Generalized Rank Minimization based Group Sparse Coding for Low-level Image Restoration via Dictionary Learning

Yunyi Li¹, Guan Gui²*, and Xiefeng Cheng¹
¹ College of Electronic and Optical Engineering & College of Microelectronics, Nanjing University of Posts and Telecommunications, Nanjing 210023, China;
² College of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210023, China;
*Corresponding Author: guiguan@njupt.edu.cn

Abstract: Recently, low-rank matrix recovery theory has been emerging as a significant progress for various image processing problems. Meanwhile, the group sparse coding (GSC) theory has led to great successes in image restoration with group contains low-rank property. In this paper, we introduce a novel GSC framework using generalized rank minimization for image restoration tasks via an effective adaptive dictionary learning scheme. For a more accurate approximation of the rank of group matrix, we proposed a generalized rank minimization model with a generalized and flexible weighted scheme and the generalized nonconvex nonsmooth relaxation function. Then an efficient generalized iteratively reweighted singular-value function thresholding (GIR-SFT) algorithm is proposed to handle the resulting minimization problem of GSC. Our proposed model is connected to image restoration (IR) problems via an alternating direction method of multipliers (ADMM) strategy. Extensive experiments on typical IR problems of image compressive sensing (CS) reconstruction, inpainting, deblurring and impulsive noise removal demonstrate that our proposed GSC framework can enhance the image restoration quality compared with many state-of-the-art methods.

Keywords: low-rank; group sparse coding; generalized rank minimization; nonconvex nonsmooth; image restoration; reweighted; singular-value function thresholding; dictionary learning.

I. Introduction

Image restoration (IR) problem [1][2][3] is a classic yet active topic in low-level vision. The goal of IR is to restore a high-quality image or sequence $\mathbf{X}$ from its various degraded observations $\mathbf{Y}$. Typical low level IR research includes image compressive sensing (CS) reconstruction [4][5], denoising [6], inpainting [7] and deblurring [8] and impulsive noise removal [9], etc. To tackle this typical ill-posed problem, the most popular method is the regularization technique

$$\min_{\mathbf{X}} \{ E(\mathbf{X}) = f(\mathbf{X}, \mathbf{Y}) + \lambda R(\mathbf{X}) \}$$

(1)

where $f(\mathbf{X}, \mathbf{Y})$ denotes the fidelity term which can penalize our desired restoration image sequence $\mathbf{X}$ far from the original degraded $\mathbf{Y}$, the second term $R(\mathbf{X})$ is the regularization term which can provide the necessary prior knowledge of image, e.g., the sparsity, smoothness or continuity, and the regularization parameter can make the tradeoff between the first fidelity term and the second
regularization term.

It is well documented that how to exploit more prior knowledge for the minimization of problem (1) is at the core, in the past several years, how to design the regularization term to exploit the prior of an image in a predefined domain has been widely studied, such as total-variation (TV) regularization [10][11], which employs the local structural patterns as the prior knowledge. The L1-norm regularization, the nonconvex penalized regularization of \( L_p \)-norm [12], Smoothly Clipped Absolute Deviation (SCAD) [13], Logarithm [14], and Minimax Concave Penalty (MCP) [15], utilize the sparsity as prior for minimization. These regularization methods have shown their great success in various sparse signal recovery [16][17][18] and image processing applications. However, these traditional regularization term for IR approaches can only exploit a few structural features of image and some important artifacts will not be preserved. On the other hand, the promising performance of the nonlocal self-similarity of patches has been well documented, which can provide more prior knowledge to improve the restoration quality by exploiting the nonlocal self-similarity features of image [19]. Lately, the sparsity and the nonlocal self-similarity features are often exploit simultaneously to product better approximation [20], one of the state-of-the-art model is the nonlocally centralized sparse representation (NCSR) model proposed in [21], which can exploit the nonlocal self-similarity of image to achieve a more accurate sparse representation coefficient, and then centralize the sparse coefficients of the degraded image to the restored image, and has shown its promising performance.

The last decades has seen the low-rank minimization-based methods for image and video restoration [22] with the development of the theory of compressive sensing and sparse representation. For any given image, the adjacent patches have similar structures, and similar patches are grouped into a matrix, such that the matrix shows the low-rank property, then matrix competition for each group can be conducted to recover the desired image. In [23], a novel and efficient framework is proposed for image restoration using group sparse coding (GSC) model, where the image can be sparsely represented as a linear combination in the domain of group. Although the proposed GSC based IR model has shown the significant improvements for image restoration, the important low-rank property of each group is not being used, and minimizing the optimization problem from the perspective of sparse vector recovery, and hence, the gap between the sparsity-based approaches and the low-rank based approaches for image restoration is still exist.

Like the \( L_0 \)-norm regularized minimization problem, it is not tractable to minimize the rank regularized problem because of the property of nonconvex and discontinuous of \( \text{rank}(\mathbf{X}) \), and is usually relaxed to the convex nuclear norm minimization (NNM) model. However, the most popular convex NNM often leads to a biased solution since NNM is usually over-shrink the singular values and treat each of them equally. It is well documented that the nuclear norm minimization is the most popular approach for rank minimization, such as the weighted nuclear norm (WNN), the truncated nuclear norm (TNN) [24], the Schatten-\( p \) nuclear norm (Sp -NN) [25], and the generalized nonconvex nonsmooth functions on singular values (e.g., SCAD, MCP and Logarithm, e.t.,) [26][27]. In practice, it is the fact that the larger singular values often quantify the main preserved information, to tackle this problem, various low-rank relaxations have been proposed, an alternative and popular relaxations is the nonconvex and nonsmooth relaxations.

In this paper, motivated by the promising performance of nonconvex low-rank minimization approaches for matrix recovery, and to bridge the gap between GSC modal and existing rank minimization model, we develop a novel GSC framework for IR problem using a generalized rank
minimization scheme. Our proposed framework will not only unify the local sparsity and nonlocal similarity of image simultaneously, but also can convert the traditional sparse coding problem into the generalized low-rank minimization model, and thus can restore image using various matrix competition approaches. Our contributions in this paper can be summarized as:

1) We develop a connection from GSC to low-rank matrix minimization via an effective dictionary learning approach, where the sparse representation coefficient vector is also the singular value vector of matrix in our proposed model, then we can convert the sparsity-inducing optimization problem into the low-rank matrix minimization problem.

2) For a more accurate approximation of the rank of group matrix, we develop a generalized low-rank minimization model. Our proposed model employs a generalized and flexible weighted scheme and a generalized nonconvex nonsmooth surrogate function on singular values of the group matrix.

3) To solve the generalized rank minimization problem for GSC, we will first convert the optimization model into a derivative double nonconvex nonsmooth rank (DNNR) minimization problem, and then develop an iterative reweighted singular-value function thresholding (IRSFT) algorithm for GSC.

4) We develop an alternative direction method of multipliers (ADMM) framework to integrate the proposed GSC model into IR problem, where the sparse vector for GSC and the group matrix of desired image will be achieved simultaneously. We evaluate the proposed framework for four classic IR problems, including image CS reconstruction, inpainting, deblurring, and impulsive noise removal.

The flowchart of our proposed methods for IR problem by rank minimization based GSC is presented in the Fig. 1.

![Fig. 1. Flowchart of our proposed ADMM-GIR-SFT algorithm for image restoration problem.](image)

The rest of our paper is as follows. In the section II, we will first review the group sparse coding theory, and then build a connection from GSC to low-rank matrix recovery via an effective self-adaptive dictionary learning scheme. In the section III, we propose a weighted nonconvex nonsmooth rank relaxation function for rank minimization, and then a reweighted singular-value thresholding (IDNN) algorithm is developed for GSC problem. In the section IV, we will evaluate the effectiveness of our proposed method for various IR tasks and compare the performance with current state-of-the-art IR algorithms. Finally, a brief summary will be made in section VI.
II. From Group Sparse Coding to Low-Rank Minimization

This section will introduce the basic theory of group sparse coding, and then a self-adaptive dictionary learning strategy is introduced for each group. It should be noted that the self-adaptive dictionary scheme is crucial, we can convert the GSC problem into the low-rank matrix recovery problem via our proposed adaptive dictionary learning scheme.

2.1. Group sparse coding

Concretely, we first divide the original image \( X \in \mathbb{R}^{N \times N} \) into \( n \) overlapping patches \( X_k, k = 1, 2, \ldots, n \), with the size is \( \sqrt{B_c} \times \sqrt{B_c}, B_c < N \). For each patch \( X_k \), there exist \( c \) best matched patches, we denote the set of these patches as \( S_{x_k} \), then we search them in a given searching window with the size of \( L \times L \) using the well-known Euclidean distance as the similarity criterion. Next, these similar patches will be stacked into a group matrix with the size of \( B_c \times c \), denoted by \( X_{G_k} = [X_{G_k,1}, X_{G_k,2}, \ldots, X_{G_k,c}] \in \mathbb{R}^{B_c \times c} \), where each patch will be vectorized as \( X_{G_k,i} \in \mathbb{R}^{B_c \times 1}, i = 1, 2, \ldots, c \) as the columns. Then the constructed group matrix \( X_{G_k} \) consisting of \( c \) patches containing similar structures, the construction process of group can be expressed as

\[
X_{G_k} = c_k(X), k = 1, 2, \ldots, n
\]

(2)

where the operator \( c_k( \cdot ) \) denotes the group construction from \( X \). Fig. 2 illustrates the construction process of group.

Conversely, if we average all the group \( X_{G_k} \), we can achieved the original image \( X \) by

\[
X = \frac{1}{n} \sum_{k=1}^{n} R_k^T(X_{G_k}) / \sum_{k=1}^{n} R_k^T(1_{B_c})
\]

(3)

where \( R_k^T( \cdot ) \) denotes the transpose grouping operator, \( 1_{B_c} \in \mathbb{R}^{B_c \times 1} \) stands for a matrix with all the elements is 1, and \( / \) denotes an element-wise division of two matrix.

Fig. 2. Illustrations of generating every group.

According to sparse coding theory, it is expected that the coefficient vector is as sparse as possible with the group \( X_{G_k} \) can be faithfully represented by the dictionary \( D_{G_k} \). Assume a given dictionary \( D_{G_k} = [d_{G_k,1}, d_{G_k,2}, \ldots, d_{G_k,m}] \in \mathbb{R}^{(B_c \times c) \times m} \), the representation coefficient vector can be achieved by

\[
\hat{\alpha}_{G_k} = \arg \min_{\alpha_{G_k}} \frac{1}{2} \|X_{G_k} - D_{G_k} \alpha_{G_k}\|_2^2 + \lambda \|\alpha_{G_k}\|_0 \quad k = 1, 2, \ldots, n
\]

(4)

in which \( \alpha_{G_k} \in \mathbb{R}^{m \times 1} \) is the achieved sparse coefficient vector. It should be noted that every atom \( d_{G_k,i} \in \mathbb{R}^{B_c \times c}, i = 1, 2, \ldots, m \) in the dictionary \( D_{G_k} \) is a matrix with the same size of group \( X_{G_k} \). Then the image group will be represented sparsely by

\[
X_{G_k} = D_{G_k} \hat{\alpha}_{G_k} = \sum_{i=1}^{m} \alpha_{G_k,i} d_{G_k,i}
\]

(5)

where \( D_{G_k} = [d_{G_k,1}, \ldots, d_{G_k,m}] \in \mathbb{R}^{(B_c \times c) \times m} \) denotes the dictionary, and \( \hat{\alpha}_{G_k} = [\hat{\alpha}_{G_k,1}, \hat{\alpha}_{G_k,2}, \ldots, \hat{\alpha}_{G_k,m}] \in \mathbb{R}^{m \times 1} \) is the desired vector.
2.2. Self-adaptive dictionary learning strategy

To obtain an adaptive dictionary $D_{G_k}$ for each group $X_{G_k}$, in this paper, we adopt a self-adaptive dictionary learning scheme [28] for each group, we can learn the adaptive dictionary $D_{G_k}$ from $X_{G_k}$ directly. We first conduct the singular value decomposition (SVD) of $X_{G_k}$ by

$$X_{G_k} = U_{G_k} \Sigma_{G_k} V_{G_k}^T = \sum_{i=1}^{m} \sigma_{X_{G_k}} u_{x_{G_k}} v_{x_{G_k}}^T$$

where $m = \min(B_s, c) \quad U_{G_k} = [u_{x_{G_k}^{1}}, u_{x_{G_k}^{2}}, \ldots, u_{x_{G_k}^{m}}] \quad V_{G_k} = [v_{x_{G_k}^{1}}, v_{x_{G_k}^{2}}, \ldots, v_{x_{G_k}^{m}}]$ and $\Sigma_{G_k} = \text{diag}(\sigma_{x_{G_k}^{1}}; \sigma_{x_{G_k}^{2}}; \ldots; \sigma_{x_{G_k}^{m}})$. Then, the atom $d_{G_k}$ of the dictionary $D_{G_k}$ can be obtained by

$$d_{G_k} = u_{x_{G_k}} v_{x_{G_k}}^T, \quad i = 1, 2, \ldots, m$$

Finally, we can achieve the self-adaptive dictionary for each group by

$$D_{G_k} = [d_{G_k}^{1}, d_{G_k}^{2}, \ldots, d_{G_k}^{m}]$$

For simplify, if we concatenate all the corresponding dictionary $D_{G_k}$ as $D_G$, and the original entire image $X$ can be represented by

$$X = D_G \circ \alpha_G$$

where $\alpha_G$ denotes the concatenation of all $D_{G_k}$. As a result, the learned self-adaptive dictionary will build the link between the GSC problem and rank minimization problem.

