ABSTRACT

Low-rank and sparse decomposition (LRSD) has attracted wide attention in video foreground-background separation and many other fields. However, the traditional LRSD methods have many tough problems, such as the problems of the low accuracy of the surrogate functions of rank and sparsity, ignoring the spatial information of the videos and sensitivity to noise, etc. To deal with these problems, this paper proposes the generalized nuclear norm and structured sparse norm (GNNSSN) method based LRSD for video foreground-background separation, which introduces the generalized nuclear norm (GNN) and the structured sparse norm (SSN) to approximate the rank function and the $l_0$-norm of the LRSD method. In addition, we extend our proposed model to a robust model against noise for practical applications, and we called the extended method as the robust generalized nuclear norm and structured sparse norm (RGNNSSN) method. At last, we use the alternating direction method of multipliers (ADMM) to solve our proposed two methods. Experimental results and discussions on video foreground-background separation demonstrate that our proposed two methods have better performances than other LRSD based foreground-background separation methods.

INDEX TERMS

Low-rank and sparse decomposition, generalized nuclear norm, structured sparse norm, alternating direction method of multipliers, foreground-background separation.
2) it does not handle square matrices well; 3) the principal elements obtained by PCA method are not optimal; 4) it does not consider the time and spatial information of videos perfectly; 5) it assumes that variables obey Gaussian distribution, when variables do not obey Gaussian distribution, scaling and rotation will occur. To deal with these problems, many improved methods [13], [14] have been proposed, however, none of them can solve the problems of assumption that variables in PCA method must obey Gaussian distribution.

In order to solve this problem of PCA method, low-rank and sparse decomposition (LRSD) method [15]–[20] has been proposed and used widely in many fields.

LRSD is a major and widely method which poses a big challenge in the video foreground-background separation. The LRSD decomposes the observation matrix M into low-rank and sparse matrices, and its model is shown as follows:

\[
\min_{B,F} \text{rank}(B) + \tau \|F\|_0 \\
\text{s.t. } M = B + F
\]  

(1)

where \(\text{rank}(B)\) is the rank of the matrix \(B \in \mathbb{R}^{m \times n}\), \(\|F\|_0\) is the \(l_0\)-norm of the matrix \(F \in \mathbb{R}^{m \times n}\), i.e., is the number of non-zero elements of the matrix \(F\), and \(\tau > 0\) is the penalty factor. However, as we all know that the equation (1) is an NP-hard problem, which is difficult to solve directly. The traditional method to deal with this problem is the principal component pursuit (PCP) method [20], [21]. This method uses the nuclear norm and the \(l_1\)-norm to approximate the rank function and the \(l_0\)-norm of the LRSD model, respectively, and it can be shown as follows:

\[
\min_{B,F} \|B\|_* + \tau \|F\|_1 \\
\text{s.t. } M = B + F
\]  

(2)

where \(\|B\|_*\) is the nuclear norm of the matrix of the matrix \(B\), i.e., the sum of the singular values of the matrix \(B\), \(\|F\|_1\) is the \(l_1\)-norm of the matrix \(F\), and \(\tau > 0\) is the penalty factor.

Recently, researchers have shown that the PCP method has strong theoretical guarantees for the video foreground-background separation [18], endoscopic images highlight removing and so on. For example, Li et al. [22] proposed Adaptive-RPCA which can remove the specular reflections in endoscopic image sequences well. However, researchers also found that the PCP method has many defects [19], [20] as follows: 1) the nuclear norm in the PCP method is the \(l_1\)-norm of the singular value essentially, which processes the singular value uniformly, and always generates the problem of over-punishment of the singular value; 2) the calculation cost of the PCP method is relatively high; 3) the surrogate function in the PCP method is a biased estimate which has low approximation accuracy. 4) the PCP method always ignores the spatial information of video; 5) when the video contains noise, the effectiveness of the PCP method is limited.

In recent years, many enhanced methods have been proposed to address these problems of the PCP method. Hu et al. proposed the truncated nuclear norm method (TNN) [23] for the problem of over-penalty of the PCP method. This method uses the truncated kernel norm to approximate the rank function of the LRSD method, and the experimental results have shown that the TNN method has better performances than the PCP method in some scenarios. However, some smaller singular values are discarded in the TNN method, so it is easy to be lost some information of video; In order to deal with the high complexity and noise sensitivity of the PCP method, Zhou et al. proposed the Go Decomposition method (GoDec) [24]. This method changes the LRSD model to foreground, background, and noise items, and also uses the bilateral random projections (BRP) to replace the singular value decomposition (SVD) in the LRSD method. A large number of experiments for the GoDec method have shown that the method reduces the calculation cost, however, the performance of this method is poor in some complex scenarios of video foreground-background separation.

Recently, to deal with the problem of the low accuracy of the convex surrogate function in the PCP method, researchers have gradually studied the non-convex surrogate function [25]–[38]. For example, Yang et al. adopt a general dual step-size to solve the non-convex and non-smooth problem which is called NNP method [32]. Wen et al. proposed a Low-rank and sparse decomposition (LRSD) based the generalized non-convex regularization (GNR) method and applied it to video foreground-background separation [35]. Ma et al. [39] showed that the nuclear norm minimization is the convex relaxation of low rank minimization which leads to good denoising results and applied the weighted nuclear norm minimization to image denoising. Based on this fact, Yang et al. proposed the nonconvex nonsmooth weighted nuclear norm (NNWNN) [25] and the non-convex low-rank and sparse decomposition (NonLRSD) methods [28]. The NNWNN and the NonLRSD methods use nonconvex nonsmooth weighted nuclear norm and the non-convex generalized singular value threshold operator to approximate the rank function in the LRSD model, respectively. Lots of simulation experiments have shown that the effectiveness of the video foreground-background separation of the NNWNN and NonLRSD methods are better than the other methods which using convex surrogate function.

The nuclear norm based methods always consider few spatial connections of videos. These methods may fail in real scenarios, especially for handling a dynamic background. Mairal et al. introduced the optimization for structured sparsity [40]. Liu et al. proposed the low-rank and structured sparse decomposition method (LSD) [41]. This method uses a structured sparse induction norm (SSIN) to approximate the sparsity function which adds some structured information when handling the background regions or foreground movements with varying scales. Experimental results have shown that the LSD method has better performances than other methods especially for some practical scenarios with background movements which are with influence of lighting changes, background fluctuations, etc. Ye et al. proposed...
the motion-assisted matrix restoration (MAMR) [42] method for video foreground-background separation, and extended this method to the robust motion-assisted matrix restoration method (RMAMR) [42]. The two methods introduced the motion information to the LRSD method. The experiments have shown that the two methods not only have better performances for the video foreground-background separation, but also have strong anti-noise characteristics. Although significant progress has been gained during these years, more researches work will be further made.

