Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

TSCp-Regularized Low-Rank and Invariant Sparse Decomposition for Moving Object Detection under Complex Illumination Variations

Siying Liu
Chongqing University
Jianwen Xu
x.jw@cqu.edu.cn

Chongqing University

Research Article

Keywords:

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11791488/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

TSCp-Regularized Low-Rank and Invariant Sparse Decomposition for Moving Object Detection under Complex Illumination Variations

Siying Liu[1] and Jianwen Xu[2*]

[1]College of Mathematics and Statistics, Chongqing University, Chongqing, 404100, China.
[2*]National Elite Institute of Engineering, Chongqing University, Chongqing, 404100, China.

*Corresponding author(s). E-mail(s): xjw@cqu.edu.cn;

Abstract

Low-rank and sparse decomposition methods have attracted considerable attention in background subtraction and moving object detection, where backgrounds are often modeled as low-rank components and moving objects are modeled as sparse outliers. However, in time-lapse and motion-triggered image sequences, severe illumination variations, moving shadows, and background disturbances can be easily incorporated into the sparse term, resulting in false detections and missed detections. Shakeri and Zhang (2017) introduced the Low-Rank and Invariant Sparse Decomposition (LISD) framework to deal with this problem, where illumination-invariant representations and dynamic prior maps are employed to distinguish moving objects from illumination changes. In this paper, in order to alleviate the over-shrinkage of dominant background components caused by the conventional nuclear norm, we proposed three alternative versions of the LISD based on the truncated nuclear norm by Hu et al. (2012), the weighted nuclear norm by Gu et al. (2014) and the truncated Schatten capped p-norm (TSCp) by Yang and Xu (2025). In addition, experiments on five image sequences under complex illumination conditions are performed and it is demonstrated that the TSCp-LISD achieves the best performances.

Keywords: Moving object detection, Low-rank regularization, Illumination variation, Sparse decomposition

1 Introduction

Moving object detection is a fundamental computer vision task widely used in intelligent surveillance and wildlife conservation. Its accuracy directly affects high-level vision tasks including object tracking and behavior analysis. Actually, motion-triggered and time-lapse cameras are commonly employed for long-term monitoring. These sequences have large frame gaps and sharp illumination changes. Illumination variations and moving shadows introduce numerous false detections, making foreground detection under complex lighting a challenging task. In fact, low-rank and sparse decomposition are quite popular methods for moving object detection, where static backgrounds and moving targets are respectively modeled as low-rank matrices and sparse pixel outliers. For example, Candès et al. [1] proposed the Principal Component Pursuit (PCP) method, where

the nuclear norm is introduced for rank approximation and the ℓ_1-norm is adopted for sparsity. Zhou et al. [2] proposed the detecting contiguous outliers in the low-rank representation method (DECOLOR) with connectivity constraints to preserve foreground spatial integrity. Liu et al. [3] developed the LSD method with structured sparsity to maintain foreground continuity. Shakeri and Zhang [4] introduced the LISD method, where illumination-invariant representations and dynamic prior maps are employed to distinguish moving objects from illumination changes.

However, most of these methods adopted conventional nuclear norm regularization, which equally shrinks all singular values. This damages key background structures and fails to suppress noise and subtle illumination perturbations. Various refined low-rank regularizers are examined in the literatures, such as the truncated nuclear norm (TNN) by Hu et al. [5], the weighted nuclear norm (WNN) by Gu et al. [6] and the truncated Schatten capped p-norm (TSCp) by Yang and Xu [7]. These regularizers work well in image denoising and matrix completion, yet they are rarely integrated with illumination-aware detection frameworks. Their performance on long-interval time-lapse images with heavy lighting interference needs further study.

To tackle this issue, we embed the WNN, TNN and TSCp norm into the low-rank background regularization of the LISD framework and construct three improved detection models. In the meantime, some comparative experiments are conducted on public datasets and self-captured time-lapse sequences to illustrate performances of these methods under severe illumination variations. The rest of this paper is organized as follows. In Section 2, some related works are introduced on moving object detection based on low-rank and sparse decomposition. Then, three improved versions of the LISD are described in Section 3 and the corresponding experiment results are presented in Section 4. Finally, some concluding remarks are given in Section 5.

2 Moving object detection based on low-rank and sparse decomposition

In this section, we are to introduce some moving objection detection methods based on low-rank and sparse decomposition in the literatures. Firstly, it is assumed that each frame of the image sequence is vectorized column-wise and arranged sequentially to form the observation matrix $D \in \mathbb{R}^{m\times n}$, where m is the number of pixels in a single frame and n is the number of image frames. Then, a low-rank matrix L and a sparse matrix S are respectively employed to represent the background and the moving targets. Therefore, moving object detection can be modeled as the separation problem of low-rank background and sparse foreground.