2.3. From GSC to low-rank minimization

For practical IR problems, the original image $X$ and the corresponding groups $X_{G_k}$ are unknown, we usually need to estimate the latent image $X_{G_k}$ over the adaptive dictionary $D_{G_k}$ from its degraded observation $Y_{G_k} \in \mathbb{R}^{B_s \times c}$ by the following group sparse coding model

$$\tilde{\alpha}_{G_k} = \arg \min_{\alpha_{G_k}} \frac{1}{2} \|Y_{G_k} - D_{G_k} \alpha_{G_k}\|_2^2 + \lambda \|\alpha_{G_k}\|_0, \quad k = 1, 2, \ldots, n.$$  \hspace{1cm} (10)

where $\alpha_{G_k}$ denotes the sparse coefficient vector over the dictionary $D_{G_k}$, and $\lambda$ denotes the regularization parameter, then our desired matrix (group) can be reconstructed by $\tilde{X}_{G_k} = D_{G_k} \tilde{\alpha}_{G_k}$.

Benefiting from the fact that the degraded group $Y_{G_k}$ and original group $X_{G_k}$ share the same coding dictionary $D_{G_k}$, and due to $X_{G_k} = D_{G_k} \alpha_{G_k}$, then we have the following equivalent relationship

$$\|\alpha_{G_k}\|_0 = \text{rank}(X_{G_k})$$

where $\text{rank}(X_{G_k})$ denotes the singular value number of matrix $X_{G_k}$. By substituting $X_{G_k} = D_{G_k} \alpha_{G_k}$ in (10), the problem (10) have the following equivalent low-rank minimization problem

$$\hat{X}_{G_k} = \arg \min_{X_{G_k}} \frac{1}{2} \|Y_{G_k} - X_{G_k}\|_F^2 + \lambda \text{Rank}(X_{G_k}), \quad k = 1, 2, \ldots, n.$$  \hspace{1cm} (12)

where $X_{G_k}$ denotes the constructed image group with low-rank property, and $Y_{G_k}$ denotes the corresponding observation group. It is worth noting that the sparse coefficient vector $\alpha_{G_k}$ is also the singular value vector of matrix $X_{G_k}$ in our proposed model. Then we can convert the sparsity-inducing optimization problem (10) into the low-rank minimization problem (12).

III. Group Sparse Coding via Generalized Rank Minimization
To approximate the rank of group matrix more accurately, this section will propose a generalized rank minimization model using a generalized and flexible relaxation function with proper and lower semi-continuity in \([0, +\infty)\). Then a derivative DNNR model is developed to minimize the resulting generalized rank minimization problem. At a result, a generalized iterative reweighted singular value function thresholding (GIR-SFT) algorithm is developed to solve the rank minimization problem for GSC.

### 3.1. Generalized rank relaxation function for GSC

The low-rank minimization problem (12) is a NP-hard problem, which is usually relaxed as a convex minimization problem, one popular convex relaxation model for GSC can be expressed as

\[
\hat{X}_{G_k} = \arg \min_{X_{G_k}} \frac{1}{2} \|Y_{G_k} - X_{G_k}\|_F^2 + \lambda \|X_{G_k}\|_*
\]

where, \(\|X_{G_k}\|_* = \sum_{i=1}^{m} |\sigma_i(X_{G_k})|\) denotes the nuclear norm, and \(\sigma_i(X_{G_k}), i = 1, 2, \ldots\), are the singular values of \(X_{G_k}\), and \(\|\cdot\|_F\) presents the Frobenius norm.

However, the NNM model of (13) usually cannot approximate the rank accurately. Recently, various rank relaxation functions have been proposed to deal with the limitation, such as the truncated nuclear norm (TNN) [24], Schatten \(p\)-norm (Sp) [25], weighted nuclear norm (WNN) [29], and so on. It is well documented that the WNN based method can often achieve the best performance [29]. What’s more, some popular nonconvex counterparts of \(L_q\)-norm on singular value have shown great potentials to improve the rank minimization performance [26][27], typical nonconvex surrogate functions including \(L_p\)-norm [12], SCAD [13], Logarithm [14], MCP [15]. By combining the weighted strategy and the nonconvex counterparts of \(L_q\)-norm, this paper proposed a more generalized and flexible rank relaxation function for rank approximation of the group matrix denoted by

\[
\rho_{w_k}(\sigma(X_{G_k})) = \sum_{i=1}^{r} w_{k,i} \rho(\sigma_{k,i}), \quad r = \min(B_s, c)
\]

where the weighting vector \(w_k = (w_{k,1}, w_{k,2}, \ldots, w_{k,r})\) is the weighting vector with \(w_{k,1} \leq w_{k,2} \leq \cdots \leq w_{k,r}\), \(\sigma(X_{G_k}) = (\sigma_{k,1}, \sigma_{k,2}, \ldots, \sigma_{k,r})\) denotes the singular value vector with \(\sigma_{k,1} \leq \sigma_{k,2} \leq \cdots \leq \sigma_{k,r}\), the function \(\rho(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+\) is the proper and lower semi-continuous function, and is nondecreasing on \([0, +\infty)\). It should be note that the nonconvexity of the function \(\rho(\cdot)\) is often weaker than traditional nonconvex functions, e.g., \(L_p\)-norm, MCP and SCAD. According to the definition of the rank relaxation function, \(\rho_{w} (\cdot)\) will be more flexible with different \(w_l\) and \(\rho(\cdot)\). When \(\rho(\cdot)\) is the absolute function, \(\rho_{w} (\cdot)\) becomes the traditional nuclear norm and the weighted nuclear norm with all \(w_l = 1\) and not all \(w_{k,i} = 1\), respectively. When \(\rho(\cdot)\) is the \(L_p\)-norm with \(0 < p < 1\), \(\rho_{w} (\cdot)\) will become the Schatten \(p\)-nuclear norm and the weighted Schatten \(p\)-nuclear norm with all \(w_{k,i} = 1\) and not all \(w_{k,i} = 1\), respectively. Moreover, when the weighting vector \(w\) with partial \(w_{k,i} = 0\), \(\rho_{w} (\cdot)\) will become the truncated nuclear norm and the truncated Schatten \(p\)-nuclear norm with \(p = 1\) and \(0 < p < 1\), respectively. A summarization for these special cases can be found in the Table 1.

**Table 1.** Different relaxation functions with different weight \(w\) and relaxation function \(\rho\).
After substituting the nuclear norm using the proposed rank relaxation function (14), the GSC problem can be converted into the following generalized rank minimization problem

$$\tilde{X}_{G_k} = \arg \min_{X_{G_k}} \frac{1}{2} \| Y_{G_k} - X_{G_k} \|_F^2 + \lambda \rho \left( \sigma (X_{G_k}) \right)$$

(15)

Our proposed model can estimate the rank of each group with a high accurate and nearly unbiased solution, one popular example is the weighting vector $w_k$ is inversely proportional to the corresponding singular values, e.g., $\omega_{k,i} = 1/|\sigma_i(X_{G_k})| + \varepsilon$ [29]. Then the problem (15) can be described as the following more generalized optimization problem,

$$\tilde{X}_{G_k} = \arg \min_{X_{G_k}} f(X_{G_k}) + \lambda \rho \left( \sigma (X_{G_k}) \right)$$

(16)

where the fidelity function $f(\cdot)$ is continuously differentiable for any $G_1, G_2 \in \mathbb{R}^{p \times q}$ with

$$\| V \|_F (G_1) - f(G_2) \|_F \leq L \| G_1 - G_2 \|_F$$

(17)

where $L > 0$ denotes the Lipschitz constant, moreover, $f(\cdot)$ is possibly nonconvex.

Actually, according to the reweighted strategy [26] and the super-gradient properties [26], the problem (16) can be intrinsically derived from the following nonconvex nonsmooth rank minimization problem [30][31], e.g.,

$$\tilde{X}_{G_k} = \arg \min_{X_{G_k}} f(X_{G_k}) + \lambda \sum_{i=1}^r h \left( \rho \left( \sigma_i(X_{G_k}) \right) \right)$$

(18)

where $h(\rho(\cdot))$ denotes the relaxation function, and in this paper, we choose $h(\cdot) = \rho(\cdot)$ without loss of the generality, in which, $\rho(\cdot)$ denotes a function with lower semi-continuous property in $[0, +\infty)$. As a result, the problem (18) can be substituted by

$$\tilde{X}_{G_k} = \arg \min_{X_{G_k}} f(X_{G_k}) + \lambda \sum_{i=1}^r \rho \left( \sigma_i(X_{G_k}) \right)$$

(19)

in which, there exist two relaxation function $\rho(\cdot)$, thus (19) can be known as double nonconvex nonsmooth rank (DNNR) minimization problem.

### 3.2 Iterative reweighted strategy for DNNR minimization based GSC

The DNNR minimization problem (19) is more difficult to resolve than traditional convex NNM problem. To solve this problem, we need to convert the DNNR minimization problem into the iteratively reweighted minimization problem benefit from the antimonotone property of relaxation function $\rho(\cdot)$ and the supergradient function $\partial \rho(\cdot)$. We first give the following theorem.

**Theorem 3.1.** [26] Let $\rho(\cdot): \mathbb{R}^n \rightarrow \mathbb{R}$ be concave with the monotonically nondecreasing property, and its supergradient $\partial \rho(\cdot)$ is monotonically nonincreasing on $[0, +\infty)$, then we have

$$\rho(s_1) - \rho(s_2) \leq w_{s_2} (s_1 - s_2)$$

(20)

where $w_{s_2} \in \partial \rho(s_2)$, $s_1$ and $s_2$ denote two any given values.

Since the monotonically nondecreasing property of $\rho(\cdot)$ and the monotonically nonincreasing property of the supergradient $\partial \rho(\cdot)$.
of \( \partial \rho(\cdot) \), for any given \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \), we have \( \rho(\sigma_1) \geq \rho(\sigma_2) \geq \cdots \geq \rho(\sigma_r) \) and \( \partial \rho(\rho(\sigma_1)) \leq \partial \rho(\rho(\sigma_2)) \leq \cdots \leq \partial \rho(\rho(\sigma_r)) \), that is to say, the function of \( \partial \rho(\rho(\cdot)) \) is monotonical nonincreasing on \([0, +\infty)\). Then according to the Theorem 3.1, if we substitute the value \( s_1 \) and \( s_2 \) by \( \rho(s_1) \) and \( \rho(s_2) \), and substrate the weighting \( w_{z_2} \in \partial \rho(s_2) \) by \( w_{\rho(s_2)} \in \partial \rho(\rho(s_2)) \). Then we have the following theorem.

**Theorem 3.2.** [31] Let \( \rho(\cdot): \mathbb{R}^n \rightarrow \mathbb{R} \) be concave with the monotonical nondecreasing property, its supergradient \( \partial \rho(\cdot) \) is monotonically nonincreasing on \([0, +\infty)\), and the double supergradient \( \partial \rho(\rho(\cdot)) \) is monotonically nonincreasing on \([0, +\infty)\), then we have

\[
\rho(\rho(s_1)) - \rho(\rho(s_2)) \leq w_{\rho(s_2)}(\rho(s_1) - \rho(s_2))
\]

(21)

where \( w_{\rho(s_2)} \in \partial \rho(\rho(s_2)) \), \( s_1 \) and \( s_2 \) denote two any given values.