This paper proposes the generalized nuclear norm and structured sparse norm (GNNSSN) method based LRSD which uses the generalized nuclear norm (GNN) and the structured sparse norm (SSN) [35][38] to approximate the rank function and the \( l_0 \)-norm of the LRSD model, respectively. In addition, in order to enhance the robust of the GNNSSN method for videos with noise, we extend this method to the robust generalized nuclear norm and structured sparse norm (RGNNSSN) model. Finally, we use the alternating direction method of multipliers (ADMM) [38], [40] to solve the above two proposed models and apply the two methods to the video foreground-background separation without noise and with noise for simulation experiments. Experiments have shown that our proposed methods have higher superiority than other LRSD based methods.

The rest of this paper is organized as follows. The second section of this paper proposes the GNNSSN and RGNNSSN models, and uses the alternating direction method of multipliers to solve the above two proposed models. In the third section, simulation experiments are carried out and the experimental results are exhibited and analyzed. At last, in the fourth section, we give the summarized remarks.

II. THE PROPOSED TWO METHODS

A. THE GENERALIZED NUCLEAR NORM AND STRUCTURED SPARSE NORM (GNNSSN) METHOD

In this section, in order to enhance the approximation of low-rank components in the LRSD model, this paper uses a generalized non-convex proxy function \( \min_{i=1}^{\min(m,n)} g(\sigma_i(B)) \) to approximate the rank function of the LRSD model, which is called generalized nuclear norm (GNN). In addition, we use the structured sparse norm (SSN) [40], [41] to approximate \( l_0 \)-norm of the LRSD model for modeling the spatial information in sparse outliers. The SSN is expressed as \( \Phi(F) \). Therefore, the GNNSSN model proposed in this paper can be expressed as follows:

\[
\Phi(F) = \sum_{d \in D} \max_{j \in d} |F_{j}| = \sum_{d \in D} \|F_d\|_\infty \tag{4}
\]

where \( F \) is a sparse signal that satisfies a structured distribution, \( F_j \) is the value of the \( j \)-th element of \( F \), \( \| . \|_\infty \) is the infinite norm, i.e. the maximum absolute value of all elements. \( D \) is a collection of predefined group distributions, where each group distribution is \( d \), and \( F_d \) is a subset of \( F \) with group distribution \( d \).

In the following, we use the alternating direction method of multipliers (ADMM) method to solve the GNNSSN model. The augmented Lagrangian function corresponding to problem (3) is given as follows:

\[
\mathcal{L}(B, F, Y, \mu) = \min_{i=1}^{\min(m,n)} g(\sigma_i(B)) + \tau \Phi(F)
\]

\[
- \langle Y, B + F - M \rangle + \frac{\mu}{2} \|B + F - M\|_F^2 \tag{5}
\]

The solution of the above problem can be transformed to the following problem:

\[
\text{prox}_{\lambda g}^\tau(E) = \min_B \lambda \sum_{i=1}^{m} g(\sigma_i(B)) + \frac{1}{2} \|B - E\|_F^2 \tag{7}
\]

where \( E = M - F^k + \frac{Y_k}{\mu_k} \). The above problem can be solved by the following formulation:

\[
\min_{\sigma_i(B) \geq \cdots \geq \sigma_m(B) \geq 0} \left( \lambda g(\sigma_i(B)) + \frac{1}{2} (\sigma_i(B) - \sigma_i(E))^2 \right) \tag{8}
\]

where \( \sigma_i(B) \geq \cdots \geq \sigma_m(B) \geq 0 \) is the singular values of \( B \). It is equal to solve the following problem for each \( e = \sigma_i(E) \):

\[
\text{prox}_{\lambda g}(e) = \min_{b \geq 0} \lambda g(b) + \frac{1}{2} (b - e)^2 \tag{9}
\]
An optimal solution of above problem is $\hat{b} = 0$ or $\hat{b} = \max \{ b | \hat{a}g(b) + b - e = 0, 0 \leq b \leq e \}$, which can be solved by the method in [36], where $\hat{a}(\cdot)$ denotes the subgradient operator.

In order to give the details of the proximal mapping of the concave $g(\cdot)$ function, we introduce the proximity operator for three popular concave $g(\cdot)$ functions, including the $l_p - \text{norm}(0 < p < 1)$, the SCAD and the MCP functions.

(1) $l_p - \text{norm}(0 < p < 1)$: the function of the $l_p - \text{norm}(0 < p < 1)$ is given as follows:

$$g(b) = |b|^p, \quad 0 < p < 1$$

Except for two known cases of $p = 1/2$ and $p = 2/3$, the $l_p - \text{norm}(0 < p < 1)$ does not has a closed-form expression, but it can be solved as:

$$\text{prox}_{\lambda g}(e) = \begin{cases} 0, & |e| < \nu \\ \{0, \text{sign}(e)\beta\}, & |e| = \nu \\ \text{sign}(e)y^*, & |e| > \nu \end{cases}$$

where $\beta = [2(1 - p)\lambda]^{1/(2-p)}$, $\nu = \beta + p\lambda\beta^{p-1}$, $y^* \in (\beta, |e|)$ is the solution of $h(y) = py^{p-1} + y/\lambda - |e|/\lambda = 0$. Because $h(y)$ is the convex and differentiable, in the case of $|e| > \nu$, $y^*$ can be efficiently computed by an iterative algorithm.

(2) SCAD: the function of the SCAD is given as follows:

$$g(b) = \begin{cases} 0, & |e| < \nu \\ \{0, \text{sign}(e)\beta\}, & |e| = \nu \\ \text{sign}(e)y^*, & |e| > \nu \end{cases}$$

where $\beta = [2(1 - p)\lambda]^{1/(2-p)}$, $\nu = \beta + p\lambda\beta^{p-1}$, $y^* \in (\beta, |e|)$ is the solution of $h(y) = py^{p-1} + y/\lambda - |e|/\lambda = 0$. Because $h(y)$ is the convex and differentiable, in the case of $|e| > \nu$, $y^*$ can be efficiently computed by an iterative algorithm.