Candès et al. [1] proposed the well-known Principal Component Pursuit (PCP) model for low-rank and sparse decomposition, where the nuclear norm is adopted as the convex relaxation of the matrix rank function and the ℓ_1 norm is chosen to be the sparse term. The corresponding optimization problem for PCP could be written as

$$\begin{aligned} \min_{L,S} \ & \|L\|_* + \lambda\|S\|_1, \\ \text{s.t.} \quad & D = L + S, \end{aligned} \tag{1}$$

where $\|L\|_* = \sum_i \sigma_i(L)$ denotes the nuclear norm, and $\sigma_i(L)$ represents the i-th singular value, $\|S\|_1 = \sum_{i,j} |S_{ij}|$ denotes the element-wise ℓ_1 norm, $\lambda > 0$ is the balancing parameter. However, moving targets often exhibit locally continuous regional structures within a single frame. Consequently, relying only on the ℓ_1 norm may lead to incomplete foreground, and make it hard to effectively distinguish targets from large-scale illumination variations.

To characterize the grouping or spatial structural properties of the foreground, researchers have introduced block sparsity, structured sparsity and connectivity constraints into the sparse term. In addition, Guyon et al. [8] proposed a low-rank and block-sparse decomposition model to characterize foreground regions via block-sparse constraints. For scenarios with overlapping groups, Mairal et al. [9] investigated the corresponding structured sparsity regularization and its optimization methods. Liu et al. further built the

LSD model [3]. The corresponding optimization problem is given by

$$\begin{aligned} \min_{L,S} \ & \|L\|_* + \lambda\Phi(S), \\ \text{s.t.} \quad & D = L + S, \end{aligned} \tag{2}$$

where $\Phi(S)$ denotes the structured sparse norm. Each foreground frame is partitioned into a set of overlapping local pixel groups $\mathcal{G}$, that is

$$\begin{aligned} \Phi(S) &= \sum_{j=1}^{n} \sum_{g \in \mathcal{G}} \|S_{g,j}\|_\infty \\ &= \sum_{j=1}^{n} \sum_{g \in \mathcal{G}} \max_{i \in g} |S_{ij}|, \end{aligned} \tag{3}$$

where $S_{g,j}$ denotes the foreground coefficient belonging to pixel group g in the j-th frame. Unlike the element-wise ℓ_1 norm that penalizes each pixel independently, the intra-group norm ℓ_∞ is determined by the maximum response within the group, followed by summation over all groups. This encourages the foreground to be distributed in a locally structured manner.

Another category of methods is to impose spatial connectivity constraints on the foreground support region directly. For example, for the DECOLOR [2] method, a binary foreground mask $M \in \{0,1\}^{m \times n}$ is introduced, where $M_{ij} = 1$ indicates whether the i-th pixel in the j-th frame belongs to the foreground. Thus, the optimization model is

$$\begin{aligned} \min_{L,S,M} \ & \tfrac{1}{2}\|D - L - S\|_F^2 + \lambda_L \|L\|_* \\ & + \lambda_S \|S \odot M\|_1 + \lambda_M \Psi(M), \end{aligned} \tag{4}$$

where $\odot$ denotes the element-wise product, and $\Psi(M)$ represents the spatial prior for the foreground support region.

The effect of spatial structure constraints depends on the spatial structure assumption of targets. For sequences collected at long time intervals, relying on foreground shape continuity is still not enough to reliably separate moving targets from illumination changes. Thus, for scenes with strong illumination variations, the observation matrix can be split into the following three components: background, foreground and illumination change, that is

$$D = L + S + C, \tag{5}$$

where L is the low-rank background, S denotes moving targets, and C stands for the illumination variation or shadow component.

Since both S and C may appear as sparse anomalies, it is usually hard to separate them only by sparsity. Shakeri and Zhang proposed the LISD model [4], where the prior graph Φ was used to represent illumination-invariant and various different penalties are chosen to indicate the foreground and illumination components based on the prior. The corresponding optimization model could be written as

$$\begin{aligned} \min_{L,S,C} \ & \|L\|_* + \lambda(\|S\|_1 + \|C\|_1) \\ & + \gamma \sum_i (\|G_i^\perp S_i\|_F^2 + \|G_i C_i\|_F^2), \\ \text{s.t.} \quad & D = L + S + C, \end{aligned} \tag{6}$$

where $G_j = \mathrm{diag}(\sqrt{\Phi_{:,j}})$, $G_j^\perp = \mathrm{diag}(\sqrt{1 - \Phi_{:,j}})$. The value of Φ tells if a pixel is more likely to belong to the moving targets or the illumination variation.

In fact, the matrices S and C explain different types of observation changes for the LISD model, which could reduce the risk that light fluctuations are wrongly detected as foreground. However, the traditional nuclear norm is still be used for the background term and we shall discuss some alternative versions of the LISD based on other norms recently proposed in the literatures.