Then according to the Theorem 3.2, in this paper, since the relaxation function of \( \rho(\cdot) \) is concave on \([0, +\infty)\), then according to the monotonical nonincreasing property of double supergradient \( \partial \rho(\rho(\cdot)) \). As a result, the problem (19) can be optimized by the following updating scheme, e.g.,

\[
X_{t+1}^{i} = \arg \min_{X_{Gk}^{i}} f(X_{Gk}^{i}) + \lambda \sum_{t=1}^{r} \rho \left( \rho \left( \sigma_i(X_{Gk}^{i}) \right) \right)
\]

\[
= \arg \min_{X_{Gk}^{i}} f(X_{Gk}^{i}) + \lambda \sum_{t=1}^{r} \{ \rho \left( \rho \left( \sigma_i(X_{Gk}^{i}) \right) \right) + w_{t,i} \rho \left( \sigma_i(X_{Gk}^{i}) \right) \}
\]

\[
= \arg \min_{X_{Gk}^{i}} f(X_{Gk}^{i}) + \lambda \sum_{t=1}^{r} w_{t,i} \rho \left( \sigma_i(X_{Gk}^{i}) \right)
\]

(22)

where

\[
\rho \left( \rho \left( \sigma_i(X_{Gk}^{i}) \right) \right) \leq \rho \left( \rho \left( \sigma_i(X_{Gk}^{i}) \right) \right) + w_{t,i} \rho \left( \sigma_i(X_{Gk}^{i}) \right) - \rho \left( \sigma_i(X_{Gk}^{i}) \right)
\]

(23)

with \( w_{t,i} \in \partial \rho \left( \rho \left( \sigma_i(X_{Gk}^{i}) \right) \right) \) and \( w_{t,i} \leq w_{t,i} \leq \cdots \leq w_{k,r} \), in which, the iteratively weighting vector can be updating by \( w_{t,i} \in \partial \rho \left( \rho \left( \sigma_i(X_{Gk}^{i}) \right) \right) \), and \( X_{Gk}^{i} \) denotes the \( k \)-th iteration of variable \( X_{Gk}^{i} \).

Secondly, we need to linearize the fidelity function \( f(X_{Gk}^{i}) \) at the \( k \)-th iteration point \( X_{Gk}^{i} \) by adding the proximal term by

\[
f(X_{Gk}^{i}) = f(X_{Gk}^{i}) + \langle \nabla f(X_{Gk}^{i}), X_{Gk}^{i} - X_{Gk}^{i} \rangle + \frac{\gamma}{2} \left\| X_{Gk}^{i} - X_{Gk}^{i} \right\|_F^2
\]

(24)

where \( \gamma \) is a constant with \( \gamma > L_g \). It should be noted that a proper choice of \( \gamma \) will guarantee the algorithmic convergence. Similar to IRNN [26] and GPG [27], we can update the \( X_{Gk}^{i+1} \) by

\[
X_{Gk}^{i+1} = \arg \min_{X_{Gk}^{i}} \left\{ \frac{\gamma}{2} \left\| X_{Gk}^{i} - X_{Gk}^{i} \right\|_F^2 + \left\langle \nabla f(X_{Gk}^{i}), X_{Gk}^{i} - X_{Gk}^{i} \right\rangle + f(X_{Gk}^{i}) + \lambda \sum_{t=1}^{r} w_{t,i} \rho \left( \sigma_i(X_{Gk}^{i}) \right) \right\}
\]

(25)

As a result, the problem (19) can be expressed as the following iterative reweighted formulation, e.g.,

\[
X_{Gk}^{i+1} = \arg \min_{X_{Gk}^{i}} \left\{ \frac{\gamma}{2} \left\| X_{Gk}^{i} - B_{Gk}^{i} \right\|_F^2 + \lambda \sum_{t=1}^{r} w_{t,i} \rho \left( \sigma_i(X_{Gk}^{i}) \right) \right\}
\]

(26)

where \( B_{Gk}^{i} = X_{Gk}^{i} - \frac{1}{\gamma} \nabla f(X_{Gk}^{i}) \), and \( w_{t,i} \in \partial \rho \left( \rho \left( \sigma_i(X_{Gk}^{i}) \right) \right) \).
3.3. GIR-SFT algorithm

We next give GIR-SFT algorithm to optimize the problem (26). Essentially, to solve (26) is equivalent to solving a proximal operator. Some previous works have been studied to solve the (26) for some special cases of the relaxation functions, such as the iteratively reweighted nuclear norm (IRNN) algorithm (Lu et al. [26]). Here, we will develop our algorithm for GSC proposed in Zhang et al [31], which can be a extend version of Lu’s IRNN method (Lu et al. [26]). We will present that, if the singular-value function thresholding operator is monotone, the problem (26) can be resolved by the weighted operator.

**Theorem 3.3.** [27] Let a function \( \rho(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R} \) such that a proximal operator \( \text{Prox}_{w_i, \rho}^{\delta_i}(\cdot) \) is monotone, i.e., \( \text{Prox}_{w_i, \rho}^{\delta_i}(\delta_i) > \text{Prox}_{w_i, \rho}^{\delta_2}(\delta_2) \) for any \( \delta_1 > \delta_2 \). For any given \( \lambda > 0 \), \( B \in \mathbb{R}^{m \times n}, r = \min(m, n) \), the weighting values with \( w_1 \leq w_2 \leq \cdots \leq w_r \). Let \( B = U \text{Diag}(\delta(B))V^T \) is the SVD of \( B \), and the singular values satisfy \( \delta_1(B) \geq \delta_2(B) \geq \cdots \geq \delta_r(B) \), then the solution to

\[
X = \arg\min_X \frac{1}{2} \|X - B\|_F^2 + \lambda \sum_{i=1}^r w_i \rho(\sigma_i(X))
\]  

(27)

is

\[
X^* = U \text{Diag}(\delta^*_i)V^T
\]

(28)

where \( \delta^*_i \) can be obtained by

\[
\delta^*_i \in \text{Prox}_{w_i, \rho}^{\delta_i}(\delta_i(B)) = \arg\min_{\sigma_i(X) - \delta_i(B)} \frac{1}{2} (\sigma_i(X) - \delta_i(B))^2 + \lambda w_i \rho(\sigma_i(X))
\]

(29)

It follows from the *Theorem 3.3.*, considering the problem (26), since \( w_{k,1} \leq w_{k,2} \leq \cdots \leq w_{k,r} \), and \( \rho(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R} \) and the proximal operator \( \text{Prox}_{w_i, \rho}^{\delta_i}(\cdot) \) is monotone. Then the optimal solution of our optimization problem (26) can be eventually achieved by

\[
X_{Gk}^{t+1} = U \text{Diag} \left( \text{Prox}_{w_i, \rho}^{\delta_i}(B_{Gk}^t) \right) V^T
\]

(30)

where \( \text{Prox}_{w_i, \rho}^{\delta_i}(B_{Gk}^t) \) denotes the element-wise operator, e.g.,

\[
\text{Prox}_{w_i, \rho}^{\delta_i}(B_{Gk}^t) = \arg\min_{\sigma_i(X) - \delta_i(B_{Gk}^t)} \frac{1}{2} (\sigma_i(X) - \delta_i(B_{Gk}^t))^2 + \lambda w_k \rho(\sigma_i(X))
\]

(31)

where \( \delta_1(B_{Gk}^t) \geq \delta_2(B_{Gk}^t) \geq \cdots \geq \delta_r(B_{Gk}^t) \).

We next give the closed-form solution of (31). It should be noted that the problem (31) is a weighted version with a nonconvex relaxation function \( \rho(\cdot) \). In this paper, we employ the popular \( L_p \)-function with \( p = 1/2 \) and \( p = 2/3 \) as the relaxation surrogates. The reason is due that the \( L_p \)-function is more flexible, moreover, two special cases of \( p = 1/2 \) and \( p = 2/3 \) have be demonstrated their high efficiency and can earn their closed-form solutions [32].

(1) \( L_p \)-norm with \( p = 1/2 \), then the problem (31) is reduced to

\[
\sigma_i(X_{Gk}) = \arg\min_{\sigma_i(X) - \delta_i(B_{Gk}^t)} \frac{1}{2} (\sigma_i(X) - \delta_i(B_{Gk}^t))^2 + \xi_i \left( \sigma_i(X_{Gk}) \right)^{1/2}
\]

(32)

then the closed-form of (31) can be defined by [33]
\[
\sigma^*_i(X_{G_k}) = \begin{cases} 
\frac{2}{3} \sigma_i(X_{G_k}) \left(1 + \cos \left(\frac{2\pi}{3} \varphi \left(\sigma_i(X_{G_k})\right)\right)\right), & |\sigma_i(X_{G_k})| > T \\
0, & \text{otherwise}
\end{cases}
\]

where \( \varphi(\sigma_i(X_{G_k})) = \cos^{-1}\left(\frac{i}{4} \left|\sigma_i(X_{G_k})\right|^{-3/2}\right) \). \( \xi_i = (\lambda w_{k,l}^\ell) \), and the threshold value \( T = \frac{3\sqrt{2}}{4} (2\xi_i)^{2/3} \).

(2) \( L_p \)-norm with \( p = 2/3 \), then the problem (31) is reduced to

\[
\sigma^*_i(X_{G_k}) = \arg \min_{\delta, \rho} \frac{1}{2} \left(\sigma_i(X_{G_k}) - \delta_i \left(B_{G_k}^\ell\right)\right)^2 + \xi_i \left(\sigma_i(X_{G_k})\right)^{2/3}
\]

then the closed-form of (31) can be defined by [32]

\[
\sigma^*_i(X_{G_k}) = \begin{cases} 
\left(\theta \left(\sigma_i(X_{G_k})\right) + \sqrt{2|\sigma_i(X_{G_k})|/\theta \left(\sigma_i(X_{G_k})\right)} - |\sigma_i(X_{G_k})|^{2/3}, & |\sigma_i(X_{G_k})| > T \\
0, & \text{otherwise}
\end{cases}
\]

where \( \theta(\sigma_i(X_{G_k})) = \frac{2}{\sqrt{3}} (2\xi_i)^{1/4} \left(cosh(\frac{1}{2} arccosh(\frac{2^2}{3} (2\xi_i)^{-3/2} (\sigma_i(X_{G_k})^2)))\right)^{1/2}, \) \( \xi_i = (\lambda w_{k,l}^\ell) \)

and \( T = \frac{2\sqrt{3}}{3} (2\xi_i)^{3/4} \) denotes the threshold value.

It should be noted that the sparse coefficient vector \( a_{G_k} \) in problem (10) is the singular value vector of matrix \( X_{G_k} \) in our proposed model (see subsection 2.3), hence we can achieved the \( a_{G_k}^{t+1} \) simultaneously from (26) over the self-adaptive dictionary \( D_{G_k} \). The whole algorithm for GSC can be summarized in the Algorithm 1.