(3) MCP: the function of the MCP is given as follows:

$$g(b) = \eta \int_0^{|b|} \max(1 - t/\gamma \eta, 0) dt$$

where $\gamma > 1$. The proximal mapping of the MCP function is:

$$\text{prox}_{\lambda g}(e) = \begin{cases} 0, & |e| < \lambda \eta \\ \text{sign}(e)(|e| - \lambda \eta), & \lambda \eta < |e| \leq \alpha \eta \\ e, & |e| > \alpha \eta \end{cases}$$

In summary, the problem (6) can be solved by the following formulation:

$$B^{k+1} = \text{prox}_{\lambda g}(M - F^k + \frac{Y^k}{\mu_k})$$

$$= \text{U}^k \text{Diag}(\text{prox}_{\frac{\lambda}{\mu_k}}(\sigma(M - F^k + \frac{Y^k}{\mu_k}))) (Y^k)^T$$

Second, fixing $B$, $Y$ and $\mu$, and updating $F$, we can get $F^{k+1}$ as:

$$F^{k+1} = \arg \min_{\text{arg min } F} \mathcal{L}(B^{k+1}, F, Y^k, \mu_k)$$

$$= \arg \min_{\text{arg min } F} \tau \Phi(F) - \langle Y^k, B^{k+1} + F - M \rangle$$

$$+ \frac{\mu_k}{2} \| B^{k+1} + F - M \|_F^2$$

$$= \arg \min_{\text{arg min } F} \tau \frac{\mu_k}{2} \| F - (M - B^{k+1} + \frac{Y^k}{\mu_k}) \|_F^2$$

Algorithm 1 The GNNSSN Method

1. **Initialization:** Given $\tau > 0$, $\mu_0 > 0$, $\mu_{\text{max}} > \mu_0$, $\rho > 1$, starting points $B^0 = 0$, $F^0 = 0$, $Y_0 = \frac{M}{\max(||I_n|, \sqrt{\text{nnz}(M)\|M\|_{\infty}})}$, and iteration index $k = 0$;

2. **Update B:** Updating $B^{k+1}$ according equation (9);

3. **Update F:** Updating $F^{k+1}$ according equation (12);

4. **Update Y:** Updating $Y^{k+1}$ according equation (13);

5. **Update \mu:** Updating $\mu_{k+1}$ according equation (14);

6. **Terminate or set \text{set } k := k + 1:** If the termination conditions $\frac{\|M - B^{k+1} - F^{k+1} - \mu_{k+1} \|_F}{\|M\|_F} \leq \varepsilon$ are met, the process iteration terminates, where $\varepsilon$ is a small value. Else, returns to step (2).
### B. The Robust Generalized Nuclear Norm and Structured Sparse Norm (RGNNSSN) Method

Noise is always appeared in the actual videos, so we extend the GNNSSN method to the RGNNSSN method and apply the proposed method to noisy video foreground-background separation. The model of RGNNSSN method is shown as follows:

\[
\min_{B,F} \sum_{i=1}^{\min(m,n)} g(\sigma_i(B)) + \tau_1 \Phi(F) + \tau_2 \|M - B - F\|_F^2 \tag{15}
\]

where \( M \) is observation matrix corresponding to video, \( B \) is the background of the video, \( F \) is the foreground of the video, \( \tau_1, \tau_2 \) is the penalty factor, and \( \| \cdot \|_F \) is the Frobenius norm. When \( \tau_2 \to \infty \), the model (15) degenerates into the model (3). Model (15) can be transformed into a constraint problem as shown in equation (16):

\[
\min_{B,F} \sum_{i=1}^{\min(m,n)} g(\sigma_i(B)) + \tau_1 \Phi(F) + \tau_2 \|G\|_F^2 \\
\text{s.t. } M = B + F \tag{16}
\]

In the following section, we use the alternating direction method of multipliers (ADMM) to solve the proposed RGNNSSN model. The augmented Lagrangian function of the model (16) is given as follows:

\[
\mathcal{L}(B, F, G, Y, \mu) = \min_{\min(m,n)} \sum_{i=1}^{\min(m,n)} g(\sigma_i(B)) + \tau_1 \Phi(F) + \tau_2 \|G\|_F^2 \\
+ \mu_k (B + F + G - M) \tag{17}
\]

The solution of problem (17) is divided into the following steps:

First, fixing \( B, G, Y \) and \( \mu \), and updating \( B \), we can get \( B^{k+1} \) as:

\[
B^{k+1} = \arg\min_B \mathcal{L}(B, F^k, G^k, Y^k, \mu_k) \\
= \arg\min_B \sum_{i=1}^{\min(m,n)} g(\sigma_i(B)) - \langle Y^k, B + Y^k + G^k - M \rangle \\
+ \frac{\mu_k}{2} \|B + F^k + G^k - M\|_F^2 \\
= \arg\min_B \sum_{i=1}^{\min(m,n)} g(\sigma_i(B)) \\
+ \frac{\mu_k}{2} \|B - (M - F^k - G^k + \frac{Y^k}{\mu_k})\|_F^2 \\
= \text{prox}_{\frac{\mu_k}{2} \| \cdot \|_F^2} (M - F^k - G^k + \frac{Y^k}{\mu_k}) \tag{18}
\]

Second, fixing \( B, G, Y \) and \( \mu \), and updating \( F \), we can get \( F^{k+1} \) as:

\[
F^{k+1} = \arg\min_F \mathcal{L}(B^{k+1}, F, G^k, Y^k, \mu_k) \\
= \arg\min_F \tau_1 \Phi(F) - \langle Y^k, B^{k+1} + F + G^k + M \rangle \\
+ \frac{\mu_k}{2} \|B^{k+1} + F + G^k - M\|_F^2 \\
= \tau_1 \Phi(F) \\
+ \frac{\mu_k}{2} \|F - (M - B^{k+1} - G^k - \frac{Y^k}{\mu_k})\|_F^2 \\
= \text{prox}_{\frac{\mu_k}{2} \| \cdot \|_F^2} (M - B^{k+1} - G^k + \frac{Y^k}{\mu_k}) \tag{19}
\]

Then, \( B, F, Y \) and \( \mu \), and updating \( G \), we can get \( G^{k+1} \) as:

\[
G^{k+1} = \arg\min_G \mathcal{L}(B^{k+1}, F^{k+1}, G, Y^k, \mu_k) \\
= \arg\min_G \tau_2 \|G\|_F^2 - \langle Y^k, B^{k+1} + F^{k+1} + G - M \rangle \\
+ \frac{\mu_k}{2} \|B^{k+1} + F^{k+1} + G - M\|_F^2 \\
= \tau_2 \|G\|_F^2 \\
+ \frac{\mu_k}{2} \|G - (M - B^{k+1} - F^{k+1} + \frac{Y^k}{\mu_k})\|_F^2 \\
= \text{prox}_{\frac{\mu_k}{2} \| \cdot \|_F^2} (M - B^{k+1} - F^{k+1} + \frac{Y^k}{\mu_k}) \tag{20}
\]

Eq. (20) can be solved by using the following closed expression:

\[
G^{k+1} = \frac{\mu_k}{2 \mu_k + \tau_2} (M - B^{k+1} - F^{k+1} + \frac{Y^k}{\mu_k}) \tag{21}
\]

Finally, the multiplier \( Y \) and penalty parameters \( \mu \) are updated as follows:

\[
Y^{k+1} = Y^k - \mu_k (B^{k+1} + F^{k+1} + G^{k+1} - M) \tag{22}
\]

\[
\mu_{k+1} = \min \{ \rho \mu_k, \mu_{\text{max}} \} \tag{23}
\]

In summary, the steps of the RGNNSSN method are given as follows:
C. CONVERGENCE ANALYSIS FOR THE PROPOSED ALGORITHMS

In this section, we give the convergence analysis of our proposed algorithms.

