of Geophysics and Engineering*, vol. 12, no. 6, pp. 1024–1035, 2015.

[11] T. Zhang, H. Wu, Y. Liu, L. Peng, C. Yang, and Z. Peng, "Infrared small target detection based on non-convex optimization with lp-norm constraint," *Remote Sensing*, vol. 11, article number 559, 2019.

[12] S. S. Rawat, S. K. Verma, and Y. Kumar, "Infrared small target detection based on bi-nuclear norm minimization," *Multimedia Tools and Applications*, vol. 83, pp. 34 111–34 135, 2024.

[13] R. Chartrand, "Exact reconstruction of sparse signals via nonconvex minimization," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707–710, 2007.

[14] R. Chartrand and V. Staneva, "Restricted isometry properties and nonconvex compressive sensing," *Inverse Problems*, vol. 24, no. 3, p. 035020, 2008.

[15] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for

compressive sensing," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3869–3872.
[16] M. Elad, *Sparse and Redundant Representations*. Springer New York, 2010.
[17] J. Liu, T. Z. Huang, I. W. Selesnick, X.-G. Lv, and P.-Y. Chen, "Image restoration using total variation with overlapping group sparsity," *Information Sciences*, vol. 295, pp. 232–246, 2015.
[18] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, pp. 1956–1982, 2008.
[19] F. Nie, H. Wang, H. Huang, and C. Ding, "Joint schatten p-norm and l_p-norm robust matrix completion for missing value recovery," *Knowledge and Information Systems*, vol. 42, no. 3, pp. 525–544, 2015.
[20] P. Y. Chen and I. W. Selesnick, "Translation-invariant shrinkage/thresholding of group sparse signals," *Signal Processing*, vol. 94, pp. 476–489, 2014.
[21] S. D. Deshpande, M. H. Er, R. Venkateswarlu, and P. Chan, "Max-mean and max-median filters for detection of small targets," in *Signal and Data Processing of Small Targets 1999*, pp. 74–83.
[22] B. Xiong, X. Huang, and M. Wang, "Local gradient field feature contrast measure for infrared small target detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 3, pp. 553–557, 2021.
[23] M. Zeng, J. Li, and Z. Peng, "The design of top-hat morphological filter and application to infrared target detection," *Infrared Physics & Technology*, vol. 48, no. 1, pp. 67–76, 2006.
[24] T. Liu, J. Yang, B. Li, C. Xiao, Y. Sun, Y. Wang, and W. An, "Nonconvex tensor low-rank approximation for infrared small target detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–18, 2022.
[25] F. Wu, H. Yu, A. Liu, J. Luo, and Z. Peng, "Infrared small target detection using spatiotemporal 4-d tensor train and ring unfolding," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–22, 2023.