**Algorithm 1**: Proposed iterative reweighted singular value function thresholding (GIR-SFT) algorithm for GSC by

\[
\hat{X}_{G_k} = \arg \min_{X_{G_k}} f(X_{G_k}) + \lambda \rho_w \left(\sigma(X_{G_k})\right)
\]

**Input**: The Observation \( Y \), the degradation matrix \( H \), \( D_{G_k} \);

**Initialization**: \( t = 0, \gamma = L_y, X^{0}_{G_k}, w_k^{0}; \)

**For** \( t = 1, 2, \cdots \) **do**

**While** not converge **do**

1. Updating \( B_{G_k}^\ell \) by \( X_{G_k}^\ell - \frac{1}{\gamma} \nabla f(X_{G_k}^\ell); \)

2. Updating \( X_{G_k}^{t+1}, a_{G_k}^{t+1} \) by solving (26);

3. Updating the weighting vector \( w_k^{t+1} \) by \( w_k^{t+1} \in \partial \rho \left(\sigma(X_{G_k}^{t+1})\right); \)

**End**

**Output**: \( \hat{X}_{G_k}, a_{G_k}^{t+1}. \)

---

**IV. Integrating Group Sparse Coding to Image Restoration via ADMM**
In this section, we will integrate the GSC framework to the image restoration problem. Considering the following image degraded model

\[ \mathbf{b} = \mathbf{A} \mathbf{x} + \mathbf{n} \]  

where \( \mathbf{b} \in \mathbb{R}^{N \times N} \) denotes the degraded observation, \( \mathbf{A} \in \mathbb{R}^{N \times N} \) denotes the degradation matrix, \( \mathbf{x} \in \mathbb{R}^{N \times N} \) and \( \mathbf{n} \) are the desired image and the additive noise. Different \( \mathbf{A} \) will cause different IR task, when \( \mathbf{A} \) is a compressed sampling operator, the IR problem becomes compressive sensing \([34]\), and an identity matrix of \( \mathbf{A} \) with entries either 1 or 0 will often cause image inpainting problem \([35]\). Suppose the original image \( \mathbf{X} \in \mathbb{R}^{N} \) (also \( \mathbf{X} \in \mathbb{R}^{N \times N} \)) can be represented by the sparse coefficient vector in the domain \( \Psi \), denotes as \( \mathbf{a} = \Psi \mathbf{X} \), that is \( \mathbf{X} = \mathbf{D} \mathbf{a} \), where \( \mathbf{D} \) denotes the corresponding dictionary, which can be known or learned from the images, and the IR problem can be described as the following generalized optimization model

\[ \hat{\mathbf{a}} = \arg \min_{\mathbf{a} \in \mathbb{Z}^N} \frac{1}{2} \| \mathbf{b} - \mathbf{A} \mathbf{D} \mathbf{a} \|_2^2 + \lambda R(\mathbf{a}) \]  

(37)

where \( R(\mathbf{a}) \) denotes regularization term, which measures the sparsity degree of the image and can provide prior knowledge for minimization, such as \( \| \mathbf{a} \|_p \), the parameter \( \lambda \) denotes the regularization parameter. Then our desired image can be reconstructed by \( \hat{\mathbf{X}} = \mathbf{D} \hat{\mathbf{a}} \).

4.1 ADMM framework for Image restoration via nonconvex weighted group sparse coding

According to the alternative direction method of multipliers (ADMM) framework, we introduce an auxiliary variable \( \mathbf{z} \) to the problem (37),

\[ \hat{\mathbf{a}} = \arg \min_{\mathbf{a} \in \mathbb{Z}^N} \frac{1}{2} \| \mathbf{b} - \mathbf{A} \mathbf{D} \mathbf{a} \|_2^2 + \lambda R(\mathbf{a}), \text{ s.t. } \mathbf{z} = \mathbf{D} \mathbf{a} \]  

(38)

Without confusion, we have the following iterative steps:

\[ \begin{align*}
\mathbf{z}^{(t+1)} &= \arg \min_{\mathbf{z}} \frac{1}{2} \| \mathbf{b} - \mathbf{A} \mathbf{z} \|_2^2 + \frac{\mu}{2} \| \mathbf{z} - \mathbf{D} \alpha^{(t)} - \mathbf{u}^{(t)} \|_2^2 \\
\alpha^{(t+1)} &= \arg \min_{\mathbf{a}} \frac{\mu}{2} \| \mathbf{z}^{(t+1)} - \mathbf{D} \mathbf{a} - \mathbf{u}^{(t)} \|_2^2 + \lambda R(\mathbf{a}) \\
\mathbf{u}^{(t+1)} &= \mathbf{u}^{(t)} - \left( \mathbf{z}^{(t+1)} - \mathbf{D} \alpha^{(t+1)} \right)
\end{align*} \]  

(39)

Then our optimization problem can be split into two subproblems of \( \mathbf{z} \) and \( \alpha \).

4.1.1 \( \mathbf{z} \)-subproblem

The \( \mathbf{z} \)-subproblem is a quadratic problem, which has a closed-form solution expressed as

\[ \mathbf{z} = (\mathbf{A}^T \mathbf{A} + \mu \mathbf{I})^{-1} (\mathbf{A}^T \mathbf{b} + \mu (\mathbf{D} \alpha + \mathbf{u})) \]  

(40)

where \( \mathbf{I} \) denotes the identity matrix. It is efficient to achieve the solution by (40) for image inpainting problem without computing the matrix inverse because of the specific structure in observation matrix \( \mathbf{A} \). However, it will be too time-consuming for CS reconstruction problem. To avoid the computing of matrix inverse, here, a gradient descent method is adopted to solve the \( \mathbf{z} \)-subproblem for CS reconstruction by

\[ \tilde{\mathbf{z}} = \mathbf{z} - \eta \mathbf{d} \]  

(41)

where the parameter \( \eta \) denotes the optimal step size, and \( \mathbf{d} \) denotes the gradient direction of

\[ \frac{1}{2} \| \mathbf{b} - \mathbf{A} \mathbf{z} \|_2^2 + \frac{\mu}{2} \| \mathbf{z} - \mathbf{D} \mathbf{a} - \mathbf{u} \|_2^2 \],

and we have

\[ \mathbf{d} = \mathbf{A}^T \mathbf{Az} - \mathbf{A}^T \mathbf{b} + \mu (\mathbf{z} - \mathbf{D} \mathbf{a} - \mathbf{u}). \]  

(42)

4.1.2 \( \alpha \)-subproblem

After achieving the \( \mathbf{z} \), the \( \mathbf{x} \) subproblem can be expressed as

\[ \alpha^{(t+1)} = \arg \min_{\mathbf{a}} \frac{\mu}{2} \| \mathbf{R}^{(t+1)} - \mathbf{D} \mathbf{a} \|_2^2 + \lambda R(\mathbf{a}) \]  

(43)
where \( \mathbf{R}^{(t+1)} = \mathbf{z}^{(t+1)} - \mathbf{u}^{(t)} \) can be regarded as the degraded observation of \( \mathbf{z}^{(t+1)} \). According to the theory described in the section 2.3, and by substituting \( \mathbf{X} = \mathbf{D}\mathbf{a} \), the optimization (43) becomes

\[
\mathbf{X}^{(t+1)} = \arg\min_{\mathbf{X}} \frac{\mu}{2} \| \mathbf{R}^{(t+1)} - \mathbf{X} \|_2^2 + \lambda R(\mathbf{X})
\]  

(44)

**Theorem 4.1** [36] Assume \( \mathbf{R}, \mathbf{X} \in \mathbb{R}^{N \times \sqrt{N}}, \mathbf{R}_{G_k}, \mathbf{X}_{G_k} \in \mathbb{R}^{R \times c} \), we define \( \mathbf{E} = \mathbf{R} - \mathbf{X} \) as the error matrix, and its each element is \( \mathbf{E}(j), j = 1, \cdots, \sqrt{N} \times \sqrt{N} \). If each element \( \mathbf{E}(j) \) is independent and the corresponding distribution of \( \mathbf{E} \) with zero mean and variance \( \delta^2 \), Then for any \( \epsilon > 0 \), we have the following relationship between \( \| \mathbf{R} - \mathbf{X} \|_2^2 \) and \( \| \mathbf{R}_{G_k} - \mathbf{X}_{G_k} \|_F^2 \), that is

\[
\lim_{N \to +\infty} P \left\{ \frac{1}{N} \| \mathbf{R} - \mathbf{X} \|_2^2 = \frac{1}{K} \sum_{k=1}^{n} \| \mathbf{R}_{G_k} - \mathbf{X}_{G_k} \|_F^2 < \epsilon \right\} = 1
\]  

(45)

To improve the performance for IR problem, we can generate the corresponding group matrix by grouping the similar patches, e.g., \( \mathbf{R}_{G_k} \), and \( \mathbf{X}_{G_k} \). As a result, we have the following equivalent equation with the probability (near to 1) according to **Theorem 4.1**, e.g.,

\[
\| \mathbf{R} - \mathbf{X} \|_2^2 = \frac{N}{K} \sum_{k=1}^{n} \| \mathbf{R}_{G_k} - \mathbf{R}(\mathbf{X}_{G_k}) \|_F^2
\]  

(46)

Then the problem (44) can be transformed into the following optimization problem,

\[
\mathbf{x}^{(t+1)} = \arg\min_{\mathbf{x}_{G_k}} \frac{1}{2} \sum_{k=1}^{m} \| \mathbf{x}_{G_k} - \mathbf{R}_{G_k} \|_F^2 + \tau \sum_{k=1}^{m} R(\mathbf{x}_{G_k})
\]  

(47)

where \( \tau = \frac{\lambda R}{\mu} \) with \( K = n \times c \times B \), and \( \mathbf{x}_{G_k} \in \mathbb{R}^{R \times c} \) denotes the group matrix with low-rank property. Accordingly, after combining our proposed generalized rank minimization based GSC model, the optimization (47) can be converted into the low-rank matrix recovery problem, that is

\[
\mathbf{x}^{(t+1)} = \arg\min_{\mathbf{x}_{G_k}} \frac{1}{2} \sum_{k=1}^{m} \| \mathbf{x}_{G_k} - \mathbf{R}_{G_k} \|_F^2 + \tau \sum_{k=1}^{m} \rho_{w_k} \left( \mathbf{R}(\mathbf{x}_{G_k}) \right)
\]  

(48)

where \( \rho_{w_k}(\mathbf{R}(\mathbf{x}_{G_k})) = \sum_{l=1}^{r} \rho \left( \mathbf{R}(\mathbf{x}_{G_k}) \right) \) denotes the proposed relaxation function. As a result, the optimization problem (48) can be split into \( m \) subproblems, and the \( k, (k = 1, 2, \cdots, m) \)-th subproblem can be shown as

\[
\mathbf{x}_{G_k}^{(t+1)} = \arg\min_{\mathbf{x}_{G_k}} \frac{1}{2} \| \mathbf{x}_{G_k} - \mathbf{R}_{G_k} \|_F^2 + \tau \sum_{l=1}^{r} \rho \left( \mathbf{R}(\mathbf{x}_{G_k}) \right)
\]  

(49)

With this, the optimization problem (48) can be solved by the our proposed GIR-SFT algorithm in **Algorithm 1**. It should be noted that problem (49) and (19) are two same minimization problems exactly with \( f(\mathbf{x}_{G_k}) = \frac{1}{2} \| \mathbf{x}_{G_k} - \mathbf{R}_{G_k} \|_F^2 \), therefore, the \( \mathbf{a} \)-subproblem in IR can be resolved via our proposed nonconvex weighted GSC model. Then the solution for problem (49) can be updated by

\[
\mathbf{x}_{G_k}^{t+1} = \arg\min_{\mathbf{x}_{G_k}} \frac{1}{2} \| \mathbf{x}_{G_k} - \mathbf{R}_{G_k}^{t} \|_F^2 + \frac{\tau}{\gamma} \sum_{k=1}^{m} w_{k,t} \rho \left( \mathbf{R}(\mathbf{x}_{G_k}^{t}) \right)
\]  

(50)

where \( \mathbf{C}_{G_k}^{t} = \mathbf{x}_{G_k}^{t} - \frac{1}{\gamma} \nabla f(\mathbf{x}_{G_k}^{t}) = \left( 1 - \frac{1}{\tau} \right) \mathbf{x}_{G_k}^{t} + \frac{1}{\gamma} \mathbf{R}_{G_k}^{t} \), and \( w_{k,t} \in \partial \rho \left( \mathbf{R}(\mathbf{x}_{G_k}^{t}) \right) \).
Then we can achieve the closed-form solution of (49) by **Algorithm 1**, i.e.,

\[ X_{G_k}^{t+1} = U_{R_{G_k}} \text{Prox}_{\rho}^\delta \left( \Theta (C_{G_k}) \right) V_{R_{G_k}}^T, \quad i = 1, \cdots, \min(B_c, c) \]  

(51)

where \( C_{G_k} = U_{R_{G_k}} \Sigma_{R_{G_k}} V_{R_{G_k}}^T \) denotes its singular value decomposition (SVD), \( \Sigma_{G_k} = \text{diag} (\xi_{G_k}, \cdots , \xi_{G_k, \min(B_c, c)}) \) and the operator \((\theta)_+ = \max(\theta, 0)\). As mentioned before, the sparse coefficient vector \( \alpha_{G_k} \) is the singular value vector of matrix \( X_{G_k} \) in our proposed model, hence we can achieved the \( \alpha_{G_k}^{t+1} \) simultaneously from (51).