Lemma 1: Let \( \{ (B^k, F^k, Y^k) \} \) be a sequence generated by the GNNSSN algorithm, if \( \sum_{k=1}^{\infty} \mu_k < +\infty \), then \( \{ (B^k, F^k, Y^k) \} \) is bounded.

Proof: The \( k \)-th iterate generated by the GNNSSN algorithm is characterized by the following formulation:

\[
0 \in \partial \left( \sum_{i=1}^{\min(m,n)} g\left( \sigma_i \left( B^{k+1} \right) \right) \right) - \left[ Y^k - \mu_k \left( B^{k+1} + F^k - M \right) \right]
\]

\[ Y^{k+1} = Y^k - \mu_k \left( B^{k+1} + F^k - M \right) \]

Eq.(24) can be transformed to the Eq. (25).

\[
0 \in \partial \left( \sum_{i=1}^{\min(m,n)} g\left( \sigma_i \left( B^{k+1} \right) \right) \right) - \left[ Y^k - \mu_k \left( B^{k+1} + F^k - M \right) \right]
\]

\[ Y^{k+1} = Y^k - \mu_k \left( B^{k+1} + F^k - M \right) \]

(25)

we can get \( Y^{k+1} \in \partial \left( \tau \Phi \left( F^{k+1} \right) \right) \) based second equation of (25).

Because \( \Phi \left( F^{k+1} \right) = \sum_{d \in D} F_{d}^{k+1} \| \infty \) is bounded, the \( \{ Y^{k+1} \} \) sequence is bounded. On the other hand, we have

\[
L \left( B^{k+1}, F^{k+1}, Y^k, \mu_k \right)
\]

\[ L \left( B^{k+1}, F^k, Y^k, \mu_k \right) \leq L \left( B^{k+1}, F^k, Y^k, \mu_k \right) \]

\[ = L \left( B^{k+1}, F^k, Y^{k-1}, \mu_{k-1} \right) + \left( Y^{k-1} - Y^k, B^k + F^k - M \right) \]

\[ + \frac{\mu_k - \mu_{k-1}}{2} \left\| B^k + F^k - M \right\|_F^2 \]

\[ = L \left( B^{k+1}, F^k, Y^{k-1}, \mu_{k-1} \right) + \frac{1}{2} \left\| Y^{k-1} - Y^k \right\|_F^2 \]

\[ + \frac{\mu_k - \mu_{k-1}}{2 \mu_{k-1}} \left\| Y^{k-1} - Y^k \right\|_F^2 \]

\[ = L \left( B^{k+1}, F^k, Y^{k-1}, \mu_{k-1} \right) + \frac{\mu_k + \mu_{k-1}}{2 \mu_{k-1}^{2}} \left\| Y^{k-1} - Y^k \right\|_F^2 \]

(26)

Because of the bounded \( \{ Y^k \} \), we have

\[
\sum_{k=1}^{\infty} \frac{\mu_k}{2} + \frac{\mu_{k-1}}{2} \leq \sum_{k=1}^{\infty} \frac{2 \mu_k}{2 \mu_{k-1}} = \sum_{k=1}^{\infty} \frac{\mu_k}{\mu_{k-1}} < +\infty
\]

(27)

We can get \( \{ L \left( B^{k+1}, F^{k+1}, Y^k, \mu_k \right) \} \) is upper bounded. Because \( Y^k = Y^{k-1} - \mu_{k-1} \left( B^k + F^k - M \right) \), and then we can get:

\[
\sum_{i=1}^{\min(m,n)} g\left( \sigma_i \left( B^k \right) \right) + \tau \Phi \left( F^k \right)
\]

\[ = L \left( B^k, F^k, Y^{k-1}, \mu_{k-1} \right) + \left( Y^{k-1} - 1, B^k + F^k - M \right) \]

\[ - \frac{\mu_{k-1}}{2} \left\| B^k + F^k - M \right\|_F^2 \]

\[ = L \left( B^k, F^k, Y^{k-1}, \mu_{k-1} \right) + \left( Y^{k-1} - \frac{1}{\mu_{k-1}} \right) \left( Y^{k-1} - Y^k \right) \]

\[ - \frac{1}{2 \mu_{k-1}} \left\| Y^{k-1} - Y^k \right\|_F^2 \]

\[ = L \left( B^k, F^k, Y^{k-1}, \mu_{k-1} \right) - \frac{1}{2 \mu_{k-1}} \left\| Y^k \right\|_F^2 - \frac{1}{2 \mu_{k-1}} \left\| Y^{k-1} \right\|_F^2 \]

(28)

which is also upper bounded. So \( \{ B^k \} \) and \( \{ F^k \} \) are bounded. Therefore, the sequence \( \{ (B^k, F^k, Y^k) \} \) is bounded.

Theorem 2: Let \( \{ \mu_k \} \) \( (\mu_k > 0) \) be an increasing sequence and \( \{ (B^k, F^k, Y^k) \} \) be a sequence generated by GNNSSN algorithm, if \( \sum_{k=1}^{\infty} \mu_k < +\infty \), and then any accumulation point \( (B^*, F^*, Y^*) \) of the sequences \( \{ (B^k, F^k, Y^k) \} \) is a stationary point of problem (3).

Proof: The sequence \( (B^k, F^k, Y^k) \) generated by GNNSSN algorithm is bounded, so there is an accumulation point \( (B^*, F^*, Y^*) \) and a convergent subsequence \( (B^j, F^j, Y^j) \) such that

\[
\lim_{j \to +\infty} (B^j, F^j, Y^j) = (B^*, F^*, Y^*)
\]

From the third equation of (25), we can get

\[
B^{k+1} + F^k - M = \frac{1}{\mu_k} (Y^k - Y^{k-1})
\]

Besides, the fact that \( \{ Y^k \} \) is bounded, and \( \{ \mu_k \} \) is the increasing sequence, we have \( \lim_{j \to +\infty} (B^j, F^j, Y^j) = 0 \), that is \( (B^*, F^*, Y^*) = 0 \).

Because \( \lim_{j \to +\infty} F^j = F^* \), \( \lim_{j \to +\infty} B^j = B^* \), and \( \min(m,n) \), we get

\[
\sum_{i=1}^{\min(m,n)} g\left( \sigma_i \left( B^k \right) \right) \text{ is convex, there exists } N^{k+1} \in \partial \left( \sum_{i=1}^{\min(m,n)} g\left( \sigma_i \left( B^j \right) \right) \right)
\]

\[ - \frac{\mu_k}{2 \mu_{k-1}} \left\| Y^{k-1} - Y^k \right\|_F^2 \]

\[ = \left( \frac{\mu_k + \mu_{k-1}}{2 \mu_{k-1}^{2}} \left\| Y^{k-1} - Y^k \right\|_F^2 \right) \]

(26)

such that

\[
N^{k+1} - Y^k + \mu_k \left( B^{k+1} + F^* - M \right) = 0
\]

Let \( j \to \infty \), based the upper semi-continuous property of the subdifferential, there exists \( N^* \in \partial \left( \sum_{i=1}^{\min(m,n)} g\left( \sigma_i \left( B^* \right) \right) \right) \) such that

\[
N^* - Y^* + \mu_k (B^* + F^* - M) = 0, \text{ that is } N^* - Y^* = 0.
\]

Thus \( B^* \) is a stationary point of the first equation of (24).
In a similar way, we get $F^*$ is a stationary point of the second equation of (24). Thus we can get that:

$$
\begin{align*}
0 & \in - \partial \left( - \sum_{i=1}^{\min(m, n)} g \left( \sigma_i \left( B^* \right) \right) \right) - Y^* \\
0 & \in \partial (\lambda \| F \|_1) - Y^* \\
B^* + F^* - M = 0
\end{align*}
$$

(29)

Therefore, $(B^*, F^*, Y^*)$ is a stationary point of problem (3).

