IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 10, 2020, accepted February 23, 2020, date of publication February 28, 2020, date of current version March 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2977273

Moving Object Detection Based on Non-Convex RPCA With Segmentation Constraint

ZIXUAN HU[1], YONGLI WANG[1], RUI SU[2], XINXIN BIAN[1], HONGCHAO WEI[1], AND GUOPING HE[3]

[1]College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China
[2]Qingdao Urban Planning & Design Research Institute, Qingdao 266071, China
[3]Shandong Computing Center, Qilu University of Technology, Jinan 250353, China

Corresponding author: Yongli Wang (wangyongli@sdkd.net.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 11901359, and in part by the Shandong Provincial Natural Science Foundation under Grant ZR2019QA017.

ABSTRACT Recently, robust principal component analysis (RPCA) has been widely used in the detection of moving objects. However, this method fails to effectively utilize the low-rank prior information of the background and the spatiotemporal continuity prior of the moving object, and the target extraction effect is often poor when dealing with large-scale complex scenes. To solve the above problems, a new non-convex rank approximate RPCA model based on segmentation constraint is proposed. Firstly, the model adopts the low-rank sparse decomposition method to divide the original video sequence into three parts: low-rank background, moving foreground and sparse noise. Then, a new non-convex function is proposed to better constrain the low-rank characteristic of the video background. Finally, based on the spatiotemporal continuity of the foreground object, the video is segmented by the super-pixel segmentation technology, so as to realize the constraint of the motion foreground region. The augmented Lagrange multiplier method is used to solve the model. Experimental results show that the proposed model can effectively improve the accuracy of moving object detection, and has better visual effect of foreground object detection than existed methods.

INDEX TERMS Moving object detection, robust principal component analysis, non-convex rank approximation, video segmentation.

I. INTRODUCTION

As a hot spot in computer vision research, moving object detection has a wide range of practical applications in video surveillance [1], military investigation [2], medical image processing [3] and many other fields. The main purpose of moving object detection is to extract the moving object from the video image and obtain the feature information of the moving object, such as color, shape, contour and so on, which lays the foundation for the follow-up work of target tracking, recognition and classification. However, due to the background noise, illumination change, shadow and other factors in the scene, most current algorithms can not effectively deal with the moving object detection in real complex background. Therefore, it is necessary to study more effective methods to detect the moving objects in complex background.

Robust principal component analysis (RPCA) [4] is one of the most important methods for moving object detection. Compared with the traditional methods, RPCA has some obvious advantages in detection accuracy and robustness. For example, RPCA model can separate the background and the moving object at the same time. Besides, it can separate the background and the moving object directly from the video without inputting the clean background video as the background training sample. RPCA is now widely used in the field of image processing, such as video background recovery [5]–[7], moving object detection [8], [9], image denoising [10]–[12] and so on. The traditional RPCA model can be described as the following optimization problem:

$$\min_{L,S} rank(L) + \lambda \|S\|_0 \quad s.t. \quad X = L + S\,, \tag{1}$$

where $\|\cdot\|_0$ is the L_0 norm of matrix (that is, the number of non-zero elements in matrix S).

However, since the rank function and the L_0 norm of the matrix are non-convex and discrete, the solution of model (1) is NP-hard. Hence, most existed theoretical researches adopt nuclear norm as the convex approximation

The associate editor coordinating the review of this manuscript and approving it for publication was Long Wang.

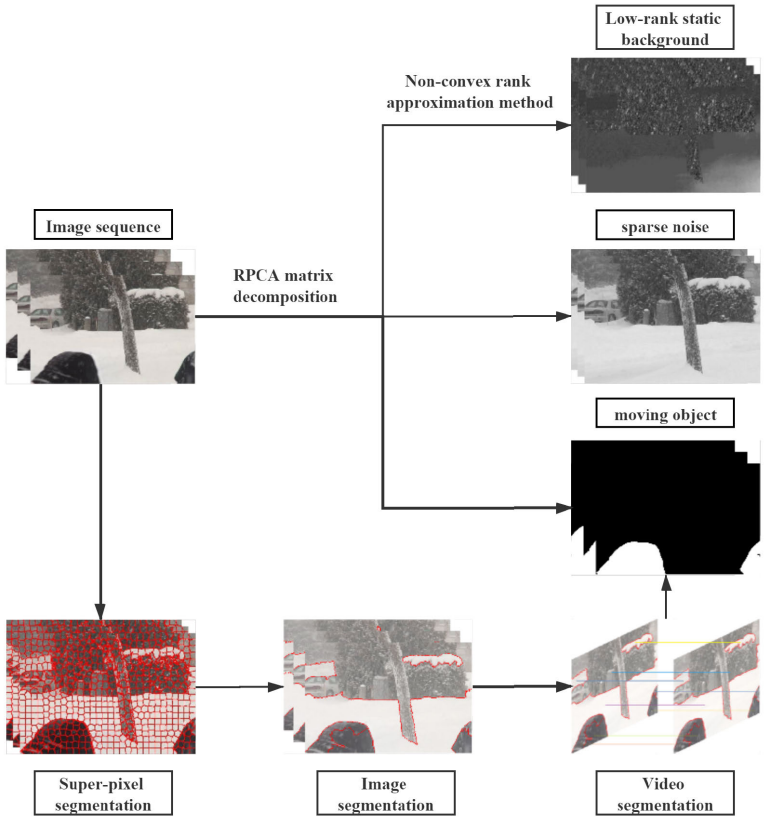

FIGURE 1. Illustration of the proposed NCSC-RPCA model for moving object detection.

of the rank function, and use L_1 norm to replace L_0 norm [13]. Then the above RPCA model can be relaxed as a convex optimization problem:

$$\min_{L,S} \|L\|_* + \lambda\|S\|_1 \quad s.t. \quad X = L + S\,, \tag{2}$$

where $\|L\|_* = \sum_{i=1}^{r} \sigma_i$, σ_i is the singular value of the matrix L and $\|S\|_1 = \sum_{i=1}^{m}\sum_{j=1}^{n} |s_{ij}|$.

Note that the traditional RPCA model is in fact extended from the low-rank matrix model of compressed sensing, it only considers the characteristics of low-rank and sparsity in the perspective of video matrix elements, and does not consider the internal spatiotemporal correlation between matrix elements. Hence, there are still some shortcomings in the traditional and some improved RPCA model. Firstly, for the low-rank part, scholars usually use the nuclear norm as the convex approximation of the rank function of the matrix [14]. However, the nuclear norm simply adds up all the singular values of the matrix. If one or more singular values of the matrix are too large, the nuclear norm will overestimate the rank of the matrix, which will affect the recovery of the low rank matrix. In order to overcome the above problems, scholars tried to use different non-convex functions to approximate the rank function [15]–[23], and the experiment results showed that the non-convex methods could extract clearer foreground targets than the traditional convex methods. Secondly, for many complex moving objects, L_1 norm is often not a good way to describe the real foreground region. However, if the spatiotemporal continuity between moving object can be reasonably utilized to eliminate the noise part which is inconsistent with the characteristic of moving object, the moving object can be detected robustly. Based on this, some scholars considered extending the RPCA model by introducing spatial continuity constraints in the foreground, such as [24]–[28]. In particular, the group sparsity constraints used in [29]–[36] are very helpful for handling the scenes with dynamic background, and can extract cleaner foreground targets than traditional RPCA. Recently, some moving object detection methods based on deep convolutional neural network have been proposed. Sultana *et al.* [37] presented a fusion based foreground detection algorithm which exploits color as well as depth information, with the help of Generative Adversarial Network (GAN). After that, Sultana *et al.* [38] presented a solution based on conditional Generative Adversarial Network (cGAN) with an architecture of robust deep auto-encoder trained with occlusion free moving object detection scenario. In reference [39], a unified method based on GAN and image inpainting is proposed. These methods can work well in many complex situations, but the training process of GAN requires a large number

of sample data, which will directly affect the training effect.