3 The improved low-rank regularized LISD models

3.1 The improved models

In order to introduce the improved versions of the LISD in this subsection, we are firstly to discuss some norms recently proposed in the literatures. For the matrix $L \in \mathbb{R}^{m \times n}$, assume that the singular values satisfy $\sigma_1(L) \geq \sigma_2(L) \geq \cdots \geq \sigma_q(L) \geq 0$, where $q = \min(m, n)$.

The truncated nuclear norm proposed by Hu et al. [5] is defined as

$$\|L\|_r = \sum_{i=r+1}^{q} \sigma_i(L), \quad q = \min(m,n), \tag{7}$$

where r is the rank of the background matrix and it is obvious that the truncated nuclear norm discards the first r large singular values.

Moreover, through assigning different weight w_i to different singular value $\sigma_i(L)$, Gu et al. [6] examined the following weighted nuclear norm:

$$\|L\|_{w,*} = \sum_i w_i \sigma_i(L). \tag{8}$$

Very recently, by combination of the truncation strategy and Schatten capped p-norm by Li et al. [10], Yang and Xu [7] considered the following truncated Schatten capped p-norm:

$$\|L\|_{S_p,\tau,k} = \left(\sum_{i=k+1}^{q} \min(\sigma_i(L), \tau)^p \right)^{1/p} \tag{9}$$

To tackle the defect of uniform shrinkage for the traditional nuclear norm in the framework of LISD, we are to introduce three low-rank regularization norms to build alternative moving target detection models. The unified optimization objective could be rewritten as follows:

$$\begin{aligned} \min_{L,S,C} \ & \mathcal{R}(L) + \lambda(\|S\|_1 + \|C\|_1) \\ & + \gamma \sum_i (\|G_i^{\perp} S_i\|_F^2 + \|G_i C_i\|_F^2), \\ \text{s.t.} \quad & D = L + S + C, \end{aligned} \tag{10}$$

where $\mathcal{R}(L)$ is the low-rank regularizer, corresponding to improved models based on the TNN, WNN and TSCp norms. The proposed models, namely the TNN-LISD, the WNN-LISD and the TSCp-LISD, keep the illumination prior's separation ability and inherit the advantages of the different regularizers. Actually, our experimental results in Section 4 do demonstrate that the proposed methods could significantly improve foreground detection in the case of moving shadows and severe illumination changes.

3.2 The optimization algorithm

In this subsection, we are to introduce the optimization algorithm for problem (10). In particular, the inexact augmented Lagrange multiplier (IALM) method by Lin et al. [11] is used to iteratively solve the unified optimization problem and the linearized alternating direction method with adaptive penalty (LADMAP) by Lin et al. [12] is chosen for the corresponding sparse subproblems.

Firstly, the augmented Lagrangian function could be given by

$$\begin{aligned} & \mathcal{L}(L,S,C,Y;\mu) \\ & = \mathcal{R}(L) + \lambda(\|S\|_1 + \|C\|_1) \\ & + \gamma \sum_i (\|G_i^{\perp} S_i\|_F^2 + \|G_i C_i\|_F^2) \\ & + \langle Y, D - L - S - C \rangle \\ & + \tfrac{\mu}{2} \|D - L - S - C\|_F^2, \end{aligned} \tag{11}$$

where Y is the Lagrange multiplier matrix, $\mu > 0$ is the penalty parameter and $\langle , \rangle$ denotes the matrix norm, namely $\langle A, B \rangle = \text{trace}(A^T B)$ for matrices A and B.

Secondly, we are to alternately fixe other variables and solve the corresponding subproblems of L, S and C, respectively. In addition, the Lagrange multiplier is updated until the convergence condition is met. The iterative procedures could be summarized as follows:

Step 1: updating L^{k+1}.

Note that the augmented Lagrangian reduces to be the following form:

$$L^{k+1} = \arg\min_L \mathcal{R}(L) + \frac{\mu^k}{2} \|L - M^k\|_F^2, \tag{12}$$

where

$$M^k = D - S^k - C^k + \frac{Y^k}{\mu^k}. \tag{13}$$

This subproblem (12) yields a closed-form solution by the common singular value thresholding algorithm. Thus, the matrix M^k could be written as $M^k = U\Sigma V^{\top}$, and U is the left singular vector matrix and V is the right singular vector matrix. The singular value update rules of the three regularizers are as follows:

For the TNN-LISD, we have

$$L^{k+1} = U\,\mathrm{diag}(\sigma_1, \ldots, \sigma_r, \mathcal{S}_{1/\mu}(\sigma_{r+1}), \ldots)V^T, \tag{14}$$

where the first r largest singular values are discarded and soft-thresholding shrinkage on the rest singular values are performed.

For the WNN-LISD, we have

$$L^{k+1} = U\,\mathrm{diag}(\mathcal{S}_{w_1/\mu}(\sigma_1), \mathcal{S}_{w_2/\mu}(\sigma_2), \ldots)V^T, \tag{15}$$

where $\mathcal{S}_\epsilon(\sigma) = \mathrm{sign}(\sigma) \cdot \max(|\sigma| - \epsilon, 0)$ is the soft-thresholding shrinkage operator.