In the same way, we can also prove the convergence of the RGNSSN algorithm according to the above method.

### III. EXPERIMENTAL RESULTS AND ANALYSIS FOR VIDEO FOREGROUND-BACKGROUND SEPARATION

#### A. EXPERIMENT DATASETS, ENVIRONMENT, COMPARED ALGORITHMS AND PARAMETERS

In order to evaluate the efficiency of the proposed GNNSSN and RGNSSN methods, a large number of simulation experiments are performed for the video foreground-background separation. The simulation experiments contain two major parts: one is the simulation experiment of GNNSSN method which deals with the noisless video foreground-background separation, and the other is simulation experiment of the RGNSSN method which deals with the noisy video foreground-background separation. The experiment datasets, experiment environment and experiment parameters are shown as follows:

1) EXPERIMENT DATASETS

The experiment datasets include 12 videos of CDnet database [44] and I2R database [45], and every video contains 30 frames. The information of one video frame of each video is given as follows: airport (176 × 144), ShoppingMall (106 × 128), Bootstrap (160 × 120), WaterSurface (160 × 120), Escalator (160 × 130), Fountain (160 × 130), Lobby (160 × 128), Curtain (160 × 128), backdoor (320 × 240), highway (320 × 240), office (320 × 240), and busstation (320 × 240).

2) EXPERIMENT ENVIRONMENT

The experimental environment includes hardware and software environments. The program run on the computer with Intel (R) Core (TM) i5- 5287U CPU E5-2609, 2.90GHz frequency, 8GB memory, and the experimental software environment is MATLAB 2018a.

3) COMPARED ALGORITHMS

In this section, we compared the proposed GNNSSN method with the methods including LSD [41], MAMR [42], NonLRSD [27], Godec [24], TNRR [23], NNWNN [25], PCP [20], NNP [32], GNR [35] and we compared the proposed RGNSSN method with the methods including LSD [41], RMAMR [42], NonLRSD [27], Godec [24], TNRR [23], NNWNN [25], PCP [20], NNP [32], GNR [35]. In addition, we add Gaussian noise with a variance of 25 to the mentioned 12 test videos for the simulation experiment of RGNSSN method.

4) EXPERIMENT PARAMETERS

A large number of experiments have demonstrated that the two methods proposed in this paper gains the best performances under the following parameters: The experiment parameters of the GNNSSN and the RGNSSN methods are given as follows: $\rho = 1.5$, $\mu_0 = \frac{1.25}{\| M \|_2}$, $\mu_{\max} = 10^7$, $\tau = \frac{1}{\| \max(m, n) \|_1}$, $\tau_1 = \frac{1}{\| \max(m, n) \|_1}$, $\tau_2 = 10$.

In model (3) and (15), $g(\cdot) : R^+ \rightarrow R^+$ is a concave, closed, proper and lower semi-continuous function, for example, the Logarithm penalty, the $l_p$-norm($0 < p < 1$), SCAD penalty function, MCP penalty function, Geman penalty function and the Laplace penalty function. In order to choose the best $g(\cdot)$ function, we carry out a large number of experiments in the same condition while the penalty function is different. This paper takes Airport video as an example to show the effect of video background separation based on various non-convex penalty functions. The F-measure value is a quantitative indicator used to measure the effect of video foreground-background separation. Its expression is given as follows:

$$
F = \frac{2rp}{r + p}
$$

where $r = \frac{TP}{TP+FP}$ is the accuracy rate which is used to measure precision, $r = \frac{TP}{TP+FN}$ is the recall rate which is used to measure recall. $TP$ is the number of correct pixels judged as the foreground, $FP$ is the number of wrong pixels judged as the background, and $FN$ is the number of pixels misjudged as the background. The F-measure value is between 0 and 1. The larger the value, the better the effect of the video foreground-background separation.

The F-measure values based on the different Penalties are shown in Table1. We can see from the Table 1 that the Logarithm penalty gain the best effect of the video foreground-background separation when we use the GNNSSN and RGNSSN methods. So in this paper, we choose the Logarithm penalty as the $g(\cdot)$ function. And the Logarithm penalty is defined as follows:

$$
\begin{align*}
g \left( \sigma_i \left( B \right) \right) &= \frac{\eta}{\log(\vartheta + 1)} \log(\vartheta \sigma_i \left( B \right) + 1)
\end{align*}
$$

where $\eta$ and $\vartheta$ are parameters. A large number of experiments show that we can get the best performances when $\eta = 40\| M \|_\infty$ and $\vartheta = 1.1$.

In Eq. (4), the structured sparse norm $\Phi(\cdot)$ is defined as $l_\infty$-norm over groups of variables, and when $D$ is the set of singletons, we get back the $l_1$-norm. However, the $l_2$-norm is also used in [40]. So in this paper, we give some experiments to demonstrate why we choose the $l_\infty$-norm. For example, the F-measure values of the airport video foreground-background separation are 0.61692 by GNNSSN methods with $l_\infty$-norm, 0.59263 by GNNSSN methods with $l_2$-norm, 0.61677 by RGNSSN methods with $l_\infty$-norm, 0.59089 by RGNSSN methods with $l_2$-norm. We can see from these experiments that the $l_\infty$-norm which is piecewise linear get better performance and we choose it.
In this paper, $D$ in Eq. (4) is a collection of group distributions, where each group distribution is $d \in D$. The choice of $d \in D$ will affect the performance of the experiments. In this section, we use many experiments to evaluate the performances of the ADMM with different $d \in D$ (overlapping-patch groups of $1 \times 1$, $3 \times 3$, $6 \times 6$, $9 \times 9$, $12 \times 12$, $15 \times 15$, $18 \times 18$) and choose the best $d \in D$. We choose the airport video as an example to show the performance of different $d \in D$ in the Table 2. We can see from the Table 2, the $d \in D$ which is $3 \times 3$ can get the best effect of video background separation by the methods of GNNSSN and RGNSSN. So we choose $d \in D$ which is $3 \times 3$ in this paper.

### 2. Simulation Experiments and Analysis on GNNSSN Method

In this section, in order to evaluate the efficiency of the proposed GNNSSN method, we apply this method to the experiments of the noiseless videos foreground-background separation. 