Considering the success of super-pixel segmentation technology in image processing and the advantage of non-convex rank approximation, this paper hopes to establish an effective model that can effectively detect the moving object under dynamic background based on these two techniques. First, the original video sequence is divided into three parts by low-rank sparse decomposition: the low-rank background, the moving foreground and the sparse noise. For the low-rank term, a new non-convex rank approximation function is proposed to overcome the defect of overestimation as well as time-consuming in the solution of the nuclear norm. For the moving foreground term, it is regarded as a significant moving object in the video, occupying a small part of the space continuous region in the video frame. In addition, note that the moving object in the video foreground is continuous in the time domain, this paper considers introducing the spatiotemporal continuity of the video foreground as the foreground prior. In order to characterize the spatiotemporal continuity of the moving object, this paper takes the result of video segmentation, based on the super-pixel segmentation technology, as a group sparse constraint. Thus, an improved RPCA model based on the segmentation constraint and the new non-convex rank approximation function (NCSC-RPCA) is proposed. The NCSC-RPCA model is solved by the augmented Lagrange multiplier method. The experimental results of seven dynamic videos in the CDnet2014 database show that the model proposed in this paper can effectively improve the accuracy of moving target extraction and has better visual effects of foreground target extraction than existed methods.

The paper is organized as follows. In the next section, some related works are introduced. Section 3 gives the proposed NCSC-RPCA model and the solution framework. The experiment results and comparison with existed models are given in section 4. Section 5 is the conclusion.

II. RELATED WORKS

In this section, we would like to provide a brief survey for stable RPCA decomposition, non-convex approaches and segmentation constraints.

A. STABLE RPCA DECOMPOSITION

The traditional RPCA model has achieved good results when dealing with the foreground/background separation of static background videos. However, when processing the dynamic background videos, the original RPCA model often extracts the noise as the foreground, which greatly affects the extraction accuracy of the foreground target. In view of this defect, Zhou *et al.* [40] proposed SPCP model, and experimentally verified that the model can achieve stable and accurate recovery of low-rank term with less noise. After that, 3WD model proposed by Oreifej *et al.* [41] and the Tri-decom model proposed by Jin *et al.* [42] both use three-term decomposition technique to deal with the image data corrupted by both large sparse noise and small dense noise. In particular, considering the superiority of the three-term decomposition in noise removal, Yang *et al.* [43] proposed an effective vehicle counting system for detecting and tracking vehicles in complex traffic scenarios. In addition, based on reference [40], Yin *et al.* [44] proposed to express the convex optimization problem with an inseparable non-convex regularization formula to recover from the superposition of the low-rank matrix and the sparse matrix observed in the additive white Gaussian noise. The experimental results show that the stable RPCA decomposition can effectively remove the noise, so as to extract a cleaner prospect.

B. NON-CONVEX APPROACHES

The original RPCA model is a nonlinear, non-convex and discontinuous combinatorial optimization problem, and its solution is NP hard. In order to solve the problem easily, Wright *et al.* [13] proposed using nuclear norm as a convex approximation of the rank function. But there are two main problems in using nuclear norm to approximate the rank function. On the one hand, the nuclear norm is defined as the sum of all singular values of the matrix. In case the first several singular values of the matrix far exceed the real rank, this will lead to the problem of over-estimation of the rank and affects the real experimental results. In order to overcome this defect, some non-convex functions have been proposed to approximate the rank function, such as the weighted nuclear norm [15], truncated nuclear norm [16], non-convex gamma norm [17], non-convex laplace norm [18] and so on. In particular, Yang *et al.* [19] proposed a general optimization model for extracting background and foreground from a surveillance video. This model can be nuclear-norm-free, and can incorporate different possibly non-convex sparsity inducing regularization functions for extracting the foreground. Experimental results show that RPCA models based on non-convex rank approximation have better foreground/background separation effect and shorter operation time. On the other hand, the solution of RPCA models based on nuclear norm approximation need singular value decomposition in each iteration, which leads to high computational complexity. In order to overcome this defect, Zhou and Tao [45] proposed GoDec algorithm, which adopts the idea of random optimization to avoid singular value decomposition and greatly improves the computational efficiency. However, due to its random optimization, the extraction accuracy is not high. Netrapalli *et al.* [20] proposed to solve RPCA model by non-convex alternating projection method, which has low computational complexity and fast convergence speed. Balcan *et al.* [21] proposed a new analytical framework for the non-convex matrix decomposition problem, and verified the exact recoverability and strong duality of the matrix decomposition problem through the framework. Wen *et al.* [23] proposed an approximate block coordinate descent(BCD) algorithm based on generalized non-convex regularization and principal component analysis. The experimental results show that the performance of these

non-convex methods is better than that of the traditional convex methods.

C. SEGMENTATION CONSTRAINTS

The above improved RPCA models are mostly based on the assumption that the video background is static or quasi-static, but this situation is often not true in real video images, especially when there is a dynamic background in the video such as waves, snowflakes, swinging leaves and so on. Note that if the spatiotemporal continuity between moving object can be reasonably used to eliminate the noise part that is inconsistent with the characteristic of the moving object, then the robustness of the moving object detection can be performed. Based on this consideration, some scholars consider making use of the spatiotemporal key information contained in video and image, that is, to extend RPCA model by imposing some additional constraints. In order to improve the separation effect of video foreground and background, Yang et al. [24] estimated the foreground with significant motion by dense optical flow method and got the real value m of binary template. Then they added this matrix into the optimization model and proposed the motion assisted matrix recovery (MAMR) model. After that, Sobral *et al.* [25] further improved the model on the basis of [24] and [41] and proposed a double-constrained RPCA model, named SCM-RPCA (Shape and Confifidence Map based RPCA). In addition, Ebadi *et al.* [26] considered applying block sparsity to the foreground and introduced an extremely efficient ''SVD-free'' technique to solve the problem of algorithm complexity. Recently, with the wide application of super-pixel technology in the field of image processing, some scholars proposed to use super-pixel segmentation technology to group pixels with similar appearance, space or time adjacent, so as to establish the spatiotemporal continuity between pixels. At present, existed super-pixel methods include normalized cuts [46], meanshift [47], turbo-pixel [48], SLIC [49] and so on. Ebadi and Izquierdo [27] introduced a new sparsity criterion and group structure sparsity constraint to the foreground part, and proposed a dynamic tree-structure sparse RPCA model. Javed *et al.* [28] proposed a matrix decomposition method based on super-pixel combining maximum norm regularization and structural sparsity constraints. Javed *et al.* [29] incorporated two different manifold regularizations on the sparse component based on the local and global invariance assumption, and proposed a novel Super-pixel based Spatiotemporal Manifold Structured-Sparse-RPCA (SSMS-RPCA) algorithm for moving object detection. Li *et al.* [30] proposed a model combining the segmentation constraint and the saliency constraint, which can effectively detect slow-moving targets. Chen *et al.* [31] proposed to use the structured L_p-regularized low-rank representation (SLLR) to enhance the detected foreground while suppressing the noise and dynamic background. Zheng *et al.* [33] proposed a background subtraction method based on multi-scale structure low-rank sparse factorization, and considered both appearance consistency and spatial compactness in the framework. By introducing a priori of spatiotemporal continuity into the foreground object, the above model greatly improves the foreground extraction effect in dynamic background video. Moreover, some additional structural information is incorporated into the low-rank matrix approximation model. Javed *et al.* [35] proposed a motion-aware regularization of graphs on low-rank component for video background modeling. Experimental results show that properly using the spatiotemporal key information contained in the video can significantly improve the extraction effect of foreground moving objects.