### 4.2. Summary of the proposed algorithm

When all the groups \( \{X_{G_k}\}, k = 1, 2, \cdots, n \) are known, then the latent image \( X \) can be reconstructed by

\[ X = \sum_{k=1}^{n} R_k^T (X_{G_k}) / \sum_{k=1}^{n} R_k^T (1_{B_c \times c}) \]  

(52)

where the \( R_k^T (\cdot) \) denotes the transpose grouping operator, \( 1_{B_c} \) denotes a column matrix with the size of \( B_c \times c \) and all elements being 1, the operator \( ./ \) denotes an element-wise division of two matrix. The proposed whole algorithm of nonconvex weighted GSC for IR via ADMM can be summarized as the **Algorithm 2**.

---

**Algorithm 2**: Proposed ADMM-GIR-SFT for IR problem

**Input**: The Observation \( Y \), the degradation matrix \( H \);

**Initialization**: \( c, B_c, t = 0, \lambda, \mu, \gamma, u, z^{(0)} \),

for

if \( A \) is a mask operator

Updating \( z \) using the Eq. (39);

elseif \( A \) is a blur operator

Updating \( z \) using the Eq. (39);

else \( A \) is a random projection operator

Updating \( z \) using the Eq. (40) and (41);

end

Computing \( R^{(t+1)} = z^{(t+1)} - u^{(t)} \);

Constructing the groups \( \{R_{G_k}\} \);

for each group \( R_{G_k} \)

Construct adaptive dictionary \( D_{G_k} \) using Eq. (6), (7) and (8);

Reconstruct \( X_{G_k}^{(t+1)} \) and computing \( \alpha_{G_k}^{(t+1)} \);

end for

Updating \( D^{(t+1)} \) by concatenating all \( D_{G_k} \);

Updating \( \alpha^{(t+1)} \) by concatenating all \( \alpha_{G_k}^{(t+1)} \);

Computing \( u^{(t+1)} \);

Computing \( X^{(t+1)} \) by concatenating all the dictionaries \( \{X_{G_k}\} \);

\( t = t + 1 \);

end

**Output**: The reconstructed image \( X \).
V. Experimental Results

In this section, we will employ the typical $L_p, p = 1/2, 2/3$ as relaxation of nuclear norm to evaluate the effectiveness of our proposed method for typical IR tasks compared with several state-of-the-art competing IR algorithms. We also analyze the convergence of our proposed algorithm on various IR problems. To measure the reconstruction performance quantitatively, two popular metrics of PSNR and feature similarity (FSIM) [37] will be calculated. The widely used test images are employed to evaluate our proposed method presented in Fig. 3. All the best results will be highlighted in bold in the table, and all the simulation experiments are conducted in a personal computer with Intel (R) Core (TM) i7-6770HQ @ 2.6GHz CPU with 16 GB memory and a Windows 10 operating system.

![Experimental image set.](image)

Fig. 3. Experimental image set. (a), Typical 256 × 256 gray natural images: Barbara, Boats, Cameraman, Elain, Foreman, House, Leaves, Monarch, Starfish and Straw. (b), Typical 256 × 256 color images. Top: Barbara, Butterfly, Castle, Clock, Cowboy, Girl, House, Light; Bottom: Mickey, Peepers, Starfish, Vegetable, Zebra, Bike, Fence, Flower and Parrots. (c), Typical 512 × 512 gray natural images: Couple, Hill and Man.

5.1 Compressive Sensing

CS based image reconstruction technology aims to capture high-quality images from a small number of under-sampling random measurements, in which, one of the main technical challenges is how to obtain high-quality images while reducing the number of measurements. To evaluated the validation of our proposed method, four representative competing CS reconstruction algorithms is employed for comparisons, include the algorithms of BCS [38], SGSR [39], ALSB [36] and JASR [40]. We empirically set the similar patch parameter $c = 60$, the patch size of $\sqrt{B_c} \times \sqrt{B_c}$ is
selected as $6 \times 6$, the block size is $32 \times 32$, and the searching window of $L \times L$ is set to be $20 \times 20$, and $\varepsilon = 0.1$ for all of our CS experiments. The other parameter selection of $(\mu, \gamma)$ or $(\mu, p)$ is list in the Table 2 for different sub-sampling rate and surrogate functions. Table 3 and 4 list all achieved PSNR and FSIM results of four competing algorithms and our proposed method under different sampling rates of 0.1, 0.2, 0.3 and 0.4. We can observe that our proposed rank minimization based GSC method can obtain higher PSNR values and FSIM values than all competing methods. To make a visual comparison, we present the reconstructed results of boats and monarch from 0.1 measurements using our proposed method and other approaches, shown in the Fig. 4 and 5. These visual results demonstrate that our proposed method can reconstruct image with higher quality.

Table 2. The parameter selection for five penalty functions under different sub-rates

| Different penalties | 0.10  | 0.20  | 0.30  | 0.40  |
|---------------------|------|------|------|------|
| $p = \frac{1}{2} (\mu, \lambda)$ | $(8e - 3, 0.10)$ | $(3e - 2, 0.30)$ | $(5e - 3, 0.01)$ | $(2.5e - 2, 0.008)$ |
| $p = \frac{1}{4} (\mu, \lambda)$ | $(9e - 3, 0.04)$ | $(9e - 3, 0.01)$ | $(5e - 3, 0.0005)$ | $(2.5e - 2, 0.01)$ |

Table 3. The PSNR (dB) achieved by our proposed algorithm and four competing algorithms

| Subrate = 0.100 |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Method          | Barbara | Boats | Elaine | Foreman | House | Leaves | Monarch | Starfish |
| BCS             | 22.80  | 24.52 | 27.46  | 29.76   | 26.90 | 18.54  | 21.70   | 22.71   |
| SGSR            | 28.70  | 27.71 | 31.32  | 34.88   | 32.77 | 22.22  | 24.27   | 22.91   |
| ALSB            | 27.01  | 27.75 | 30.99  | 33.49   | 32.18 | 21.37  | 24.27   | 23.63   |
| JASR            | 29.58  | 28.59 | 32.01  | 35.61   | 33.49 | 23.62  | 25.83   | 24.39   |
| Proposed ($p = \frac{1}{2}$) | 29.99 | 29.04 | 32.39  | 36.08   | 34.07 | 25.82  | 27.32   | 25.34   |
| Proposed ($p = \frac{1}{4}$) | 29.77 | 29.12 | 32.20  | 35.95   | 34.02 | 25.93  | 27.34   | 25.58   |

| Subrate = 0.200 |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Method          | Barbara | Boats | Elaine | Foreman | House | Leaves | Monarch | Starfish |
| BCS             | 24.31  | 27.05 | 31.19  | 32.88   | 30.58 | 21.24  | 25.21   | 25.27   |
| SGSR            | 33.45  | 32.41 | 34.86  | 36.98   | 35.81 | 28.74  | 28.76   | 27.19   |
| ALSB            | 31.77  | 33.04 | 35.11  | 35.33   | 35.93 | 27.14  | 28.39   | 27.20   |
| JASR            | 34.16  | 33.21 | 35.66  | 37.87   | 36.10 | 30.24  | 30.60   | 29.10   |
| Proposed ($p = \frac{1}{2}$) | 34.62 | 34.03 | 36.15  | 38.79   | 37.18 | 31.44  | 31.79   | 29.96   |
| Proposed ($p = \frac{1}{4}$) | 34.63 | 34.09 | 36.05  | 38.61   | 37.18 | 31.36  | 31.63   | 29.80   |

| Subrate = 0.300 |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Method          | Barbara | Boats | Elaine | Foreman | House | Leaves | Monarch | Starfish |
| BCS             | 25.70  | 28.93 | 33.70  | 35.16   | 32.87 | 23.31  | 27.70   | 27.17   |
| SGSR            | 35.91  | 35.22 | 36.87  | 38.47   | 37.37 | 32.98  | 31.99   | 30.79   |
| ALSB            | 34.70  | 36.45 | 37.49  | 36.50   | 38.36 | 31.30  | 31.37   | 30.43   |
| JASR            | 36.59  | 36.08 | 36.83  | 38.54   | 38.04 | 33.70  | 33.63   | 32.33   |
| Proposed ($p = \frac{1}{2}$) | 37.13 | 37.17 | 38.13  | 41.15   | 39.42 | 35.23  | 34.75   | 33.30   |
| Proposed ($p = \frac{1}{4}$) | 37.18 | 37.11 | 38.20  | 41.11   | 39.38 | 35.17  | 34.81   | 33.26   |

| Subrate = 0.400 |
|-----------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Method          | Barbara | Boats | Elaine | Foreman | House | Leaves | Monarch | Starfish |
| BCS             | 26.30  | 29.32 | 34.23  | 36.84   | 35.70 | 24.31  | 28.70   | 28.17   |
### Table 4. The FSIM values achieved by proposed algorithm and four competing algorithms

#### Subrate = 0.100

| Method | Barbara | Boats | Elaine | Foreman | House | Leaves | Monarch | Starfish |
|--------|---------|-------|--------|---------|-------|--------|---------|----------|
| BCS    | 0.7891  | 0.8029| 0.8811 | 0.8911  | 0.8455| 0.6852 | 0.7828  | 0.8049   |
| SGSR   | 0.9147  | 0.8915| 0.9220 | 0.9393  | 0.9187| 0.8356 | 0.8371  | 0.8177   |
| ALSB   | 0.8903  | 0.8830| 0.9184 | 0.9254  | 0.9069| 0.7934 | 0.8218  | 0.8343   |
| JASR   | 0.9223  | 0.9035| 0.9282 | 0.9437  | 0.9167| 0.8799 | 0.8822  | 0.8516   |
| Proposed ($p = \frac{1}{2}$) | 0.9291 | 0.9088 | 0.9330 | 0.9509 | 0.9301 | 0.9157 | 0.9071 | 0.8758 |
| Proposed ($p = \frac{5}{6}$) | 0.9246 | 0.9086 | 0.9296 | 0.9491 | 0.9276 | 0.9149 | 0.9065 | 0.8783 |

#### Subrate = 0.200

| Method | Barbara | Boats | Elaine | Foreman | House | Leaves | Monarch | Starfish |
|--------|---------|-------|--------|---------|-------|--------|---------|----------|
| BCS    | 0.8429  | 0.8640| 0.9280 | 0.9296  | 0.9014| 0.7567 | 0.8465  | 0.8616   |
| SGSR   | 0.9615  | 0.9465| 0.9551 | 0.9598  | 0.9502| 0.9373 | 0.9132  | 0.8993   |
| ALSB   | 0.9501  | 0.9512| 0.9597 | 0.9460  | 0.9541| 0.9094 | 0.8965  | 0.8973   |
| JASR   | 0.9651  | 0.9521| 0.9603 | 0.9636  | 0.9425| 0.9516 | 0.9409  | 0.9295   |
| Proposed ($p = \frac{1}{2}$) | 0.9674 | 0.9591 | 0.9647 | 0.9699 | 0.9632 | 0.9619 | 0.9503 | 0.9409 |
| Proposed ($p = \frac{5}{6}$) | 0.9676 | 0.9596 | 0.9650 | 0.9699 | 0.9656 | 0.9621 | 0.9482 | 0.9400 |

#### Subrate = 0.300

| Method | Barbara | Boats | Elaine | Foreman | House | Leaves | Monarch | Starfish |
|--------|---------|-------|--------|---------|-------|--------|---------|----------|
| BCS    | 0.8780  | 0.8995| 0.9512 | 0.9504  | 0.9298| 0.8062 | 0.8839  | 0.8954   |
| SGSR   | 0.9762  | 0.9684| 0.9695 | 0.9711  | 0.9648| 0.9676 | 0.9469  | 0.9447   |
| ALSB   | 0.9716  | 0.9744| 0.9742 | 0.9575  | 0.9730| 0.9537 | 0.9296  | 0.9412   |
| JASR   | 0.9785  | 0.9723| 0.9661 | 0.9649  | 0.9649| 0.9719 | 0.9610  | 0.9580   |
| Proposed ($p = \frac{1}{2}$) | 0.9812 | 0.9776 | 0.9769 | 0.9822 | 0.9795 | 0.9808 | 0.9663 | 0.9661 |
| Proposed ($p = \frac{5}{6}$) | 0.9811 | 0.9772 | 0.9769 | 0.9819 | 0.9877 | 0.9804 | 0.9670 | 0.9657 |