This experiments exhibit and analyze the following foreground of videos frames extracted by different methods. They are the 2926-th frame of airport, the 1862-th frame of Bootstrap, the 2514-th frame of Lobby, the 2323-th frame of Curtain, the 1842-th frame of Bootstrap, the 1615-th frame of WaterSurface, the 2978-th frame of Escalator, the 2514-th frame of Lobby, the 2323-th frame of Curtain, the 1658-th frame of back-door, the 689-th frame of highway, the 668-th frame of office, the 305-th frame of busstation. The foregrounds of videos extracted by different methods are shown in Table 1.

For example, the foreground information of the WaterSurface video extracted by the MAMR and Godec methods has three front people. And we can also see that the extracted foreground information of the Curtain and the Office videos of high accuracy. Finally, it can be seen from Figure 1 that the extracted foreground information of the proposed GNNSSN method has less noise compared with the existing compared LRSD based methods. Especially, the foregrounds extracted from WaterSurface video by NNP and GNR methods contain more noise than the GNNSSN method. The foregrounds extracted from Fountain video extracted by LSD, MAMR, NNWNN, NonLRSD, Godec, TNN, PCP, NNP and GNR methods also contain more noise than the GNNSSN method. In summary, we can see from Figure 1 that the GNNSSN method is superior to 9 other LRSD based methods, and the extracted foreground information of this method is richer, more accurate, and less noise than 9 other LRSD based methods.

In order to demonstrate the advantages of the GNNSSN method clearly, this paper uses the F-measure value to quantify the efficiency of video foreground-background separation, and illustrates the effect of the GNNSSN method quantitatively. The F-measure values of different methods for different noiseless videos are shown in Table 3.

It can be seen from Table 3 that the F-measure values of the proposed GNNSSN method are higher than the other methods. For example, the F-measure value of the GNNSSN is 7.9% higher than PCP method for the foreground-background separation of airport, 9.7% higher than LSD method for the WaterSurface, 3.4% higher than NonLRSD method for the Escalator, 25% higher than the Godec method for the Fountain, 8.6% higher than the NNWNN method for the highway, 13% higher than the NNP method and 7% higher than the GNR method for the office. We can also see from the average F-measure values that the GNNSSN method is better than all other methods is always mixed with a lot of background information. However, the proposed GNNSSN method does not have these defects and has the characteristics of high accuracy. Finally, it can be seen from Figure 1 that the extracted foreground information of the proposed GNNSSN method has less noise compared with the existing compared LRSD based methods. Especially, the foregrounds extracted from WaterSurface video by NNP and GNR methods contain more noise than the GNNSSN method. The foregrounds extracted from Fountain video extracted by LSD, MAMR, NNWNN, NonLRSD, Godec, TNN, PCP, NNP and GNR methods also contain more noise than the GNNSSN method. In summary, we can see from Figure 1 that the GNNSSN method is superior to 9 other LRSD based methods, and the extracted foreground information of this method is richer, more accurate, and less noise than 9 other LRSD based methods.

### Table 1. F-measure values of different penalties (BEST: BOLD).

| Methods   | Logarithm | $l_p$-norm | SCAD   | MCP    | Geman  | Laplace |
|-----------|-----------|------------|--------|--------|--------|---------|
| GNNSSN    | 0.6184    | 0.59612    | 0.60905| 0.56066| 0.60498| 0.57836 |
| RGNSSN    | 0.61727   | 0.59684    | 0.60867| 0.56992| 0.60749| 0.57867 |

### Table 2. F-measure values of different $d \in D$ (BEST: BOLD).

| Methods   | $1 \times 1$ | $3 \times 3$ | $6 \times 6$ | $9 \times 9$ | $12 \times 12$ | $15 \times 15$ | $18 \times 18$ |
|-----------|--------------|--------------|--------------|--------------|----------------|----------------|----------------|
| GNNSSN    | 0.61053      | 0.61752      | 0.60615      | 0.60606      | 0.60347        | 0.60079        | 0.59693        |
| RGNSSN    | 0.61020      | 0.6184       | 0.60596      | 0.60575      | 0.60310        | 0.60055        | 0.59669        |
F-measure values, we can see that our proposed GNNSNN method is better than any other compared LRSD based methods for noiseless videos foreground-background separation.

In this experiments, we also compare the running time between the proposed GNNSNN method and other competing methods, we record the average running time on

---

**FIGURE 1.** Foreground extraction results of noiseless video.

**TABLE 3.** F-measure values of noiseless videos foreground-background separation (BEST: BOLD).

| Datasets     | GNNSNN | LSD    | MAMR   | NonLRSD | Godec   | TNN    | NNWNN  | PCP    | NNP    | GNR    |
|--------------|--------|--------|--------|---------|---------|--------|--------|--------|--------|--------|
| airport      | 0.61752| 0.58575| 0.59441| 0.57794 | 0.59462 | 0.54027| 0.56372| 0.53834| 0.57578| 0.59464|
| ShoppingMall | 0.69439| 0.69357| 0.68731| 0.69180 | 0.68743 | 0.69168| 0.69360| 0.69150| 0.68270| 0.67714|
| Bootstraw    | 0.59260| 0.58963| 0.58189| 0.58995 | 0.58234 | 0.57716| 0.58408| 0.58575| 0.58036| 0.57915|
| WaterSurface | 0.77386| 0.67647| 0.58699| 0.66100 | 0.58889 | 0.34220| 0.62461| 0.56518| 0.51214| 0.48000|
| Escalator    | 0.59344| 0.54581| 0.54002| 0.55911 | 0.52689 | 0.53044| 0.56280| 0.54172| 0.55439| 0.55928|
| Fountain     | 0.67980| 0.55261| 0.43026| 0.65331 | 0.42916 | 0.60082| 0.65338| 0.59976| 0.60885| 0.63840|
| Lobby        | 0.58518| 0.52454| 0.47686| 0.54307 | 0.47930 | 0.50408| 0.53022| 0.49644| 0.51805| 0.53750|
| Curtain      | 0.59508| 0.54011| 0.54642| 0.53959 | 0.54937 | 0.40898| 0.53355| 0.48973| 0.49621| 0.51158|
| backdoor     | 0.64499| 0.63915| 0.55925| 0.64080 | 0.55890 | 0.53972| 0.64027| 0.61493| 0.63289| 0.62190|
| highway      | 0.65859| 0.57852| 0.60055| 0.56951 | 0.60003 | 0.51131| 0.57268| 0.50796| 0.55275| 0.54870|
| office       | 0.44955| 0.37100| 0.43431| 0.38346 | 0.43468 | 0.33718| 0.37376| 0.31513| 0.37594| 0.39878|
| bussation    | 0.57099| 0.55579| 0.58879| 0.53322 | 0.08897 | 0.52241| 0.53425| 0.52201| 0.55348| 0.56243|
| average      | 0.53258| 0.48951| 0.47336| 0.49591 | 0.47289 | 0.43628| 0.49049| 0.46203| 0.47454| 0.47925|
each frame of experimental videos. For example, for the WaterSurface video, the GNNSNN method needs 0.18s, the LSD method needs 1.18s, the MAMR method needs 0.17s, the NonLRSD method needs 0.05s, the Godec method needs 0.004s, the TNN method needs 0.07s, the NNWNN method needs 0.05s, the PCP method needs 0.07s, the NNR method needs 0.19s, and the GNR method needs 0.22s. We can see from the running time mentioned above that our proposed GNNSSN method is faster than the methods of LSD, NNR, GNR. And our proposed method is slightly slower than the methods of MAMR, NonLRSD, Godec, TNN, NNWNN and PCP because of the computation of SVD in each iteration.