III. MODEL AND ALGORITHM

In this section, we first propose a new non-convex rank approximation function. Secondly, we use the super-pixel segmentation technology to segment the video and take the segmentation result as the group sparse constraint of the video foreground. Then, we combine it with the non-convex rank approximation function and propose the NCSC-RPCA model. Finally, we use the augmented Lagrange multiplier method to solve the model and get the closed form solution of each sub-problem.

A. PROPOSED NONCONVEX RANK APPROXIMATION FUNCTION

In order to better approximate the rank function and avoid the singular value decomposition process in the solution of nuclear-norm-based RPCA models, we propose a new non-convex rank approximation function in this section. Based on the summary of the existed characteristics of non-convex rank approximation functions in reference [50] and the properties of power functions, we consider the following non-convex rank approximation function in this paper:

$$rank(L) \approx \|L\|_{\varphi} = \sum_{i=1}^{m} \frac{\sigma_i(L)}{\sqrt{\sigma_i(L)^2 + \varphi}}, \tag{3}$$

where $\varphi > 0$ is a model parameter and $\sigma_i(L)$ is the i-th singular value of matrix L.

Lemma 1: The non-convex rank approximation function $\|L\|_{\varphi}$ has the following properties:

(1) $\lim_{\varphi \to 0} \|L\|_{\varphi} = rank(L)$;

(2) If $\sigma_i(L) = 0$, then $g(\sigma_i(L)) = 0$;

(3) $\|L\|_{\varphi}$ is unitarily invariant, that is, $\|L\|_{\varphi} = \|ULV\|_{\varphi}$ for any orthonormal $U \in R^{m\times m}$ and $V \in R^{n\times n}$;

(4) Positive definiteness: For any $L \in R^{m\times n}$, $\|L\|_{\varphi} \geq 0$; If and only if $L = 0$, $\|L\|_{\varphi} = 0$.

Figure 2 shows the approximation of the rank function using non-convex φ function, nuclear norm with an increasing value of σ_i. It can be seen that the nuclear norm deviates from 1 greatly with the increase of singular value $\sigma_i(L)$, while the nonconvex function can approach 1 more stably. The smaller the parameter value φ is, the better the effect of rank function approximation is. In the later experiments, it can be seen that the low rank matrix can be more accurately restored by appropriately adjusting the parameter $\varphi(0 < \varphi < 1)$.

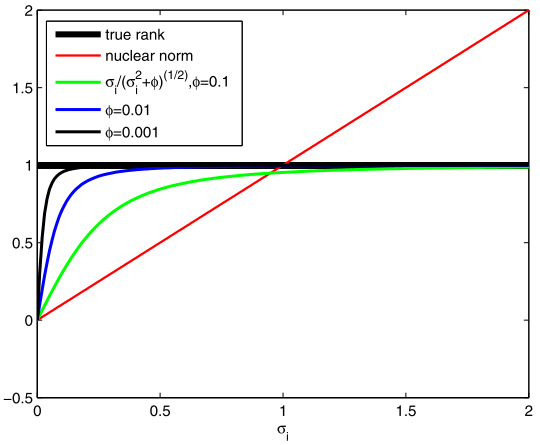

FIGURE 2. Approximation of the rank function using non-convex φ function, nuclear norm with an increasing value of σ_i.

B. ESTABLISH TEMPORAL AND SPATIAL CONTINUITY BY USING SEGMENTATION CONSTRAINT

In order to ensure that RPCA still have good foreground extraction effect in dynamic video background, we consider introducing spatiotemporal continuity in the model. In order to describe spatiotemporal continuity of the foreground object, we mainly use the super-pixel segmentation technology. The advantage of the super-pixel segmentation is that it can retain the effective information in the original image, and generally can not destroy the boundary information of the foreground object in the original image. The steps of video segmentation are as follows. Firstly, we use SLIC [49] to divide each original image into super-pixels. Secondly, we use CTM [51] to merge adjacent and similar super-pixels in a single image. In other words, the original image is divided into a series of adjacent sub-regions with similar colors, which is called image segmentation. Finally, we use Pairwise Steepest Descent of Coding Length [52] to merge the similar sub-regions between adjacent frames of the original video sequence, also known as video segmentation. In this process, whether the similar sub regions between adjacent frames can be merged depends on whether their region centers are less than the preset center threshold. The overall video segmentation framework is shown in the following Algorithm 1.

C. PROPOSED NCSC-RPCA MODEL

We can get the video pixel label matrix P through Algorithm 1. Then, by classifying the pixels belonging to the same label, we can get the video segmentation constraint C. Here, $C = C_1 \bigcup C_2 \bigcup \cdots \bigcup C_m$, where m is the number of groups. Using the non-convex function $\|L\|_\varphi$ instead of nuclear norm and considering spatiotemporal continuity of the foreground, we introduce the video segmentation constraint C into the foreground part of the video and get the following model:

$$\min_{L,S,E} \|L\|_\varphi + \lambda\|S\|_{C(2,1)} + \tau\|E\|_F^2 \quad s.t. X = L + S + E\,, \quad (4)$$

where X is the original video matrix, L is the low-rank background matrix, S is the moving foreground matrix, and E is the sparse noise matrix. $\|L\|_\varphi = \sum_{i=1}^{m} \frac{\sigma_i(L)}{\sqrt{\sigma_i(L)^2+\varphi}}$, $\|S\|_{C(2,1)} = \sum_{i=1}^{m} \sqrt{|C_i| \sum_{(j,k)\in C_i} ([S]_{jk})^2}$, $|C_i|$ represents the number of pixels in the group C_i, and $[\cdot]_{jk}$ represents the elements of the j-th row and the k-th column in the matrix.

Algorithm 1 Video segmentation

Input: a sequence of image frames$\{I_1, I_2, \cdots, I_N\}$;
Initialize: video pixel label matrix $P = 0$;
  while $k < N + 1$ **do**
    Using SLIC to calculate the super-pixel label vector SP_k of I_k;
    Using CTM to merge the super-pixels of I_k to get the pixel label vector P_k;
    if $k > 1$ **then**
      Merge the similar sub-regions in I_{k-1} and I_k;
      Update the sub-regions in I_k to the corresponding label of P_{k-1};
    end if
    Update the k-th column of P with P_k;
    $k = k + 1$;
  end while
Output: video pixel label matrix P

D. ALGORITHM

For model (4), we use the augmented Lagrange multiplier method to solve it. The augmented Lagrangian function of model (4) is as follows:

$$\begin{aligned}\mathcal{L}(L, S, E, Y, \mu) &= \|L\|_\varphi + \lambda\|S\|_{2,1} + \tau\|E\|_F^2 \\ &+\langle Y, L + S + E - X\rangle + \frac{\mu}{2}\|L + S + E - X\|_F^2\,, \quad (5)\end{aligned}$$

where $\langle\cdot,\cdot\rangle$ represents the inner product of two matrices, Y is the Lagrange multiplier and μ is a penalty parameter. Next, we will use the idea of alternating iterations to update the variables.