For the TSCp-LISD, we have

$$\mathcal{R}_{TSCp}(L) = \gamma_L \sum_{i=r+1}^{q} \min\{\sigma_i(L), \tau\}^p, 0 < p \leq 1, \tag{16}$$

where r is the number of leading singular values exempt from penalization, $\tau > 0$ is the capping threshold, p controls the shape of the nonconvex penalty, $\gamma_L > 0$ is the weight of the low-rank regularization and $q = \min(m, n)$. Then, based on the singular value decomposition of the matrix M^K, we have

$$L^{k+1} = U\,\mathrm{diag}(\widetilde{\sigma_1}, \ldots, \widetilde{\sigma_r}, \sigma_{r+1}^{k+1}, \ldots, \sigma_q^{k+1})V^T, \tag{17}$$

where σ_i^{k+1} is the $(k+1)$-th singular value updated via TSCp regularization in the i-th iteration and $i > r$.

Step 2: updating S^{k+1}.

On the other hand, the augmented Lagrangian can also reduce to:

$$\begin{aligned} S^{k+1} = \arg\min_S \ & \lambda\|S\|_1 + \gamma \sum_i \|G_i^\perp S_i\|_F^2 \\ & + \langle Y^k, D - L^{k+1} - S - C^k \rangle \\ & + \tfrac{\mu^k}{2}\|D - L^{k+1} - S - C^k\|_F^2. \end{aligned} \tag{18}$$

By the linearized approximation of LADMAP, a temporary variable for the i-th column could be given by

$$\begin{aligned} \mathrm{temp}S_i = S_i^k + \frac{1}{\eta\mu}\big[\mu(D_i - L_i^{k+1} - S_i^k - C_i^k \\ + Y_i^k/\mu) - 2\gamma(G_i^\perp)^T G_i^\perp S_i^k\big], \end{aligned} \tag{19}$$

where $\eta > 1$ is the linearized step-size parameter. Then, by applying element-wise soft-thresholding to each $\mathrm{temp}S_i$, we have

$$S_i^{k+1} = \mathcal{S}_{\lambda/(\eta\mu)}(\mathrm{temp}S_i). \tag{20}$$

Concatenate all columns to get the updated sparse foreground matrix S^{k+1}.

Step 3: updating C^{k+1}.

We can also rewrite the augmented Lagrangian as

$$\begin{aligned} C^{k+1} = \arg\min_C \ & \lambda\|C\|_1 + \gamma \sum_i \|G_i C_i\|_F^2 \\ & + \langle Y^k, D - L^{k+1} - S^{k+1} - C \rangle \\ & + \tfrac{\mu^k}{2}\|D - L^{k+1} - S^{k+1} - C\|_F^2. \end{aligned} \tag{21}$$

Similar to Step 2, C^{k+1} can also be obtained via LADMAP. That is

$$C_i^{k+1} = \mathcal{S}_{\lambda/(\eta\mu)}(\mathrm{temp}C_i), \tag{22}$$

where

$$\begin{aligned} \mathrm{temp}C_i = C_i^k + \frac{1}{\eta\mu}\big[\mu(D_i - L_i^{k+1} - S_i^{k+1} - C_i^k \\ + Y_i^k/\mu) - 2\gamma(G_i)^T G_i C_i^k\big]. \end{aligned} \tag{23}$$

Concatenate all columns to get the updated illumination variation matrix C^{k+1}.

Finally, we can find that the three improved LISD methods share the identical iterative framework and differ only in the singular value shrinkage rule. Therefore, we take the proposed TSCp-LISD as an example to present the complete algorithm procedure, and all details about TSCp-LISD are provided in the following Algorithm 1.

4 Experimental results and analysis

In this section, we are to introduce the experimental details and results. Firstly, we introduce

Algorithm 1 Iterative Low-rank and Invariant Sparse Decomposition via IALM

Input: Observation matrix D, illumination prior map Φ,
parameters $\lambda, \gamma, \gamma_L, \mu_0, \rho, \eta, r, \tau, p, \varepsilon$; maximum iteration number $K_{\max}$.