#### Subrate = 0.400

| Method | Barbara | Boats | Elaine | Foreman | House | Leaves | Monarch | Starfish |
|--------|---------|-------|--------|---------|-------|--------|---------|----------|
| BCS    | 0.9069  | 0.9246| 0.9656 | 0.9646  | 0.9491| 0.8459 | 0.9115  | 0.9213   |
| SGSR   | 0.9835  | 0.9793| 0.9784 | 0.9788  | 0.9759| 0.9799 | 0.9648  | 0.9661   |
| ALSB   | 0.9830  | 0.9838| 0.9830 | 0.9871  | 0.9820| 0.9730 | 0.9581  | 0.9642   |
| JASR   | 0.9803  | 0.9764| 0.9742 | 0.9808  | 0.9676| 0.9831 | 0.9739  | 0.9677   |
| Proposed ($p = \frac{1}{2}$) | 0.9875 | 0.9861 | 0.9844 | 0.9881 | 0.9855 | 0.9895 | 0.9794 | 0.9791 |
| Proposed ($p = \frac{5}{6}$) | 0.9872 | 0.9857 | 0.9841 | 0.9883 | 0.9859 | 0.9897 | 0.9782 | 0.9790 |
Fig. 4. Original image and seven reconstructed images by BCS, SGSR ALSB, AL SB, JASR and our proposed algorithms for 0.1 measurements of boats. (a) Original Image; (b) BCS, 24.52 dB, FSIM=0.8029; (c) SGSR, PSNR=27.71 dB, FSIM=0.8915; (d) ALSB, PSNR=27.75 dB, FSIM=0.8830; (e) JASR, PSNR=28.59 dB, FSIM=0.9035; (f) Proposed (p = 1/2), PSNR=**29.04** dB, FSIM=0.9088; (g) Proposed (p = 2/3), PSNR=**29.12** dB, FSIM=0.9086;

Fig. 5. Visual comparison of the original image and five reconstructed images by BCS, ALSB, AL SB, GSR, JASR and our proposed algorithms for 0.1 measurements of monarch. (a) Original Image; (b) BCS, PSNR=21.70 dB, FSIM=0.7828; (c) SGSR, PSNR=24.27 dB, FSIM=0.8371; (d) ALSB, PSNR=24.27 dB, FSIM=0.8218; (e) JASR, PSNR=25.83 dB, FSIM=0.8822; (f) Proposed (p = 1/2), PSNR=27.32 dB, FSIM=0.9071; (g) Proposed (p = 2/3), PSNR=27.34 dB,
FSIM = 0.9065.

5.2 Image Inpainting

For the second applications, we employ our proposed nonconvex framework for image inpainting problem. Image inpainting aims at recovering the missing or damaged pixels in images in a plausible way. In this paper, we focus on two interesting cases of restoration from text removal and partial random samples. Here, the size of each patch is $\sqrt{B_c} \times \sqrt{B_c} = 10 \times 10$ and $8 \times 8$ for text removal and partial random sample, respectively. The similar patches number $c$ is set to be 60, and the searching window of $L \times L$ is set to be $20 \times 20$, and $\varepsilon = 0.1$ for all image inpainting experiments, and other parameters selection are listed in the Table 5.

| Different penalties | 80%      | 70%      | 50%      | Embed text |
|---------------------|----------|----------|----------|------------|
| $p = \frac{1}{2} (\mu, \lambda)$ | (5e – 3, 0.1) | (1e – 3, 0.025) | (1e – 3, 0.01) | (1e – 3, 0.01) |
| $p = \frac{1}{2} (\mu, \lambda)$ | (1.5e – 1, 0.06) | (1e – 1, 0.05) | (3e – 1, 0.0007) | (1e – 2, 0.005) |

5.2.1. Image restoration from partial random samples

We first handle the IR problem from partial random samples, where the degraded matrix $A$ is generated by a random matrix. We performed experiments by randomly removing 50% or 80% of pixels in original images. We employ four state-of-the-art image inpainting algorithms for comparisons, include the algorithms of BPFA [41] (an effective sparse image representations method via Bayesian dictionary learning), IPPO [42] (an IR method based on smooth ordering of patches), Aloha [43] (an recent image inpainting approach based on a low-rank Hankel matrix Approach) and JSM [44] (an IR method using joint statistical modeling scheme). Here, we employ eight color images for experiments. The results achieved by proposed algorithm and other competing state-of-the art algorithms are listed in the Table 6 and Table 7, our proposed approach can outperform other competitive methods significantly.

| Table 6. The PSNR (dB) values achieved by our algorithm and other four competing algorithms |
|-----------------------------------------------|
| Missing pixels = 80%                         |
| Method | Butterfly | Castle | House | Girl | Light | Mickey | Vegetable | Zebra |
| BPFA   | 24.04     | 23.94  | 30.16 | 24.80 | 19.39 | 24.53   | 23.26    | 20.99 |
| IPPO   | 25.13     | 24.50  | 33.65 | 25.31 | 21.51 | 26.33   | 23.07    | 22.71 |
| Aloha  | 24.88     | 23.89  | 33.80 | 25.16 | 21.49 | 25.33   | 23.04    | 22.74 |
| JSM    | 25.58     | 24.57  | 34.28 | 25.18 | 20.20 | 24.57   | 23.29    | 21.86 |
| Proposed ($p = \frac{1}{2}$) | 26.66     | 25.03  | 35.65 | 26.01 | 22.72 | 27.09   | 23.65    | 23.24 |
| Proposed ($p = \frac{1}{2}$) | 26.39     | 24.66  | 35.15 | 25.72 | 22.07 | 26.85   | 23.46    | 22.63 |

| Missing pixels = 70%                         |
|-----------------------------------------------|
| Method | Butterfly | Castle | House | Girl | Light | Mickey | Vegetable | Zebra |
| BPFA   | 26.52     | 25.68  | 33.96 | 26.61 | 21.36 | 26.17   | 24.66    | 22.73 |
| IPPO   | 27.68     | 26.11  | 36.64 | 27.43 | 23.47 | 28.59   | 24.80    | 24.76 |
| Aloha  | 27.33     | 25.84  | 36.76 | 27.12 | 23.19 | 27.10   | 24.52    | 24.54 |
| JSM    | 27.95     | 26.63  | 36.82 | 27.21 | 23.17 | 28.22   | 24.81    | 23.97 |
| Proposed ($p = \frac{1}{2}$) | 29.19 | 26.99 | 37.28 | 27.98 | 24.45 | 29.19 | 25.28 | 25.30 |
| Proposed ($p = \frac{3}{2}$) | 29.10 | 26.74 | 37.15 | 28.02 | 24.11 | 29.12 | 25.18 | 24.91 |

| Missing pixels = 50% |
|---|
| Method | Butterfly | Castle | House | Girl | Light | Mickey | Vegetable | Zebra |
| BPFA | 30.98 | 28.83 | 39.12 | 30.58 | 25.36 | 29.43 | 27.51 | 26.35 |
| IPPO | 31.69 | 29.57 | 40.03 | 31.05 | 26.76 | 32.74 | 27.99 | 28.42 |
| Aloha | 30.78 | 28.71 | 40.56 | 30.60 | 25.83 | 30.33 | 27.06 | 27.70 |
| JSM | 31.47 | 29.48 | 40.44 | 30.59 | 26.48 | 30.69 | 27.72 | 27.70 |

| Proposed ($p = \frac{1}{2}$) | 33.23 | 30.29 | 41.80 | 32.14 | 27.53 | 33.91 | 28.51 | 29.32 |
| Proposed ($p = \frac{3}{2}$) | 33.00 | 30.25 | 41.12 | 31.93 | 27.43 | 33.59 | 28.29 | 29.17 |

| Table 7. | The FSIM comparisons of proposed algorithm and other state-of-the-art algorithms |
|---|---|---|---|---|---|---|---|---|
| Method | Butterfly | Castle | House | Girl | Light | Mickey | Vegetable | Zebra |
| Missing pixels = 80% |
| BPFA | 0.8533 | 0.8639 | 0.8987 | 0.8766 | 0.8091 | 0.8696 | 0.8672 | 0.8190 |
| IPPO | 0.9078 | 0.8818 | 0.9488 | 0.8914 | 0.8757 | 0.9099 | 0.8619 | 0.8665 |
| Aloha | 0.8586 | 0.8729 | 0.9467 | 0.8832 | 0.8644 | 0.8770 | 0.8704 | 0.8565 |
| JSM | 0.9125 | 0.8829 | 0.9467 | 0.8871 | 0.8524 | 0.9060 | 0.8665 | 0.8520 |
| Proposed ($p = \frac{1}{2}$) | 0.9289 | 0.8990 | 0.9596 | 0.9107 | 0.9086 | 0.9234 | 0.8884 | 0.8902 |
| Proposed ($p = \frac{3}{2}$) | 0.9217 | 0.8899 | 0.9429 | 0.9015 | 0.9036 | 0.9162 | 0.8813 | 0.8817 |

| Missing pixels = 70% |
| Method | Butterfly | Castle | House | Girl | Light | Mickey | Vegetable | Zebra |
| BPFA | 0.8985 | 0.9059 | 0.9447 | 0.9168 | 0.8724 | 0.9024 | 0.9046 | 0.8706 |
| IPPO | 0.9400 | 0.9162 | 0.9695 | 0.9316 | 0.9173 | 0.9406 | 0.9070 | 0.9148 |
| Aloha | 0.8998 | 0.9103 | 0.9691 | 0.9215 | 0.9066 | 0.9104 | 0.9099 | 0.9003 |
| JSM | 0.9434 | 0.9227 | 0.9672 | 0.9277 | 0.9189 | 0.9357 | 0.9083 | 0.9035 |
| Proposed ($p = \frac{1}{2}$) | 0.9530 | 0.9269 | 0.9712 | 0.9384 | 0.9342 | 0.9448 | 0.9181 | 0.9222 |
| Proposed ($p = \frac{3}{2}$) | 0.9493 | 0.9253 | 0.9632 | 0.9374 | 0.9335 | 0.9428 | 0.9183 | 0.9227 |

| Missing pixels = 50% |
| Method | Butterfly | Castle | House | Girl | Light | Mickey | Vegetable | Zebra |
| BPFA | 0.9595 | 0.9486 | 0.9809 | 0.9598 | 0.9429 | 0.9501 | 0.9468 | 0.9325 |
| IPPO | 0.9724 | 0.9576 | 0.9853 | 0.9672 | 0.9591 | 0.9719 | 0.9536 | 0.9528 |
| Aloha | 0.9414 | 0.9485 | 0.9864 | 0.9608 | 0.9463 | 0.9515 | 0.9475 | 0.9468 |
| JSM | 0.9719 | 0.9588 | 0.9853 | 0.9662 | 0.9594 | 0.9685 | 0.9536 | 0.9545 |
| Proposed ($p = \frac{1}{2}$) | 0.9788 | 0.9637 | 0.9891 | 0.9750 | 0.9670 | 0.9769 | 0.9608 | 0.9671 |
| Proposed ($p = \frac{3}{2}$) | 0.9778 | 0.9629 | 0.9873 | 0.9732 | 0.9661 | 0.9749 | 0.9594 | 0.9652 |

To further demonstrate the visual effect of our proposed nonconvex framework, Fig. 8 (c) to (k) present nine recovered images from 20% random samples of ‘zebra’ by our proposed method and other state-of-the-art competing methods, and (a) (b) is the original image and the damaged image for comparisons. It can be observed obviously that our proposed method can recover the corrupted image with higher quality and can recover more image details effectively.
Fig. 6. Visual comparison of the original image, the damaged image, and nine reconstructed images by BPFA, IPPO, Aloha, JSM and our proposed algorithms for 80% missing of zebra. (a) Original Image; (b) The damaged image; (c) BPFA, PSNR=20.99 dB, FSIM=0.8190; (d) IPPO, PSNR=22.71 dB, FSIM=0.8665; (e) Aloha, PSNR=22.74 dB, FSIM=0.8565; (f) JSM, PSNR=21.86 dB, FSIM=0.8520; (g) Proposed (\(p = 1/2\)), PSNR=23.24 dB, FSIM=0.8902; (h) Proposed (\(p = 2/3\)), PSNR=22.63 dB, FSIM=0.8817.