The basic iterative formula of augmented Lagrange multiplier method is:

$$L_{k+1} = arg\min_L \|L\|_\varphi + \frac{\mu_k}{2}\|L - (X - S_k - E\_k - \frac{Y_k}{\mu_k})\|_F^2\,, \quad (6)$$

$$\begin{aligned}S_{k+1} &= arg\min_S \frac{\lambda}{\mu_k}\|S\|_{C(2,1)} \\ &+\frac{1}{2}\|S - (X - L_{k+1} - E_k - \frac{Y_k}{\mu_k})\|_F^2\,, \quad (7)\end{aligned}$$

$$\begin{aligned}E_{k+1} &= arg\min_E \tau\|E\|_F^2 \\ &+\frac{\mu_k}{2}\|E - (X - L_{k+1} - S_{k+1} - \frac{Y_k}{\mu_k})\|_F^2\,, \quad (8)\end{aligned}$$

$$Y_{k+1} = Y_k + \mu_k(L_{k+1} - X + S_{k+1} + E_{k+1})\,, \quad (9)$$

$$\mu_{k+1} = \rho\mu_k\,. \quad (10)$$

where ρ is the step size. The purpose of ρ is to update the penalty parameter μ in each iteration.

1) SOLVING PROBLEM (6)

In order to solve problem (6), we need the following Theorem 1 [18].

Theorem 1: Let $G = Udiag(\sigma_G)V^T$ be the Singular Value Decomposition(SVD) of matrix $G \in R^{m\times n}$, where σ_G is the singular value of G. Let $F(Z) = f(\sigma(Z)) = f \circ \sigma(Z)$ be a unitarily invariant function and $\mu > 0$. Then the optimal solution of problem $\min_Z F(Z) + \frac{\mu}{2}\|Z - G\|_F^2$ is $Z^* = Udiag(\sigma^*)V^T$, where

$$\sigma^* = \underset{f,\mu}{prox}(\sigma_G) = arg\min_{\sigma\geq 0} f(\sigma) + \frac{\mu}{2}\|\sigma - \sigma_G\|_2^2 . \quad (11)$$

In (6), if we let $F(L) = \|L\|_\varphi$, $G = X - S_k - E_k - \frac{Y_k}{\mu_k}$, the following formula can be obtained according to Theorem 1:

$$L_{k+1} = Udiag(\sigma^*)V^T , \quad (12)$$

where σ^* is calculated by formula (11), and it can be obtained by the following iteration formula:

$$\sigma_{k+1} = max(0, \sigma_G - \frac{\partial f(\sigma_k)}{\mu_k}) . \quad (13)$$

2) SOLVING PROBLEM (7)

Using the following Theorem 2 proposed in [17], we can obtain the update formula of S.

Theorem 2: For a given matrix $M \in R^{m\times n}$ and a parameter $\tau > 0$, the following optimization problem

$$arg\min_S \tau\|S\|_{2,1} + \frac{1}{2}\|S - M\|_F^2 \quad (14)$$

has a closed-form solution $S^* = (S_1^*, \cdots, S_n^*)$, where

$$S_j^* = max(0, \frac{\|M_j\|_2 - \tau}{\|M_j\|_2})M_j, j = 1, \cdots, n \quad (15)$$

and M_j represents the j-th column of matrix M.

If we let $M = X - L_{k+1} - E_k - \frac{Y_k}{\mu_k}$, and define $\|[M]_{C_i}\|_2 = \sqrt{\sum_{(a,b)\in C_i}([M]_{ab})^2}$, then the solution formula of problem (7) is as follows:

$$[S]_{jk} = \begin{cases} \frac{\|[M]_{C_i}\|_2 - \frac{\lambda|C_i|}{\mu}}{\|[M]_{C_i}\|_2}[M]_{jk}, & if \ \|[M]_{C_i}\|_2 > \frac{\lambda|C_i|}{\mu}, \\ 0, & otherwise. \end{cases} \quad (16)$$

3) SOLVING PROBLEM (8)

With other parameters fixed, we can get the update formula of E as follows:

$$E_{k+1} = (1 + \frac{2\tau}{\mu_k})^{-1}(X - L_{k+1} - S_{k+1} - \frac{Y_k}{\mu_k}) \quad (17)$$

From the above conclusions, the algorithm framework for solving problem (4) can be given.

Algorithm 2 Augmented Lagrangian multiplier method for NCSC-RPCA

Input: The observed data matrix X, parameters $\lambda, \tau, \mu > 0$, the maximum iteration k_{max} and segmentation C;
Initialize: $E = 0$, $S = 0$, $Y = 0$, number of iteration $k = 0$;
while not converged **do**
Update L_{k+1} by (12);
Update S_{k+1} by (16);
Update E_{k+1} by (17);
Update Y_{k+1} by (9);
Update μ_{k+1} by (10);
Check the convergence condition: $k > k_{max}$ or $\frac{\|X - L_{k+1} - S_{k+1} - E_{k+1}\|_F^2}{\|X\|_F^2} \leq \varepsilon$;
end while
Output: The background matrix $L = L_{k+1}$, the foreground matrix $S = S_{k+1}$, the sparse noise matrix $S = S_{k+1}$.

IV. EXPERIMENTAL RESULTS AND COMPARISON

In this section, we apply the NCSC-RPCA model algorithm proposed in this paper to foreground extraction in dynamic background, and compare it with eight methods of RPCA-PCP [13], GoDec [45], 3WD [41], DECOLOR [53], MAMR [24], noncvxRPCA [17], WNNM-RPCA [15], LRSD-TNNSR [54]. This paper mainly selected the I2R dataset and CDnet2014 dataset of seven dynamic video sequences, which are fountain02, snowfall, boats, skating, wetsnow, overpass and watersurface. For each video sequence used, hundreds of consecutive frames were intercepted as the observation dataset. The methods proposed in this paper and the datasets and operating environments used in the eight comparison experiments are the same. All numerical experiments are based on PC Intel Core i5-7300hq 2.50ghz CPU, 8GB RAM environment, and MATLAB R2015b.

A. PARAMETER SETTINGS

In our model, we mainly use five parameters: μ, λ, τ, ρ and φ. In reference [55], the appropriate rules of value selection are given. In fountain02, snowfall, boats, skating, overpass and watersurface datasets, we take $\lambda = 10^{-3}$ and $\tau = 10^{-2}$. In wetsnow dataset, we select $\lambda = 10^{-3}$ and $\tau = 10^{-3}$. For penalty parameter μ, we take $\mu_0 = 3 \times 10^{-2}$. In order to accelerate the convergence speed of the algorithm, parameter $\rho = 1.1$. φ is a parameter of nonconvex function, and here $\varphi = 10^{-3}$ is chosen in the experiment. The stopping criteria for iteration are as follows:

$$Err \leq \varepsilon \ \ or \ \ Iter \geq I_{max} , \quad (18)$$

where $Err = \frac{\|X - L_{k+1} - S_{k+1} - E_{k+1}\|_F^2}{\|X\|_F^2}$, ε is the termination error of the pre-input (in our experiment, we set $\varepsilon = 10^{-3}$), Iter is the current number of iterations and I_{max} is the maximum number of iterations entered in advance (here, we set $I_{max} = 500$). In order to make different

TABLE 1. Datasets information used in experiments.