Initialization:

$L^0 = S^0 = C^0 = Y^0 = 0,\ \mu = \mu_0,\ k = 0$

while not converged do
$M = D - S^k - C^k + \mu^{-1} Y^k$
$(U, \delta, V) = \mathrm{svd}(M)$
$x_i = \delta_i, i = 1, \ldots, r$
$x_i = \arg\min_{x \geq 0} \left[\frac{1}{2}(x - \delta_i)^2 + (\gamma_L/\mu) \min(x, \tau)^p \right], \quad i = r+1, \ldots, q$
$L^{k+1} = U \,\mathrm{diag}(x) V^T$
for each column $i = 1, \ldots, n$ do
$\mathrm{temp}S_i = S_i^k + [\mu(D_i - L_i^{k+1} - S_i^k - C_i^k) + Y_i^k - 2\gamma(G_i{}^{\perp})^T G_i{}^{\perp} S_i^k]/(\eta\mu)$
$S^{k+1} = \mathcal{S}_{\lambda/(\eta\mu)}(\mathrm{temp}S)$, $\mathrm{temp}S = [\mathrm{temp}S_1, \ldots, \mathrm{temp}S_n]$
for each column $i = 1, \ldots, n$ do
$\mathrm{temp}C_i = C_i^k + [\mu(D_i - L_i^{k+1} - S_i^{k+1} - C_i^k) + Y_i^k - 2\gamma G_i^T G_i C_i^k]/(\eta\mu)$
$C^{k+1} = \mathcal{S}_{\lambda/(\eta\mu)}(\mathrm{temp}C)$, $\mathrm{temp}C = [\mathrm{temp}C_1, \ldots, \mathrm{temp}C_n]$
$R^{k+1} = D - L^{k+1} - S^{k+1} - C^{k+1}$
$Y^{k+1} = Y^k + \mu R^{k+1}$
$\mu = \rho\mu; k = k + 1$
converged if $\|R^{k+1}\|_F / \|D\|_F < \varepsilon$
end while

Output: L^k, S^k, C^k

the experimental setup and evaluation metrics adopted for foreground detection.

4.1 Experimental setup and evaluation metrics

For moving foreground detection of sequences under illumination variations, two groups of experiments with different illumination variations are considered.

On the one hand, we consider two image sequences named Cd1 for outdoor scene and Cd2 for indoor scene, both of the two sequences are derived from the public CDnet dataset [13]. The frames are continuous and only moderate illumination fluctuations and slowly moving shadows are contained, leading to low interference.

On the other hand, we also examine other three sequences named People, Wildlife1 and Wildlife2 with complex illumination variations.

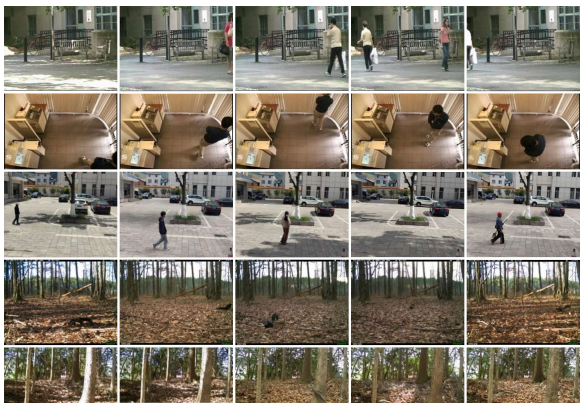

Fig. 1 Representative frames of the five types of test sequences

Table 1 Characteristics of the evaluated sequences

Sequences	Size × number of frames
Cd1	[320,240]×113
Cd2	[720,480]×124
People	[360,271]×115
Wildlife1	[360,270]×57
Wildlife2	[360,176]×116

The People sequence is a self-collected time-lapse sequence. Continuous gradual variation of overall frame brightness is observed, and foreground detection is disturbed by wind-driven drifting tree shadows. Wildlife1 and Wildlife2 are selected from the public ENA24 wildlife dataset [14], featuring a long-time span. Some targets are covered by strong light or heavy shadows, and the background shares similar with moving objects.

Subsequently, we are further to explain the details of the datasets, detection metrics and comparison methods. As shown in Fig. 1, some frames in the datasets contain only illumination variations without moving targets. For each dataset sequence, 15 frames featuring drastic illumination perturbations and moving foreground targets are selected as test samples, and the foreground ground truth is manually annotated. The characteristics of the datasets are summarized in Table 1.

The pixel-level F-measure is utilized for quantitative evaluation and is defined as:

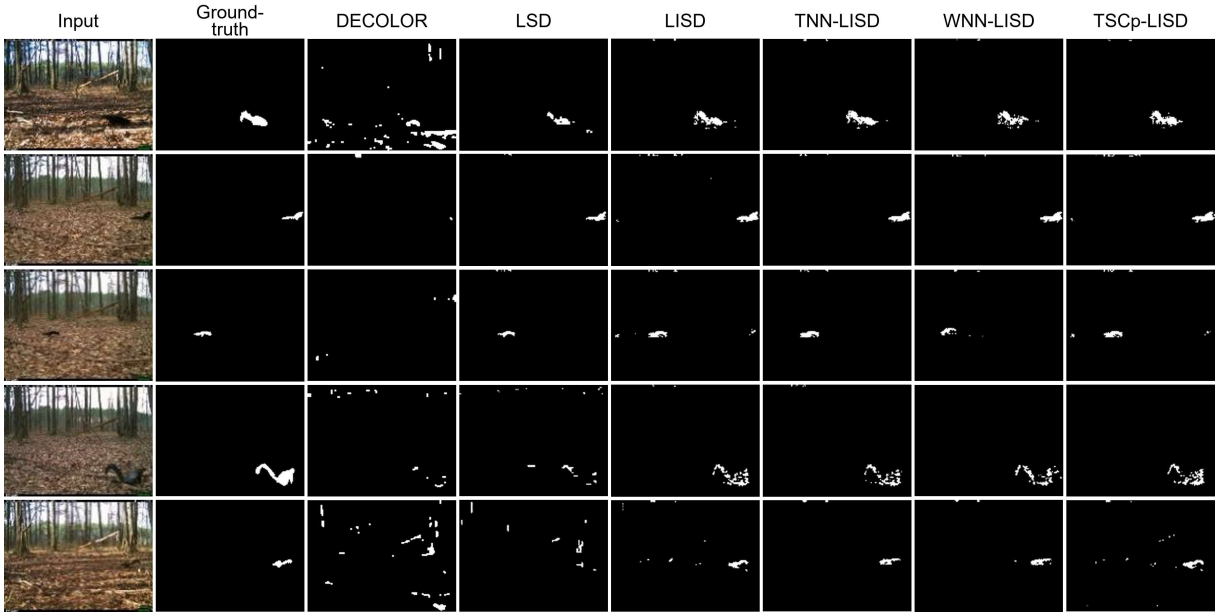

Fig. 2 Comparison of qualitative results between our method and five competing methods on selected images of sequence "Wildlife1"

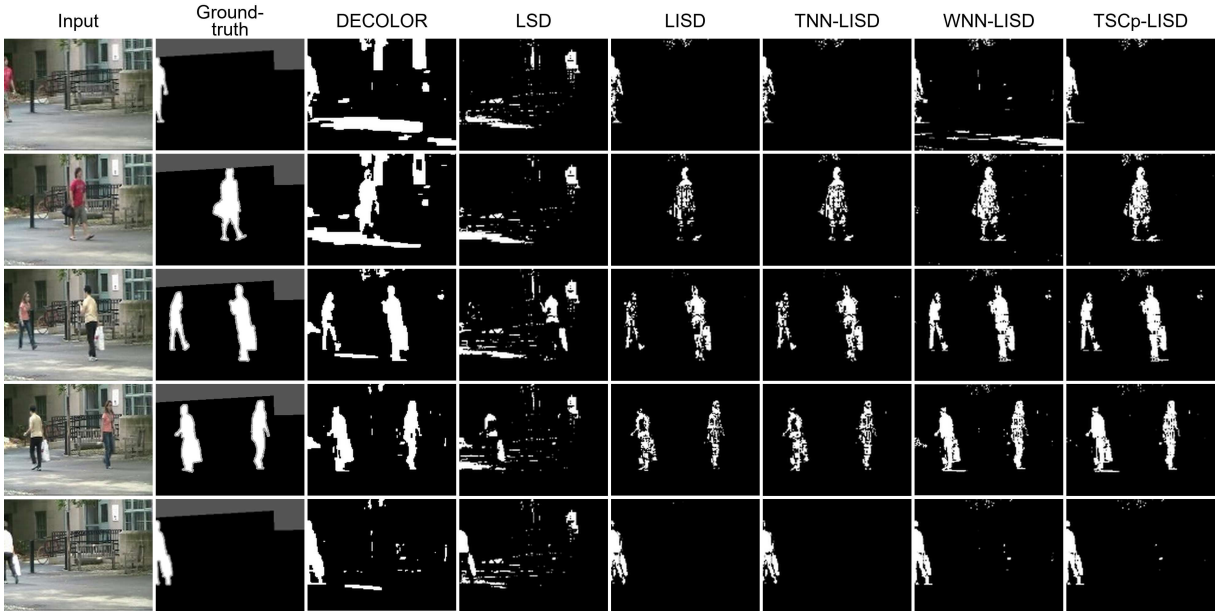

Fig. 3 Comparison of qualitative results between our method and five competing methods on selected images of sequence "Cd1"

$$\text{F-measure} = 2\frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}}. \qquad (24)$$

The accuracy and completeness of detection results are comprehensively reflected by the F-measure. A higher value indicates better foreground detection performance. Meanwhile, the