5.2.2. Image inpainting for Text Removal

In this subsection, we will deal with another typical image inpainting problem for text removal, where the degraded matrix \(A\) is generated by a text mask. The goal of text removal is to reconstruct the original image from the degraded observations by removing the text. Table 8 and Table 9 displays our results achieved by our proposed algorithm and other four state-of-the-art algorithms, our proposed approach can outperform other competitive methods significantly. To make a visual comparison, Fig. 7 (c) to (h) present the case of text removal from a corrupted ‘Mickey’ image, (a) and (b) denote the original image and the corrupted image for comparison. From the results we can find obviously that our proposed method can reconstruct image with more details and can remove more artifacts effectively.

| Method     | Barbara | Butterfly | Castle | Clock | Cowboy | Mickey | Peppers | Starfish |
|------------|---------|-----------|--------|-------|--------|--------|---------|----------|
| BPFA       | 34.13   | 30.81     | 30.25  | 33.10 | 30.43  | 30.92  | 36.04   | 32.65    |
| IPPO       | 37.65   | 33.90     | 31.91  | 36.76 | 32.62  | 34.04  | 39.51   | 35.35    |
| Aloha      | 39.16   | 31.58     | 30.34  | 34.86 | 30.94  | 30.48  | 37.40   | 32.06    |
| JSM        | 37.75   | 33.05     | 32.26  | 35.86 | 32.41  | 32.93  | 39.31   | 35.18    |
| Proposed (\(p = 1/2\)) | **40.53** | **34.54** | **32.90** | **38.10** | **33.25** | **34.55** | **40.51** | **36.01** |
| Proposed (\(p = 2/3\)) | **40.93** | **34.67** | **32.78** | **38.29** | **33.02** | **34.56** | **40.29** | **36.20** |

**Table 8.** The PSNR (dB) values achieved by our proposed algorithm and four competing algorithms

**Table 8.** The PSNR (dB) values achieved by our proposed algorithm and four competing algorithms
Table 9. The FSIM results achieved by our proposed algorithm and four competing algorithms

| Text Removal | Method | Barbara | Butterfly | Castle | Clock | Cowboy | Mickey | Peppers | Starfish |
|--------------|--------|---------|-----------|--------|-------|--------|--------|---------|----------|
| BPFA         | 0.9781 | 0.9580  | 0.9571    | 0.9676 | 0.9669 | 0.9607 | 0.9775 | 0.9695  |
| IPPO         | 0.9839 | 0.9760  | 0.9903    | 0.9861 | 0.9805 | 0.9838 | 0.9916 | 0.9859  |
| Aloha        | 0.9906 | 0.9569  | 0.9672    | 0.9770 | 0.9741 | 0.9641 | 0.9866 | 0.9719  |
| JSM          | 0.9888 | 0.9830  | 0.9763    | 0.9838 | 0.9801 | 0.9810 | 0.9911 | 0.9558  |
| Proposed ($p = \frac{1}{2}$) | 0.9934 | 0.9849  | 0.9797    | 0.9857 | 0.9827 | 0.9853 | 0.9932 | 0.9880  |
| Proposed ($p = \frac{2}{3}$) | 0.9938 | 0.9848  | 0.9804    | 0.9898 | 0.9833 | 0.9856 | 0.9931 | 0.9884  |

5.3 Image deblurring

In this application, we focus on the case of image deblurring, where the blurred image is generated by the blur kernel and Gaussian noise with the standard deviation is $\sigma$. In this paper, three typical blur kernels are adopted in our experiments, the $9 \times 9$ Uniform blur kernel, the Gaussian blur kernel, and the Motion blur kernel. We compare our proposed nonconvex method to four recently approaches, i.e., MSEPLL [45], GSR-TNNM [28] and JSM [44]. It should be noted that the method of GSR-TNNM is a recently proposed IR method, and the MSEPLL and JSM are two state-of-the-art methods that can often achieve good results for image deblurring problem. The parameter selection for different typical blur kernels and functions are list in the Table 10. The achieved PSNR and FSIM results on all test images are list in the Table 11, and Fig. 10, 11 and 12 are presented the recovered under three blur kernels, we can find that our proposed method can produce the best results when the images are corrupted by the
uniform blur kernel and the motion blur kernel. For the case of Gaussian blur kernel, though our proposed method cannot achieve the best results in some cases, we still can achieved the higher average values of PSNRs and FSIMs, and from Fig. 12, we can see that our method can achieved much sharper image edges and cleaner textures that other methods.

Table 10. The parameter selection for different typical blur kernels and functions

| Different penalties | Uniform blur kernel | Gaussian blur kernel | Motion blur kernel |
|---------------------|---------------------|----------------------|--------------------|
| $p = \frac{1}{2} (\mu, \lambda)$ | $(1e - 2, 0.6)$ | $(8e - 3, 0.13)$ | $(1e - 2, 0.25)$ |
| $p = \frac{3}{2} (\mu, \lambda)$ | $(8e - 2, 0.006)$ | $(1.5e - 2, 0.003)$ | $(1e - 2, 0.005)$ |

Table 11. PSNR/FSIM comparisons for image deblurring

| Method       | Barbara   | Bike      | Fence     | Flower    | Parrots   | Pepper    |
|--------------|-----------|-----------|-----------|-----------|-----------|-----------|
| Deblurred    | 21.74/0.6720 | 19.56/0.6443 | 19.57/0.5949 | 23.53/0.6960 | 23.85/0.8265 | 24.03/0.7869 |
| MSEPLL       | 25.24/0.8791 | 24.15/0.8570 | 27.18/0.8990 | 28.09/0.8847 | 29.89/0.9246 | 31.23/0.9309 |
| GSR-TNNM     | 26.68/0.8857 | 24.30/0.8375 | 28.38/0.9124 | 28.10/0.8697 | 29.59/0.9049 | 30.43/0.9003 |
| JSM          | 25.68/0.8615 | 23.89/0.8368 | 27.25/0.9038 | 26.88/0.8420 | 27.87/0.8265 | 28.26/0.8458 |
| Proposed ($p = \frac{1}{2}$) | 27.64/0.9053 | 25.34/0.8645 | 29.67/0.9309 | 28.96/0.8880 | 30.65/0.9154 | 31.38/0.9168 |
| Proposed ($p = \frac{3}{2}$) | 27.40/0.9091 | 25.07/0.8609 | 29.76/0.9285 | 29.21/0.8932 | 30.87/0.9359 | 32.14/0.9332 |

Gaussian blur kernel: $\text{fspecial (“gaussian”, 25, 1.6)} \quad \sigma = \sqrt{2}$

| Method       | Barbara   | Bike      | Fence     | Flower    | Parrots   | Pepper    |
|--------------|-----------|-----------|-----------|-----------|-----------|-----------|
| Deblurred    | 22.79/0.7568 | 22.19/0.7708 | 22.14/0.6830 | 26.57/0.8285 | 26.91/0.8937 | 27.79/0.8876 |
| MSEPLL       | 23.83/0.8610 | 26.00/0.9021 | 25.97/0.8936 | 29.98/0.9200 | 31.49/0.9460 | 33.28/0.9530 |
| GSR-TNNM     | 26.76/0.9012 | 26.64/0.8854 | 27.58/0.9128 | 30.50/0.9108 | 31.91/0.9338 | 32.75/0.9279 |
| JSM          | 25.78/0.8738 | 26.65/0.8890 | 27.08/0.9028 | 30.01/0.8956 | 31.07/0.8928 | 31.98/0.9097 |
| Proposed ($p = \frac{1}{2}$) | 26.75/0.8952 | 26.65/0.8844 | 27.52/0.9111 | 30.32/0.9045 | 31.59/0.9148 | 32.32/0.9158 |
| Proposed ($p = \frac{3}{2}$) | 26.73/0.9006 | 26.85/0.8948 | 27.64/0.9150 | 30.71/0.9154 | 32.13/0.9392 | 33.21/0.9385 |

Motion blur kernel: $\text{fspecial (“motion”, 20, 45)} \quad \sigma = \sqrt{2}$

| Method       | Barbara   | Bike      | Fence     | Flower    | Parrots   | Pepper    |
|--------------|-----------|-----------|-----------|-----------|-----------|-----------|
| Deblurred    | 20.92/0.6678 | 17.99/0.6303 | 19.65/0.5641 | 22.35/0.6966 | 22.48/0.8201 | 22.17/0.7656 |
| MSEPLL       | 25.98/0.8944 | 25.54/0.8924 | 25.64/0.8763 | 28.99/0.9022 | 31.49/0.9318 | 30.67/0.9283 |
| GSR-TNNM     | 28.45/0.9072 | 25.76/0.8688 | 27.55/0.9030 | 29.63/0.8956 | 31.56/0.9197 | 30.96/0.9074 |
| JSM          | 24.95/0.8464 | 23.19/0.8342 | 25.40/0.8809 | 25.69/0.8175 | 26.48/0.7878 | 26.38/0.8071 |
| Proposed ($p = \frac{1}{2}$) | 30.14/0.9371 | 27.40/0.9070 | 29.05/0.9226 | 30.71/0.9167 | 33.13/0.9452 | 33.03/0.9448 |
| Proposed ($p = \frac{3}{2}$) | 30.38/0.9397 | 27.48/0.9070 | 29.11/0.9228 | 30.78/0.9171 | 33.20/0.9463 | 33.15/0.9463 |

Table 12. The averaged PSNR/FSIM values achieved by our proposed method and three competing methods for image inpainting corrupted by Gaussian blur kernel ($\text{fspecial (“gaussian”, 25, 1.6)} \quad \sigma = \sqrt{2}$).

| Methods  | Deblurred | MSEPLL | GSR-TNNM | JSM | Proposed ($p = \frac{1}{2}$) | Proposed ($p = \frac{3}{2}$) |
|----------|-----------|--------|----------|-----|-----------------------------|-----------------------------|
| PSNR     | 24.73     | 28.43  | 29.35    | 28.76 | 29.19                        | 29.55                        |
| FSIM     | /0.8034   | /0.9126| /0.9120  | /0.8940| /0.9043                      | /0.9173                      |
Fig. 8. Restored Fence images form corrupted image by \( 9 \times 9 \) Uniform blur kernel with \( \sigma = \sqrt{2} \) by various methods for visual comparison. (a) Original Image; (b) The deblurred and noisy image, PSNR=19.57 dB, FSIM=0.5949; (c) MSEPLL, PSNR=27.18 dB, FSIM=0.8990; (d) GSR-TNNM, 28.38 dB, FSIM=0.9124; (e) JSM, PSNR=27.25 dB, FSIM=0.9038; (f) Proposed \( (p = 1/2) \), PSNR=29.67 dB, FSIM=0.9039; (g) Proposed \( (p = 2/3) \), PSNR=29.76 dB, FSIM=0.9285.