Dataset	Image dimension	Dataset frames	Challenges
fountain02	288×432	1499	Fountain with water fall
snowfall	480×720	6500	Falling snow with snow-covered cars
boats	240×320	7999	Shimmering water
skating	360×540	3900	Falling snow
wetsnow	540×720	3500	Dense falling snow with intermittent object
overpass	240×320	3000	Swaying leaves with slowly moving person
watersurface	128×160	633	Rippling water and lingering object

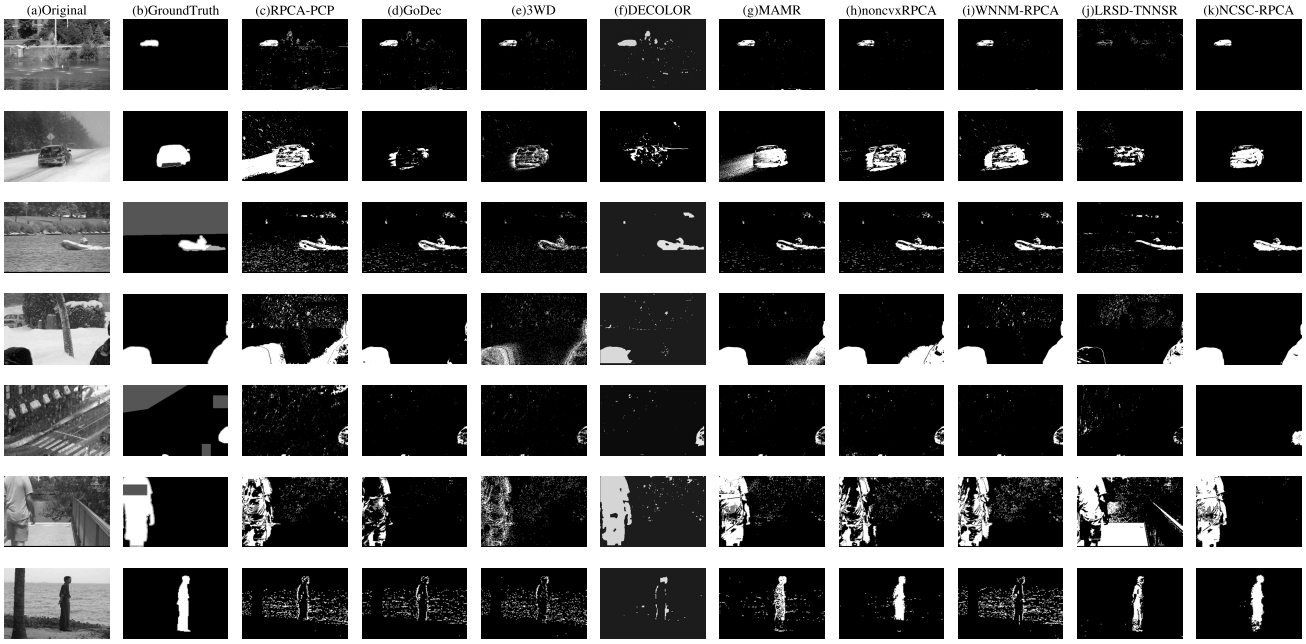

FIGURE 3. Comparison of visual effects of foreground extraction in dynamic background.

algorithms comparable, all methods adopt the same stopping criterion, that is, when the maximum number of iterations reaches 500 or the relative error is less than 10^{-3}, the calculation stops.

B. EVALUATION METRICS

Objectively, in order to evaluate the performance of each model more accurately, this paper referred to literature [30] and used recall, precision and their comprehensive evaluation index, F-measure, to evaluate the effect of dynamic video foreground detection. The formula is as follows:

$$Precision = \frac{TP}{TP + FP}, \tag{19}$$

$$Recall = \frac{TP}{TP + FN}, \tag{20}$$

$$F - measure = 2\frac{Precision \times Recall}{Precision + Recall}, \tag{21}$$

where TP=true positive indicates that the foreground pixels are correctly labeled as foreground, FP=false positive refers to the background pixels incorrectly labeled as foreground and FN=false negative corresponds the fact that foreground pixels are incorrectly labeled as background. All the evaluation indices are between 0 and 1. The higher the index value is, the more accurate the result is.

C. EXPERIMENTAL COMPARISON OF DYNAMIC BACKGROUND DATASETS

In order to verify the effectiveness of NCSC-RPCA model in moving object detection, this paper mainly selects seven dynamic video datasets for experiments. The specific experiments are shown in Figure 3.

Figure 3 shows the comparison results of visual effects of foreground extraction in dynamic background, where (a) represents original video frames selected from seven different dynamic background videos, (b) represents the ground truth of seven original video frames, (c) - (k) represent the foreground results extracted by different algorithms. Through the comparison, we can see that for seven dynamic background test datasets, the model proposed in this paper can

TABLE 2. Comparison of objective results of foreground detection by different methods (Best: bold. Second best: underline).

Dataset	Index	RPCA-PCP	GoDec	3WD	DECOLOR	MAMR	noncvx-RPCA	WNNM-RPCA	LRSD-TNNSR	NCSC-RPCA
fountain02	Re	0.915	0.905	0.729	**1**	0.818	0.720	0.767	0.204	0.903
	Pr	0.402	0.593	0.794	0.445	0.857	0.933	0.932	0.366	**1**
	Fm	0.559	0.716	0.760	0.616	0.837	0.813	0.841	0.262	**0.949**
snowfall	Re	0.640	0.203	0.308	0.175	0.716	0.536	0.630	0.374	**0.724**
	Pr	0.281	0.830	0.636	0.527	0.529	0.592	0.645	0.859	**0.995**
	Fm	0.391	0.326	0.415	0.263	0.609	0.563	0.637	0.521	**0.838**
boats	Re	0.721	0.722	0.418	0.735	0.546	0.660	0.627	0.310	**0.805**
	Pr	0.384	0.548	0.603	0.862	0.756	0.696	0.656	0.599	**0.952**
	Fm	0.501	0.623	0.494	0.794	0.634	0.678	0.641	0.409	**0.872**
skating	Re	0.838	0.599	0.478	0.524	0.926	0.949	**1**	0.407	0.994
	Pr	0.408	**0.999**	0.511	0.877	0.940	0.673	0.913	0.708	0.998
	Fm	0.549	0.749	0.494	0.654	0.933	0.787	0.954	0.517	**0.996**
wetsnow	Re	0.523	0.328	0.308	0.592	0.417	0.469	0.406	0.205	**0.811**
	Pr	0.262	0.814	0.568	0.648	0.717	0.677	0.721	0.331	**0.832**
	Fm	0.349	0.467	0.400	0.619	0.527	0.554	0.520	0.253	**0.822**
overpass	Re	0.577	0.298	0.306	**0.927**	0.735	0.445	0.578	0.541	0.808
	Pr	0.783	0.937	0.741	0.936	0.899	0.643	0.647	0.323	**0.993**
	Fm	0.664	0.453	0.433	**0.932**	0.809	0.526	0.611	0.405	0.891
watersurface	Re	0.129	0.096	0.093	0.133	0.560	0.692	0.125	0.605	**0.779**
	Pr	0.166	0.223	0.149	0.695	0.805	0.915	0.133	**0.988**	0.967
	Fm	0.145	0.134	0.114	0.223	0.660	0.788	0.129	0.750	**0.863**

TABLE 3. Comparison of foreground detection running time (in seconds) among different algorithms (Best: bold. Second best: underline).