C. SIMULATION EXPERIMENTS AND ANALYSIS ON RGNNSSN METHOD
In order to verify the effectiveness of the RGNNSSN method, we add a Gaussian noise with a variance of 25 to test videos, and use the RGNNSSN, LSD, RMAMR, NonLRSD, Godec, TNN, NNWNN and PCP methods to deal with the foreground-background separation of all test videos. We randomly choose the following frames for analysis: the 2926-th frame of airport, the 1862-th frame of ShoppingMall, the 1842-th frame of Bootstrap, the 1615-th frame of WaterSurface, the 2978-th frame of Escalator, the 1494-th frame of Fountain, the 2514-th frame of Lobby, the 2323-th frames of Curtain, the 1658-th frame of backdoor, the 689-th frame of highway, the 668-th frame of office, the 305-th frame of busstation. The Foreground extraction results of noisy video are shown in Figure 2.

As can be seen from the Figure 2, the proposed RGNNSSN method has more significant effectiveness on noisy videos foreground-background separation than the 9 other LRSD based methods. For example, compared with the 9 other methods, the proposed RGNNSSN method can completely extract the luggage in the hands of a passenger in the airport, the information of the foreground person of the WaterSurface, the information of the lecturer of the Curtain, and the information of walking people in backdoor. Compared with other methods, the RGNNSSN method extracts foreground with less noise such as WaterSurface, Fountain, Escalator, Curtain, and highway videos. It can be seen from the foreground of
the WaterSurface, Lobby, Curtain, and office, our proposed RGNNSSN method does not regard the background as foreground. In summary, the RGNNSSN method for noisy videos foreground-background separation has obvious advantages than the 9 other LRSD based methods.

In addition, in order to quantify the effectiveness of the RGNNSSN method for noisy video foreground-background separation, we calculate the F-measure values of the extracted foreground of our RGNNSSN method and the 9 other LRSD based methods. The F-measure values of each method are shown in Table 4.

It can be seen from Table 4 that our proposed RGNNSSN method has a bigger F-measure value than the other compared methods for noisy video foreground-background separation. For example, we can also see that the average F-measure value of RGNNSSN method is 15.7% higher than the RMAMR method, and 5% higher than the NNP and GNR methods, and so on. So we can get the conclusion that the RGNNSSN method has more superiority than the 9 other LRSD based methods.

In this experiments, the processing time between the RGNNSSN method and the compared methods for every experimental videos are recorded. For example, the processing time of the RGNNSSN method for each frame of WaterSurface is 0.13s, the LSD method needs 1.18s, the RMAMR method needs 0.36s, the NonLRSD method needs 0.04s, the Godec method needs 0.004s, the TNN method needs 0.05s, the NNWNN method needs 0.04s, the PCP method needs 0.05s, the NNR method needs 0.19s, and the GNR method needs 0.16s. We can see that our proposed RGNNSSN method is slowly faster than the methods of LSD, RMAMR, NNR, GNR.

### REFERENCES

[1] S. Prativadibhayankaram, H. Van Luong, T. H. Le, and A. Kaup, “Compressive online robust principal component analysis with optical flow for video foreground-background separation,” in Proc. 8th Int. Symp. Inf. Commun. Technol. SoICT, 2017, pp. 385–392.

[2] K. Huang, C. Zhu, and G. Li, “Robust salient object detection via fusing foreground and background priors,” in Proc. 25th IEEE Int. Conf. Image Process. (ICIP), Oct. 2018, pp. 2341–2345.

[3] Z. Yang, L. Fan, Y. Yang, Z. Yang, and G. Gui, “Generalized nuclear norm and Laplacian scale mixture based low-rank and sparse decomposition for video foreground-background separation,” Signal Process., vol. 172, Jul. 2020, Art. no. 107527.

[4] C. Stauffer and W. E. L. Grimson, “Adaptive background mixture models for real-time tracking,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., Jun. 1999, pp. 246–252.

[5] Z. Zivkovic, “Improved adaptive Gaussian mixture model for background subtraction,” in Proc. 17th Int. Conf. Pattern Recognit. ICPAR, Aug. 2004, pp. 28–31.

[6] O. Barnich and M. V. Droogenbroeck, “ViBe: A universal background subtraction method for video sequences,” IEEE Trans. Image Process., vol. 20, no. 6, pp. 1709–1724, Dec. 2011.

[7] N. Kwak, “Principal component analysis based on L1-norm maximization,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 9, pp. 1672–1680, Sep. 2008.

[8] S. Nakajima, M. Sugiyama, and S. D. Babacan, “Variational Bayesian sparse additive matrix factorization,” Mach. Learn., vol. 92, nos. 2–3, pp. 319–347, Sep. 2013.

[9] C. Tian, Q. Zhang, J. Zhang, G. Sun, and Y. Sun, “2D-PCA representation and sparse representation for image recognition,” J. Comput. Theor. Nanosci., vol. 14, no. 1, pp. 829–834, Jan. 2017.

[10] M. Sahu, Y. Sharma, and D. Sharma, “Feature compression using PCA on motor imagery classifications,” in Proc. 3rd Int. Conf. Internet Things Connected Technol. (ICIoTCT), 2018, pp. 111–117.
[11] C. Guzel Turhan and H. S. Bilge, “Class-wise two-dimensional PCA method for face recognition,” IET Comput. Vis., vol. 11, no. 4, pp. 286–300, Jun. 2017.

[12] Z. Lin, M. Chen, and Y. Ma, “The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices,” 2010, arXiv:1009.5055. [Online]. Available: http://arxiv.org/abs/1009.5055

[13] J. Qiao, Y. Yang, S. He, S. Gu, L. Zhang, and R. W. H. Lau, “Joint image denoising and disparity estimation via stereo structure PCA and noise-tolerant cost,” Int. J. Comput. Vis., vol. 124, no. 2, pp. 204–222, Sep. 2017.

[14] X. Deng, X. Tian, S. Chen, and C. J. Harris, “Nonlinear process fault diagnosis based on serial principal component analysis,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 3, pp. 560–572, Mar. 2018.