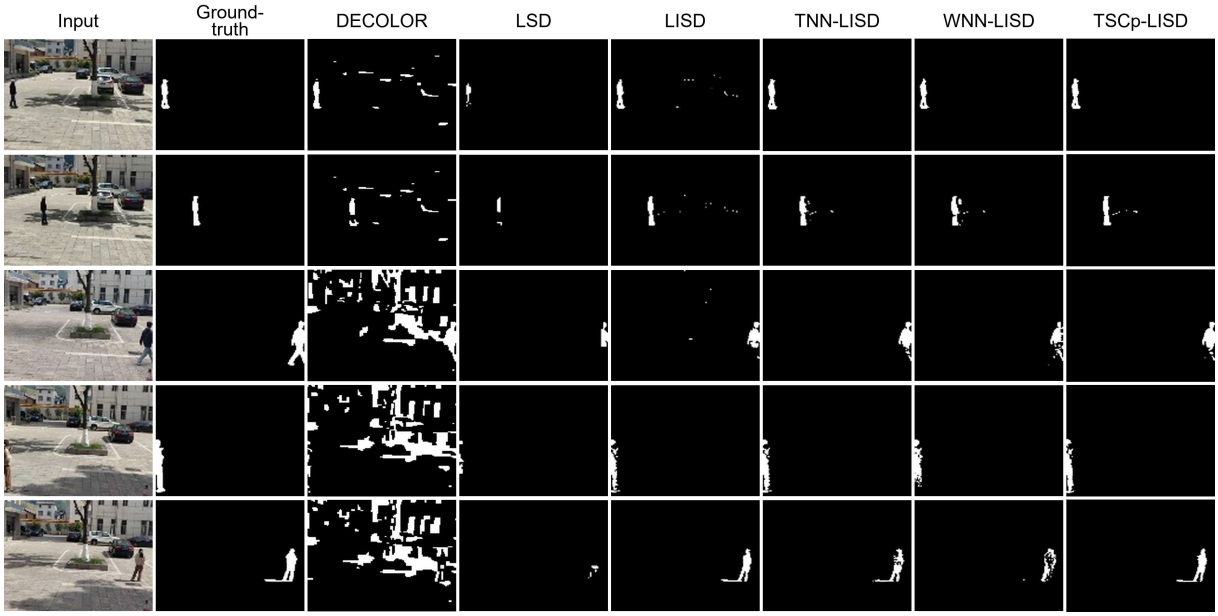

Fig. 4 Comparison of qualitative results between our method and five competing methods on selected images of sequence "People"

standard deviation of frame-wise F-measure values is calculated to evaluate the detection stability of the algorithm across different frames. A smaller standard deviation implies less performance fluctuation of the algorithm under scenarios with drastic illumination variations.

4.2 Analysis of experimental results

In this subsection, the experimental results for various models are presented for comparison, including DECOLOR, LSD, LISD, WNN-LISD, TNN-LISD and TSCp-LISD.

In particular, Wildlife1 is selected as an example for demonstration, and five frames are extracted for intuitive comparison, as illustrated in Fig. 2.

Drastic illumination variations are evident in the first and fifth rows in Fig. 2. The DECOLOR and LSD methods have limited ability to distinguish genuine moving targets from illumination perturbations effectively, suffering from false detections and missed detections to varying degrees. The LISD method can separate illumination variations and moving targets relatively well. The improved LISD models exhibit better performance than the first three methods and the TSCp-LISD model achieves the best performance.

Furthermore, the results for sequences Cd1 and People are presented to further compare these models. The corresponding visualized comparison results are shown in Fig. 3 and Fig. 4, respectively.

As shown in Fig. 3, partial human regions are detected by the DECOLOR model, yet extensive illumination variations are misclassified as foreground. The LSD method has obvious missed detections. Similar phenomena can be observed in Fig. 4.

From the qualitative results, it can be concluded that incomplete foreground regions, background false detections and missed detections are prone to appear in DECOLOR and LSD. The LISD model achieves relatively good detection performance. TNN-LISD, WNN-LISD and TSCp-LISD methods further improve the integrity of moving targets, and TSCp-LISD exhibits superior stability.

In addition, the quantitative metrics results of the competing methods on the five sequences are reported in Table 2, where each entry illustrates the mean F-measure/standard deviation.

The numerical results demonstrate that the improved methods based on LISD can provide

Table 2 F-measure results of each method on the five datasets

Sequence	DECOLOR	LSD	LISD	TNN-LISD	WNN-LISD	TSCp-LISD
Cd1	0.7141/0.2841	0.0568/0.0422	0.6251/0.1197	0.7669/0.1292	0.8520/0.1846	**0.8742/0.1627**
Cd2	0.4777/0.3474	0.4345/0.0622	0.6440/0.1460	0.6564/0.0872	0.6321/0.0872	**0.7717/0.0996**
People	0.1630/0.2101	0.4086/0.1670	0.7869/0.1313	0.8399/0.0918	0.7244/0.1544	**0.8677/0.0852**
Wildlife1	0.1232/0.1130	0.5060/0.2425	0.7190/0.1103	0.7275/0.1082	0.7198/0.1120	**0.7853/0.0765**
Wildlife2	0.1407/0.1773	0.2518/0.1914	0.5614/0.1442	0.5715/0.1224	0.6997/0.0991	**0.7210/0.0698**

better performance in handling such illumination changes than other competing methods in time-lapse sequences.