Dataset	RPCA-PCP	GoDec	3WD	DECOLOR	MAMR	noncvx-RPCA	WNNM-RPCA	LRSD-TNNSR	NCSC-RPCA
fountain02	309.33	398.23	39.88	122.62	41.96	**1.7**	109.37	87.30	500.01
snowfall	860.66	902.30	95.71	453.12	118.75	**4.08**	255.10	224.77	328.02
boats	134.76	242.74	26.40	121.02	31.47	**1.09**	88.29	57.10	285.10
skating	587.48	766.11	65.39	572.95	78.30	**2.88**	183.98	143.82	322.01
wetsnow	578.63	917.63	60.86	305.87	105.37	**2.47**	116.39	194.54	522.94
overpass	174.73	232.18	22.34	297.26	26.68	**1.04**	63.34	51.92	340.89
watersurface	17.56	71.15	6.29	35.65	8.38	**0.25**	15.44	14.65	8.06

extract the foreground object more completely and can effectively remove part of the noise. Specifically, for the datasets of fountain02, boats, skating, wetsnow and overpass which are affected by the dynamic background, the foreground objects extracted by the eight comparison methods all contain different levels of noise, only our method can extract clean foreground targets. In the snowfall dataset, the targets extracted by RPCA-PCP, GoDec, 3WD, DECOLOR and LRSD-TNNSR all contain large holes. At the same time, there are İ°ghostsİś in the foreground extraction of MAMR, noncvxRPCA and WNNM-RPCA while our method can extract the foreground object completely and effectively eliminate the interference of dynamic snowflakes. In the watersurface dataset, RPCA-PCP, GoDec, 3WD, DECOLOR, WNNM-RPCA can not extract the foreground contour well and there are a lot of errors in judging the noise as the foreground. MAMR, noncvxRPCA and LRSD-TNNSR have the problem of classifying the foreground pixels as background pixels. In a word, compared with the other eight methods, our method has better foreground extraction effect.

Table 2 shows the relevant evaluation indices of our method and other comparison methods. From the perspective of f-measure value in fountain02, snowfall, boats, skating, wetsnow and watersurface datasets, our method is in the best position. In the overpass dataset, our f-measure value is slightly lower than DECOLOR. Besides, the recall and precision of our method also show the advantages, which shows that the model in this paper has a good comprehensive effect in suppressing the noise generated by dynamic background and the integrity and accuracy of target extraction.

Table 3 shows the running time of each algorithm. It is obvious that noncvxRPCA has the fastest running time in processing foreground extraction problems. RPCA-PCP and GoDec run slowly because they require a lot of singular

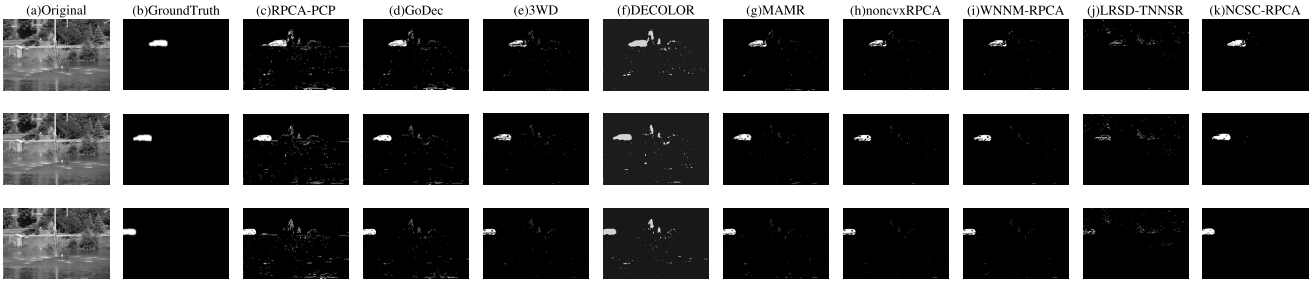

FIGURE 4. Experimental results comparison with different frames among different algorithms (fountain02).

TABLE 4. Comparison of objective results of representative frames of fountain02 dataset under different algorithms(Best:bold. Second best:underline).

Dataset	frame	Index	RPCA-PCP	GoDec	3WD	DECO-LOR	MAMR	noncvx-RPCA	WNNM RPCA	LRSD-TNNSR	NCSC-RPCA
fountain02	48	Re	0.959	0.823	0.532	**1**	0.626	0.517	0.548	0.150	0.756
		Pr	0.428	0.530	0.695	0.424	0.758	0.872	0.849	0.303	**0.972**
		Fm	0.592	0.645	0.603	0.596	0.686	0.649	0.666	0.201	**0.850**
	58	Re	0.961	0.945	0.740	**0.999**	0.858	0.745	0.796	0.268	0.913
		Pr	0.355	0.630	0.814	0.460	0.850	0.913	0.909	0.437	**0.987**
		Fm	0.519	0.756	0.775	0.630	0.854	0.821	0.849	0.332	**0.949**
	68	Re	0.851	0.951	0.741	**0.965**	0.871	0.782	0.825	0.240	0.929
		Pr	0.362	0.540	0.775	0.428	0.816	0.897	0.881	0.366	**0.998**
		Fm	0.508	0.689	0.758	0.593	0.843	0.836	0.852	0.290	**0.962**

value decomposition. Our method does not have the advantage in terms of time and the main time is spent in the process of video segmentation. Next, we select three arbitrary frames in the fountain02 dataset, namely the 48th, 58th and 68th frames, and take these three frames as test frames to compare our method with the other eight methods.

As can be seen in Figure 4, this dataset is affected by dynamic fountain. Hence, it is easy to wrongly extract the fountain as foreground. Among all the methods, we can see that LRSD-TNNSR can not extract the foreground object contour very well. There are mainly two reasons. On the one hand, the model uses truncated nuclear norm to describe the low rank part and introduces the sparse prior to the low rank part. On the other hand, the model still uses the L_1 norm to describe the foreground part, which leads to a not-so-good foreground extraction effect. At the same time, all the foreground images extracted by the eight comparison methods contain different degrees of fountain noise while our method can extract clean and almost complete foreground object.

Table 4 is a quantitative comparison of the three video frames selected in fountain02 dataset. As can be seen from Table 4, the f-measure value of our method is significantly better than the other eight methods. Table 2 and Table 4 quantify the validity of the model in this paper.

V. CONCLUSION

In order to overcome the shortcomings of traditional nuclear-norm-based RPCA model, such as the large amount of computation in the solution process and poor foreground extraction effect under dynamic background, this paper proposes an improved RPCA model based on a new non-convex rank approximation function and a segmentation constraint. The proposed new model is solved by the augmented Lagrangian multiplier method. Experimental results show that, compared with the existed methods, the improved model proposed in this paper can effectively suppress the noise generated by the dynamic background and the extracted moving objects have better comprehensive effect in the integrity and accuracy.

REFERENCES

[1] J. Kim, D. Yeom, and Y. Joo, "Fast and robust algorithm of tracking multiple moving objects for intelligent video surveillance systems," *IEEE Trans. Consum. Electron.*, vol. 57, no. 3, pp. 1165–1170, Aug. 2011.

[2] L. Li, D. Zhou, and Z. Liu, "A method of the moving object detection used in the military video monitor," *J. Transcluction Technol.*, vol. 2, 2004, doi: 10.1109/JLT.2003.821766.

[3] R. Aufrichtig and D. L. Wilson, "X-ray fluoroscopy spatio-temporal filtering with object detection," *IEEE Trans. Med. Imag.*, vol. 14, no. 4, pp. 733–746, Dec. 1995, doi: 10.1109/42.476114.