[15] E. Candès, X. Li, Y. Ma, and J. Wright, “Robust principal component analysis?: Recovering low-rank matrices from sparse errors,” in Proc. IEEE Sensor Array Multichannel Signal Process. Workshop, Oct. 2010, pp. 201–204.

[16] E. J. Candès, X. Li, Y. Ma, and J. Wright, “Robust principal component analysis?” J. ACM, vol. 58, no. 1, pp. 1–37, 2009.

[17] J. Zhao and L. Zhao, “Low-rank and sparse matrices fitting algorithm for low-rank representation,” Comput. Math. Appl., vol. 79, no. 2, pp. 407–425, Jan. 2020.

[18] J. Liu and B. D. Rao, “Robust PCA via ℓ0-ℓ1 regularization,” IEEE Trans. Signal Process., vol. 67, no. 2, pp. 535–549, Jan. 2018.

[19] 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.

[20] T. Bouwmans and E. H. Zahzah, “Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance,” Comput. Vis. Image Underst., vol. 122, pp. 22–34, May 2014.

[21] 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.

[22] R. Li, J. Pan, Y. Si, B. Yan, Y. Hu, and H. Qin, “Specular reflections removal for endoscopic image sequences with adaptive-RPCA decomposition,” IEEE Trans. Med. Imag., vol. 39, no. 2, pp. 328–340, Feb. 2020.

[23] 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.

[24] T. Zhou and D. Tao, “GoDec: Randomized low-rank & sparse matrix decomposition in noisy case,” in Proc. Int. Conf. Mach. Learn., Washington, DC, USA, 2011, pp. 33–40.

[25] Z. Yang, Z. Yang, and D. Han, “Alternating direction method of multipliers for sparse and low-rank decomposition based on nonconvex non-smooth weighted nuclear norm,” IEEE Access, vol. 6, pp. 56945–56953, 2018.

[26] 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. 182–208, Jan. 2017.

[27] K. Hosono, S. Ono, and T. Miyata, “Weighted tensor nuclear norm minimization for color image denoising,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2016, pp. 2862–2869.

[28] 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.

[29] X. Jia, X. Feng, and W. Wang, “Adaptive regularizer learning for low rank approximation with application to image denoising,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2016, pp. 3096–3100.

[30] X. Jia, X. Feng, W. Wang, C. Xu, and L. Zhang, “Bayesian inference for adaptive low rank and sparse matrix estimation,” Neurocomputing, vol. 291, pp. 71–83, May 2018.

[31] F. Xu, J. Han, Y. Wang, M. Chen, Y. Chen, G. He, and Y. Hu, “Dynamic magnetic resonance imaging via nonconvex low-rank matrix approximation,” IEEE Access, vol. 5, pp. 1958–1966, 2017.

[32] L. Yang, T. K. Pong, and X. Chen, “Alternating direction method of multipliers for a class of nonconvex and nonsmooth problems with applications to Background/Foreground extraction,” SIAM J. Imag. Sci., vol. 10, no. 1, pp. 74–110, Jan. 2017.

[33] Q. Li and G. Tang, “The nonconvex geometry of low-rank matrix optimization with general objective functions,” in Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP), Nov. 2017, pp. 1235–1239.

[34] Y. Li, Y. Lin, X. Cheng, Z. Xiao, F. Shu, and G. Gui, “Nonconvex penalized regularization for robust sparse recovery in the presence of SxS noise,” IEEE Access, vol. 6, pp. 25474–25485, 2018.

[35] F. Wen, R. Ying, P. Liu, and R. C. Qiu, “Robust PCA using generalized nonconvex regularization,” IEEE Trans. Circuits Syst. Video Technol., early access, Apr. 2, 2019, doi: 10.1109/TCSVT.2019.2908833.

[36] C. Lu, C. Zhu, and X. Xu, “Generalized singular value thresholding,” in Proc. 29th AAAI Conf. Artif. Intell., Feb. 2015, pp. 1805–1811.

[37] C. Lu, J. Tang, S. Yan, and Z. Lin, “Generalized nonconvex nonsmooth low-rank minimization,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2014, pp. 4130–4137.

[38] 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.

[39] L. Ma, L. Xu, T. Zeng, “Low rank prior and total variation regularization for image Deblurring,” J. Sci. Comput., vol. 70, no. 3, pp. 1336–1357, 2017.

[40] J. Mairal, R. Jenatton, G. Obozinski, and F. Bach, “Convex and network flow optimization for structured sparsity,” J. Mach. Learn. Res., vol. 12, pp. 2681–2720, Sep. 2011.

[41] 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.

[42] X. Ye, J. Yang, X. Sun, K. Li, C. Hou, and Y. Wang, “Foreground–Background separation from video clips via motion-assisted matrix restoration,” IEEE Trans. Circuits Syst. Video Technol., vol. 25, no. 11, pp. 1721–1734, Nov. 2015.

[43] P. Zhao, G. Rocha, and B. Yu, “The composite absolute penalties family for grouped and hierarchical variable selection,” Ann. Statist., vol. 37, no. 6A, pp. 3468–3497, Dec. 2009.

[44] N. Goyette, P.-M. Jodoin, P. Porikli, J. Konrad, and P. Ishwar, “ChangepointNet: A new change detection benchmark dataset,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, Jan. 2012, pp. 1–8.

[45] L. Li, W. Huang, I. Y.-H. Gu, and Q. Tian, “Statistical modeling of complex backgrounds for foreground object detection,” IEEE Trans. Image Process., vol. 13, no. 11, pp. 1459–1472, Nov. 2004.

**Yongpeng Yang** was born in Shandong, China, in 1984. He received the M.S. degree from the Nanjing University of Posts and Telecommunications, in 2011. He is currently a Lecturer with the Nanjing Vocational College of Information Technology. His main research interests include low-rank and sparse decomposition and multimedia information processing.

**Zhenzhen Yang** was born in Shandong, China, in 1984. She received the M.S. and Ph.D. degrees from the Nanjing University of Posts and Telecommunications, in 2011 and 2014, respectively. She was a Lecturer with the Nanjing University of Posts and Telecommunications, in 2014, where she was an Associate Professor, in 2018. Since 2015, she has been a project manager of several national projects, such as the National Natural Science Foundation of China and the China Postdoctoral Science Foundation. She is the author of more than 20 journal articles and conference papers. Her research interests include sparse and low-rank decomposition and multimedia information processing.
JIANLIN LI was born in Nanjing, China, in 1974. He is currently a Professor with the Nanjing Vocational College of Information Technology. He is also a Young and Middle-Aged Academic Leader of Qing Lan Project and the 333 Project in Jiangsu Province. His main research interests include low-rank and sparse decomposition and multimedia information processing.

LU FAN was born in Taizhou, China, in 1994. She received the B.S. degree from the North China College, Nanjing Normal University, in 2017. She is currently pursuing the M.S. degree with the Nanjing University of Posts and Telecommunications. Her research interests include sparse and low-rank decomposition and multimedia information processing.

* * *