5 Concluding remarks

This paper investigates moving-object detection in time-lapse image sequences. Based on the LISD framework, three improved models are constructed by incorporating the truncated nuclear norm, weighted nuclear norm and TSCp norm. Experiments show that the three improved models achieve different levels of improvement in background recovery and foreground extraction. The TNN-LISD and WNN-LISD methods exhibit sequence-dependent performance: each yields improved detection in some scenarios but suffers from performance degradation in others. In contrast, the TSCp-LISD method achieves the best overall performance and demonstrates more consistent improvements across all tested sequences.

Although the proposed models achieve promising detection results on the test sequences, there remains room for improvement in the separation of the illumination component and sparse foreground component. The current model adopts unified regularization parameters for these two components, which weakens the ability to distinguish illumination perturbations from genuine moving targets. In future work, independent tunable parameters can be introduced for the illumination components and the sparse foreground component to improve the decomposition accuracy of the two types of components.

Statements and Declarations

Funding

No funding was received for this work.

Data Availability

The CDnet 2012 and ENA24-detection datasets used in this study are publicly available: CDnet 2012: https://www.changedetection.net/dataset2012, and ENA24-detection: https://lila.science/datasets/ena24detection. And the self-collected People time-lapse dataset generated in this work is available from the corresponding author upon reasonable request.

Competing interests

The authors declare that they have no competing interests.

Author Contributions

Siying Liu: Methodology, Software, Writing-original draft, Visualization, Validation. Jianwen Xu: Conceptualization, Writing-review & editing, Supervision. All authors have read and agreed to the published version of the manuscript.

References

[1] Candès, E.J., Li, X., Ma, Y., et al.: Robust principal component analysis? Journal of the ACM **58**(3), 1–37 (2011). https://doi.org/10.1145/1970392.1970395.

[2] Zhou, X., Yang, C., Yu, W.: Moving object detection by detecting contiguous outliers in the low-rank representation. IEEE Transactions on Pattern Analysis and Machine Intelligence **35**(3), 597–610 (2013). https://doi.org/10.1109/TPAMI.2012.132.

[3] Liu, X., Zhao, G., Yao, J., et al.: Background subtraction based on low-rank and structured sparse decomposition. IEEE Transactions on Image Processing **24**(8), 2502–2514 (2015). https://doi.org/10.1109/TIP.2015.2419084.

[4] Shakeri, M., Zhang, H.: Moving object detection in time-lapse or motion trigger image sequences using low-rank and invariant sparse decomposition. In: Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, pp. 5133–5141 (2017). https://doi.org/10.1109/ICCV.2017.548.

[5] Hu, Y., Zhang, D., Ye, J., Li, X., He, X.: Fast and accurate matrix completion via truncated nuclear norm regularization. IEEE Transactions on Pattern Analysis and Machine Intelligence **35**(9), 2117–2130 (2013). https://doi.org/10.1109/TPAMI.2012.271.

[6] Gu, S., Xie, Q., Meng, D., et al.: Weighted nuclear norm minimization with application to image denoising. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, pp. 2862–2869. IEEE (2014). https://doi.org/10.1109/CVPR.2014.366.

[7] Yang, Z., Xu, J.: An alternative truncated Schatten capped p-norm regularization method for low-rank matrix approximation. Applied Mathematical Modelling **146**, 116186 (2025). https://doi.org/10.1016/j.apm.2025.116186.

[8] Guyon, C., Bouwmans, T., Zahzah, E.H.: Foreground detection based on low-rank and block-sparse matrix decomposition. In: 2012 IEEE International Conference on Image Processing, Orlando, pp. 1225–1228. IEEE (2012). https://doi.org/10.1109/ICIP.2012.6467087.

[9] Mairal, J., Jenatton, R., Obozinski, G., Bach, F.: Network flow algorithms for structured sparsity. In: Advances in Neural Information Processing Systems 23, Vancouver, pp. 1558–1566 (2010). https://arxiv.org/abs/1008.5209.

[10] Li, G., Guo, G., Peng, S., Wang, C., Yu, S., Niu, J., Mo, J.: Matrix completion via Schatten capped p norm. IEEE Transactions on Knowledge and Data Engineering **34**(1), 394–404 (2020). https://doi.org/10.1109/TKDE.2020.2978465.

[11] Lin, Z., Chen, M., Ma, Y.: The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. UIUC Technical Report UILU-ENG-09-2214, arXiv:1009.5055 (2010). https://arxiv.org/abs/1009.5055.

[12] Lin, Z., Liu, R., Su, Z.: Linearized alternating direction method with adaptive penalty for low-rank representation. In: Advances in Neural Information Processing Systems 24, pp. 612–620 (2011). https://arxiv.org/abs/1109.0367.

[13] Goyette, N., Jodoin, P.-M., Porikli, F., Konrad, J., Ishwar, P.: Changedetection.net: A new change detection benchmark dataset. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 1–8 (2012). https://doi.org/10.1109/CVPRW.2012.6238919.

[14] LILA Science: ENA24-detection: Eastern North America camera trap detection dataset (2024). https://lila.science/datasets/ena24detection.