[4] N. Vaswani, T. Bouwmans, S. Javed, and P. Narayanamurthy, "Robust subspace learning: Robust PCA, robust subspace tracking, and robust subspace recovery," *IEEE Signal Process. Mag.*, vol. 35, no. 4, pp. 32–55, Jul. 2018, doi: 10.1109/MSP.2018.2826566.

[5] X. Liu, G. Zhao, J. Yao, and C. Qi, "Background subtraction based on low-rank and structured sparse decomposition," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2502–2514, Aug. 2015, doi: 10.1109/TIP.2015.2419084.

[6] S. Javed, A. Mahmood, T. Bouwmans, and S. K. Jung, "Spatiotemporal low-rank modeling for complex scene background initialization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 6, pp. 1315–1329, Jun. 2018, doi: 10.1109/TCSVT.2016.2632302.

[7] T. Bouwmans, A. Sobral, S. Javed, S. K. Jung, and E.-H. Zahzah, "Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset," *Comput. Sci. Rev.*, vol. 23, pp. 1–71, Feb. 2017.

[8] X. Cao, L. Yang, and X. Guo, "Total variation regularized RPCA for irregularly moving object detection under dynamic background," *IEEE Trans. Cybern.*, vol. 46, no. 4, pp. 1014–1027, Apr. 2016, doi: 10.1109/TCYB.2015.2419737.

[9] B. Shijila, A. J. Tom, and S. N. George, "Moving object detection by low rank approximation and l1-TV regularization on RPCA framework," *J. Vis. Commun. Image Represent.*, vol. 56, pp. 188–200, Oct. 2018.

[10] Y. Chen, X. Xiao, and Y. Zhou, "Low-rank quaternion approximation for color image processing," *IEEE Trans. Image Process.*, vol. 29, pp. 1426–1439, Sep. 2020, doi: 10.1109/TIP.2019.2941319.

[11] Y. Chen, Y. Guo, Y. Wang, D. Wang, C. Peng, and G. He, "Denoising of hyperspectral images using nonconvex low rank matrix approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5366–5380, Sep. 2017, doi: 10.1109/TGRS.2017.2706326.

[12] Y. Chen, S. Wang, and Y. Zhou, "Tensor nuclear norm-based low-rank approximation with total variation regularization," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 6, pp. 1364–1377, Dec. 2018, doi: 10.1109/JSTSP.2018.2873148.

[13] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis: Exact recovery of corrupted low-rank tensors via convex optimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Vancouver, BC, Canada, Jun. 2016, pp. 2080–2088.

[14] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1–37, May 2011, doi: 10.1145/1970392.1970395.

[15] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *Int. J. Comput. Vis.*, vol. 121, no. 2, pp. 183–208, Jul. 2016, doi: 10.1007/s11263-016-0930-5.

[16] Y. Hu, D. Zhang, J. Ye, X. Li, and X. He, "Fast and accurate matrix completion via truncated nuclear norm regularization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 9, pp. 2117–2130, Sep. 2013, doi: 10.1109/tpami.2012.271.

[17] Z. Kang, C. Peng, and Q. Cheng, "Robust PCA via nonconvex rank approximation," in *Proc. IEEE Int. Conf. Data Mining*, Atlantic City, NJ, USA, Nov. 2015. pp. 211–220.

[18] Y. Chen, Y. Wang, M. Li, and G. He, "Augmented lagrangian alternating direction method for low-rank minimization via non-convex approximation," *Signal, Image Video Process.*, vol. 11, no. 7, pp. 1271–1278, Apr. 2017.

[19] L. Yang, T. Pong, and X. Chen, "Alternating direction method of multipliers for nonconvex background/foreground extraction," Jun. 2015, *arXiv:1506.07029*. [Online]. Available: https://arxiv.org/abs/1506.07029

[20] P. Netrapalli, U. Niranjan, S. Sanghavi, A. Anandkumar, and P. Jain, "Non-convex robust PCA," in *Proc. Adv. Neural Inf. Process. Syst.*, Oct. 2014, pp. 1107–1115.

[21] M. Balcan, Y. Liang, Z. Song, D. Woodruff, and H. Zhang, "Non-convex matrix completion and related problems via strong duality," *J. Mach. Learn. Res.*, vol. 20, no. 102, pp. 1–56, Apr. 2019.

[22] Z. Yang, L. Fan, Y. Yang, Z. Yang, and G. Gui, "Generalized singular value thresholding operator based nonconvex low-rank and sparse decomposition for moving object detection," *J. Franklin Inst.*, vol. 356, no. 16, pp. 10138–10154, Nov. 2019.

[23] F. Wen, R. Ying, P. Liu, and T.-K. Truong, "Nonconvex regularized robust PCA using the proximal block coordinate descent algorithm," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5402–5416, Oct. 2019, doi: 10.1109/TSP.2019.2940121.

[24] J. Yang, X. Sun, X. Ye, and K. Li, "Background extraction from video sequences via motion-assisted matrix completion," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 2411–2437, doi: 10.1109/ICIP.2014.7025493.

[25] A. Sobral, T. Bouwmans, and E.-H. Zahzah, "Double-constrained RPCA based on saliency maps for foreground detection in automated maritime surveillance," in *Proc. 12th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Karlsruhe, Germany, Aug. 2015, pp. 1–6, doi: 10.1109/AVSS.2015.7301753.

[26] S. Ebadi, V. Ones, and E. Izquierdo, "Approximated robust principal component analysis for improved general scene background subtraction," *IEEE Trans. Image Process.*, to be published.

[27] S. Ebadi and E. Izquierdo, "Foreground segmentation via dynamic tree-structured sparse RPCA," in *Proc. Eur. Conf. Comput. Vis.*, Amsterdam, The Netherlands, Oct. 2016, pp. 314–329.

[28] S. Javed, S. H. Oh, A. Sobral, T. Bouwmans, and S. K. Jung, "Background subtraction via superpixel-based online matrix decomposition with structured foreground constraints," in *Proc. IEEE Int. Conf. Comput. Vis. Workshop (ICCVW)*, Santiago, Chile, Dec. 2015, pp. 90–98, doi: 10.1109/ICCVW.2015.123.

[29] S. Javed, A. Mahmood, S. Al-Maadeed, T. Bouwmans, and S. K. Jung, "Moving object detection in complex scene using spatiotemporal structured-sparse RPCA," *IEEE Trans. Image Process.*, vol. 28, no. 2, pp. 1007–1022, Feb. 2019, doi: 10.1109/TIP.2018.2874289.

[30] Y. Li, G. Liu, Q. Liu, Y. Sun, and S. Chen, "Moving object detection via segmentation and saliency constrained RPCA," *Neurocomputing*, vol. 323, pp. 352–362, Jan. 2019, doi: 10.1016/j.neucom.2018.10.012.

[31] L. Chen, X. Jiang, X. Liu, T. Kirubarajan, and Z. Zhou, "Outlier-robust moving object and background decomposition via structured ℓ_p-regularized low-rank representation," *IEEE Trans. Emerg. Topics Comput. Intell.*, to be published.

[32] X. Liu, J. Yao, X. Hong, X. Huang, Z. Zhou, C. Qi, and G. Zhao, "Background subtraction using spatio-temporal group sparsity recovery," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 8, pp. 1737–1751, Aug. 2018, doi: 10.1109/TCSVT.2017.2697972.

[33] A. Zheng, T. Zou, Y. Zhao, B. Jiang, J. Tang, and C. Li, "Background subtraction with multi-scale structured low-rank and sparse factorization," *Neurocomputing*, vol. 328, pp. 113–121, Feb. 2019, doi: 10.1016/j.neucom.2018.02.101.

[34] S. Javed, A. Mahmood, T. Bouwmans, and S. Jung, "Superpixels-based manifold structured sparse RPCA for moving object detection," in *Proc. Brit. Mach. Vis. Conf.* London, U.K.: Imperial College London, Sep. 2017, pp. 4–7.

[35] S. Javed, S. Ki Jung, A. Mahmood, and T. Bouwmans, "Motion-aware graph regularized RPCA for background modeling of complex scenes," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Cancun, Mexico, Dec. 2016, pp. 120–125, doi: 10.1109/ICPR.2016.7899619.

[36] S. Javed, A. Mahmood, T. Bouwmans, and S. K. Jung, "Background–foreground modeling based on spatiotemporal sparse subspace clustering," *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5840–5854, Dec. 2017, doi: 10.1109/TIP.2017.2746268.

[37] M. Sultana, A. Mahmood, S. Javed, and S. Jung, "Unsupervised RGBD video object segmentation using GANs," in *Proc. ACCV*, Perth, WA, Australia, Oct. 2018, pp. 1–15.

[38] M. Sultana, A. Mahmood, T. Bouwmans, and S. Jung, "Complete moving object detection in the context of robust subspace learning," in *Proc. 3rd Workshop Robust Subspace Learn. Comput. Vis. (ICCV)*, Seoul, South Korea, Oct. 2019.

[39] M. Sultana, A. Mahmood, S. Javed, and S. K. Jung, "Unsupervised deep context prediction for background estimation and foreground segmentation," *Mach. Vis. Appl.*, vol. 30, no. 3, pp. 375–395, Nov. 2018, doi: 10.1007/s00138-018-0993-0.

[40] Z. Zhou, X. Li, J. Wright, E. Candes, and Y. Ma, "Stable principal component pursuit," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 1518–1522, doi: 10.1109/ISIT.2010.5513535.

[41] O. Oreifej, X. Li, and M. Shah, "Simultaneous video stabilization and moving object detection in turbulence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 2, pp. 450–462, Feb. 2013, doi: 10.1109/TPAMI.2012.97.

[42] M. Jin, Y. Chen, and F. Zhang, "Tri-decomposition model for image recovery," *Electron. Lett.*, vol. 54, no. 23, pp. 1322–1324, Nov. 2018.

[43] H. Yang and S. Qu, "Real-time vehicle detection and counting in complex traffic scenes using background subtraction model with low-rank decomposition," *IET Intell. Transp. Syst.*, vol. 12, no. 1, pp. 75–85, Feb. 2018, doi: 10.1049/iet-its.2017.0047.

[44] L. Yin, A. Parekh, and I. Selesnick, "Stable principal component pursuit via convex analysis," *IEEE Trans. Signal Process.*, vol. 67, no. 10, pp. 2595–2607, May 2019, doi: 10.1109/TSP.2019.2907264.

[45] T. Zhou and D. Tao, "Godec: Randomized low-rank & sparse matrix decomposition in noisy case," *Proc. 28th Int. Conf. Mach. Learn.*, Bellevue, WA, USA, Jun. 2011, pp. 1–16.

[46] A. Fabjawska, "Normalized cuts and watersheds for image segmentation," in *Proc. IET Conf. Image Process. (IPR)*, London, U.K., Jul. 2012, p. 2, doi: 10.1049/cp.2012.0440.

[47] W. Tao, H. Jin, and Y. Zhang, "Color image segmentation based on mean shift and normalized cuts," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 37, no. 5, pp. 1382–1389, Oct. 2007, doi: 10.1109/TSMCB.2007.902249.

[48] A. Levinshtein, A. Stere, K. N. Kutulakos, D. J. Fleet, S. J. Dickinson, and K. Siddiqi, "TurboPixels: Fast superpixels using geometric flows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 12, pp. 2290–2297, Dec. 2009, doi: 10.1109/TPAMI.2009.96.
[49] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
[50] C. Lu, J. Tang, S. Yan, and Z. Lin, "Nonconvex nonsmooth low rank minimization via iteratively reweighted nuclear norm," *IEEE Trans. Image Process.*, vol. 25, no. 2, pp. 829–839, Feb. 2016, doi: 10.1109/TIP.2015.2511584.
[51] A. Y. Yang, J. Wright, Y. Ma, and S. S. Sastry, "Unsupervised segmentation of natural images via lossy data compression," *Comput. Vis. Image Understand.*, vol. 110, no. 2, pp. 212–225, May 2008, doi: 10.1016/j.cviu.2007.07.005.
[52] Y. Ma, H. Derksen, W. Hong, and J. Wright, "Segmentation of multivariate mixed data via lossy data coding and compression," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 9, pp. 1546–1562, Sep. 2007, doi: 10.1109/TPAMI.2007.1085.
[53] X. Zhou, C. Yang, and W. Yu, "Moving object detection by detecting contiguous outliers in the low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 597–610, Mar. 2013, doi: 10.1109/TPAMI.2012.132.
[54] Z. Xue, J. Dong, Y. Zhao, C. Liu, and R. Chellali, "Low-rank and sparse matrix decomposition via the truncated nuclear norm and a sparse regularizer," *Vis. Comput.*, vol. 35, no. 11, pp. 1549–1566, May 2018, doi: 10.1007/s00371-018-1555-1.
[55] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. Roy. Stat. Soc., B, Stat. Methodol.*, vol. 68, no. 1, pp. 49–67, Feb. 2006.

ZIXUAN HU received the B.S. degree in applied statistics from Binzhou University, China, in 2017. She is currently pursuing the M.S. degree with the School of Mathematics and Systems Science, Shandong University of Science and Technology. Her research interests include nonconvex low-rank matrix decomposition, moving object detection, and image processing.

YONGLI WANG received the B.S. degree in operations research and control theory from the Shandong University of Science and Technology, Qingdao, China, in 2001, and the Ph.D. degree in applied mathematics from Shanghai Jiao Tong University, Shanghai, China, in 2006. Her research interests include optimization theory and method, low-rank representation, image processing, and distributed computation.

RUI SU received the B.S. degree in applied mathematics from the Shandong University of Science and Technology, Qingdao, China, in 2015, and the M.S. degree from the School of Mathematics and Systems Science, Shandong University of Science and Technology, in 2018. His research interests include Bayesian networks, actuarial insurance, data mining, and image processing.

XINXIN BIAN received the B.S. degree in applied statistics from Binzhou University, China, in 2017. She is currently pursuing the M.S. degree with the School of Mathematics and Systems Science, Shandong University of Science and Technology. Her research interests include low-rank and sparse matrix decomposition models, moving object detection, and background recovery.

HONGCHAO WEI received the B.S. degree in information and computing science from Qingdao Agricultural University, China, in 2018. He is currently pursuing the M.S. degree with the School of Mathematics and Systems Science, Shandong University of Science and Technology. His research interests include image segmentation and clustering algorithms.

GUOPING HE received the B.S. degree in computational mathematics from the Shandong University of Science and Technology, Qingdao, China, in 1982, and the M.S. and Ph.D. degrees in operations research and control theory from the Chinese Academy of Sciences, Beijing, China, in 1988 and 1995, respectively. He is currently the Vice President of the Qilu University of Technology, Jinan, China. He has over 100 research articles and is the holder of eight foundations. His research interests include nonlinear optimization theory, numerical computation, and data mining.