Fig. 9. Restored Fence images form corrupted image by Gaussian blur kernel with \( \sigma = \sqrt{2} \) by various methods for visual comparison. (a) Original Image; (b) The deblurred and noisy image, PSNR=22.14 dB, FSIM=0.6830; (c) MSEPLL, PSNR=25.97 dB, FSIM=0.8936; (d) GSR-TNNM, 27.58 dB, FSIM=0.9128; (e) JSM, PSNR=27.08 dB, FSIM=0.9028; (f) Proposed \( (p = 1/2) \), PSNR=27.52 dB, FSIM=0.9111; (g) Proposed \( (p = 2/3) \), PSNR=27.64 dB, FSIM=0.9150.
5.4 Salt and Pepper Noise Removal

As a typical impulsive noise, the salt and pepper noise (SPN) usually occurs in the procedure of the image acquisition and transmission. In this application, we will adopt our proposed algorithm to recover the corrupted image by SPN with different strength, here, three state-of-the-art algorithms are compared with our method are WESNR [46], JSM [44] and WCSR [47]. To handle this problem, an adaptive median filter [48] is first applied to the corrupted images, hence we can identify the makt matrix $A$, and we can transform the noise removal problem into the IR problem. Table 13 details the parameter selection for different penalties and the noise density. Table 14 presents all the PSNR/FSIM results on six test images, we can find that our proposed method can outperform other three methods significantly. Moreover, some visual results of ‘Cameraman’ for five algorithms are presented in the figure 9, by comparing with the competing results in (c), (d) and (e), we can obviously find that our method can achieved the most visually results, e.g., (f) and (g).

| Table 13. The parameter selection for different sub-rate and functions |
|-------------------------------------------------|
| Different penalties | 30% SPN $\mu$, $\lambda$ | 50% SPN $\mu$, $\lambda$ | 80% SPN $\mu$, $\lambda$ |
| $p = \frac{1}{2}$ ($\mu$, $\lambda$) | $(1e - 2, 0.15)$ | $(5e - 3, 0.1)$ | $(5e - 3, 0.4)$ |
| $p = \frac{2}{3}$ ($\mu$, $\lambda$) | $(1e - 1, 0.05)$ | $(1e - 2, 0.01)$ | $(2e - 2, 0.03)$ |

| Table 14. PSNR/FSIM comparison for images corrupted by salt and pepper noise |
|---------------------------------|
| 30% SPN |
| Method | Cameraman | Couple512 | Hill512 | House | Man512 | Straw |
| Noisy | 10.33/0.4542 | 10.76/0.6369 | 10.61/0.5731 | 10.75/0.4081 | 10.64/0.5864 | 10.64/0.6698 |
| WESNR | 28.15/0.9303 | 32.43/0.9800 | 33.16/0.9827 | 36.26/0.9548 | 32.99/0.9816 | 28.24/0.9442 |
**Table 1.** Visual comparison of the original image, the corrupted image by 50% salt and pepper noise, and five reconstructed images by MSEPLL, GSR-TNNM, JSM and our proposed algorithms. (a) Original Image; (b) The noisy image; (c) WESNR, PSNR=25.29 dB, FSIM=0.8901; (d) JSM, 28.60 dB, FSIM=0.8865; (e) WCSR, PSNR=28.98 dB, FSIM=0.9383; (f) Proposed (p = 1/2), PSNR=30.65 dB, FSIM=0.9606; (g) Proposed (p = 2/3), PSNR=30.66 dB, FSIM=0.9601.

| Method          | Cameraman | Couple512 | Hill512 | House | Man512 | Straw |
|-----------------|-----------|-----------|---------|-------|--------|-------|
| **50% SPN**     |           |           |         |       |        |       |
| Noisy           | 8.10/0.3610 | 8.53/0.5546 | 8.39/0.4977 | 8.53/0.4081 | 8.42/0.5864 | 8.43/0.5931 |
| WESNR           | 25.29/0.8901 | 30.40/0.9665 | 31.32/0.9675 | 34.39/0.3188 | 30.92/0.5075 | 25.77/0.9039 |
| JSM             | 28.60/0.8865 | 30.79/0.9635 | 30.78/0.9519 | 34.51/0.9140 | 30.49/0.9522 | 25.98/0.9023 |
| WCSR            | 28.98/0.9383 | 32.47/0.9842 | 33.51/0.9852 | 35.75/0.9681 | 32.51/0.0937 | 27.68/0.9479 |
| Proposed (p = 1/2) | 30.65/0.9606 | 34.27/0.9897 | 34.92/0.9887 | 40.08/0.9855 | 34.24/0.9898 | 30.35/0.9682 |
| Proposed (p = 2/3) | 30.66/0.9601 | 34.25/0.9896 | 34.86/0.9883 | 39.87/0.9841 | 34.09/0.9893 | 30.36/0.9682 |
| **80% SPN**     |           |           |         |       |        |       |
| Noisy           | 6.04/0.3012 | 6.48/0.4898 | 6.33/0.4354 | 6.44/0.2651 | 6.37/0.4422 | 6.36/0.5244 |
| WESNR           | 20.32/0.7907 | 25.91/0.9058 | 27.93/0.9210 | 27.12/0.8595 | 27.09/0.9194 | 20.96/0.7439 |
| JSM             | 24.18/0.7949 | 26.95/0.9005 | 27.83/0.8901 | 30.98/0.8737 | 27.14/0.8811 | 20.90/0.6971 |
| WCSR            | 23.15/0.8340 | 26.45/0.9302 | 28.26/0.9435 | 28.60/0.8906 | 27.26/0.9366 | 21.52/0.8250 |
| Proposed (p = 1/2) | 25.62/0.8866 | 28.85/0.9543 | 29.77/0.9526 | 34.10/0.9481 | 29.13/0.9512 | 24.32/0.8840 |
| Proposed (p = 2/3) | 25.64/0.8876 | 28.86/0.9551 | 29.91/0.9525 | 34.05/0.9451 | 29.12/0.9510 | 24.48/0.8914 |
5.5 Convergence analysis

Although our proposed regularization model can obtain better performance, it is tractable to give a theoretical proof of the convergence for our proposed algorithm since the relaxation function of nuclear norm is nonconvex. It is well documented that the behavior of PSNR curves versus the iteration number can reflect the convergence property visually. Fig. 10 (a) (b) and (c) present the PSNRs curves for different IR problems, including the image compressive sensing problem under different sub-sampling rates, image inpainting problem from partial random samples, image inpainting for text removal, and SPN removal. It is obviously that our proposed algorithm contains good convergence property and robustness.

VI. Conclusion

This paper proposed a novel group sparse coding framework via rank minimization. We first convert the group sparse coding problem into the low-rank minimization problem via an effective adaptive dictionary learning strategy. To better approximate the rank of the group matrix, we developed a novel double nonconvex and nonsmooth rank minimization framework for GSC. To solve this nonconvex optimization problem, an iteratively reweighted singular value function thresholding (GIR-SFT) based GSC algorithm is proposed. Finally, some typical IR applications
have been considered to evaluate the effectiveness and priority of our proposed method via the ADMM.

Reference

[1] J. Mairal, M. Elad, and G. Sapiro, “Sparse representation for color image restoration,” *IEEE Trans. image Process.*, vol. 17, no. 1, pp. 53–69, 2008.
[2] L. Zhang and W. Zuo, “Image Restoration: From Sparse and Low-Rank Priors to Deep Priors,” *IEEE Signal Process. Mag.*, vol. 34, no. 5, pp. 172–179, 2017.
[3] K. Zhang, W. Zuo, S. Gu, and L. Zhang, “Learning Deep CNN Denoiser Prior for Image Restoration,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 3929–3938.
[4] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
[5] X. Huang and M. Yan, “Nonconvex penalties with analytical solutions for one-bit compressive sensing,” *Signal Processing*, vol. 144, pp. 341–351, 2018.
[6] Z. Zha, X. Zhang, Q. Wang, Y. Bai, and L. Tang, “Image denoising using group sparsity residual and external nonlocal self-similarity prior,” in *2017 IEEE International Conference on Image Processing*, 2017, pp. 2956–2960.
[7] Z. Zha et al., “Analyzing the group sparsity based on the rank minimization methods,” in *IEEE International Conference on Multimedia and Expo*, 2017, no. July, pp. 883–888.
[8] J. P. Oliveira, J. M. Bioucas-Dias, and M. A. T. Figueiredo, “Adaptive total variation image deblurring: A majorization-minimization approach,” *Signal Processing*, vol. 89, no. 9, pp. 1683–1693, 2009.
[9] S. Wang et al., “Dictionary learning based impulse noise removal via L1-L1 minimization,” *Signal Processing*, vol. 93, no. 9, pp. 2696–2708, 2013.
[10] Y. Wang, J. Yang, W. Yin, and Y. Zhang, “A New Alternating Minimization Algorithm for Total Variation Image Reconstruction,” *SIAM J. Imaging Sci.*, vol. 1, no. 3, pp. 248–272, 2008.
[11] H. Zhang, L. Tang, Z. Fang, C. Xiang, and C. Li, “Nonconvex and nonsmooth total generalized variation model for image restoration,” *Signal Processing*, vol. 143, pp. 69–85, 2018.
[12] Y. Li et al., “Sparse Adaptive Iteratively-Weighted Thresholding Algorithm (SAITA) for Lp-Regularization Using the Multiple Sub-Dictionary Representation,” *Sensors*, vol. 17, no. 12, pp. 2920–2936, 2017.
[13] J. Fan and R. Li, “Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties,” *J. Am. Stat. Assoc.*, vol. 96, no. 456, pp. 1348–1360, 2001.
[14] J. H. Friedman, “Fast sparse regression and classification,” *Int. J. Forecast.*, vol. 28, no. 3, pp. 722–738, 2012.
[15] C. H. Zhang, *Nearly unbiased variable selection under minimax concave penalty*, vol. 38, no. 2, 2010.
[16] F. Wen, P. Liu, Y. Liu, R. C. Qiu, and W. Yu, “Robust sparse recovery in impulsive noise via Lp-L1 Optimization,” *IEEE Trans. Signal Process.*, vol. 65, no. 1, pp. 105–118, 2017.
[17] F. Wen, R. Ying, P. Liu, and R. C. Qiu, “Robust PCA Using Generalized Nonconvex Regularization,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. PP, no. 2, pp. 1–1, 2019.
[18] F. Wen, L. Chu, P. Liu, and R. C. Qiu, “A Survey on Nonconvex Regularization-Based Sparse
and Low-Rank Recovery in Signal Processing, Statistics, and Machine Learning,” *IEEE Access*, vol. 6, pp. 69883–69906, 2018.

[19] M. Jung, X. Bresson, T. F. Chan, and L. A. Vese, “Nonlocal mumford-shah regularizers for color image restoration,” *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1583–1598, 2011.

[20] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng, “Patch group based nonlocal self-similarity prior learning for image denoising,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 244–252.

[21] W. Dong, L. Zhang, G. Shi, and X. Li, “Nonlocally centralized sparse representation for image restoration,” in *IEEE Transactions on Image Processing*, 2013, vol. 22, no. 4, pp. 1620–1630.

[22] Y. Liu, X. Yuan, J. Suo, D. Brady, and Q. Dai, “Rank Minimization for Snapshot Compressive Imaging,” *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1–18, 2018.

[23] J. Zhang, D. Zhao, and W. Gao, “Group-Based Sparse Representation for Image Restoration,” *Image Process. IEEE Trans.*, vol. 23, no. 8, pp. 3336–3351, 2014.

[24] 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, 2013.

[25] F. Nie, H. Huang, and C. Ding, “Low-Rank Matrix Recovery via Efficient Schatten p-Norm Minimization,” in *AAAI Conference on Artificial Intelligence*, 2012, pp. 655–661.

[26] C. Lu, J. Tang, S. Yan, and Z. Lin, “Generalized nonconvex nonsmooth low-rank minimization,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2014, pp. 4130–4137.

[27] C. Lu, C. Zhu, C. Xu, S. Yan, and Z. Lin, “Generalized Singular Value Thresholding,” in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015, pp. 1805–1811.

[28] T. Geng, G. Sun, Y. Xu, and J. He, “Truncated Nuclear Norm Minimization Based Group Sparse Representation,” *SIAM J. Imaging Sci.*, vol. 11, no. 3, pp. 1878–1897, 2018.

[29] S. Gu, L. Zhang, W. Zuo, and X. Feng, “Weighted nuclear norm minimization with application to image denoising,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2014, vol. 2, no. 1, pp. 2862–2869.

[30] 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, 2016.

[31] H. Zhang, C. Gong, J. Qian, B. Zhang, C. Xu, and J. Yang, “Efficient Recovery of Low-Rank Matrix via Double Nonconvex Nonsmooth Rank Minimization,” *IEEE Trans. Neural Networks Learn. Syst.*, pp. 1–10, 2019.

[32] W. Cao, J. Sun, and Z. Xu, “Fast image deconvolution using closed-form thresholding formulas of $L^p(q=1/2,2/3)$ regularization,” *J. Vis. Commun. Image Represent.*, vol. 24, no. 1, pp. 31–41, 2013.

[33] Z. Xu, H. Zhang, Y. Wang, X. Chang, and Y. Liang, “$L_{1/2}$ Regularization: A Thresholding Representation Theory and a Fast Solver,” *IEEE Transactions Neural Netw. Learn. Syst.*, vol. 23, no. 7, pp. 1013–1027, 2012.

[34] Y. Li, F. Dai, X. Cheng, L. Xu, and G. Gui, “Multiple-prespecified-dictionary sparse representation for compressive sensing image reconstruction with nonconvex regularization,” *J. Franklin Inst.*, vol. 356, no. 4, pp. 2353–2371, 2019.
[35] C. Guillemot and O. Le Meur, “Image Inpainting: Overview and Recent Advances,” *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 127–144, 2014.

[36] J. Zhang, C. Zhao, D. Zhao, and W. Gao, “Image compressive sensing recovery using adaptively learned sparsifying basis via L0 minimization,” *Signal Processing*, vol. 103, pp. 114–126, 2014.

[37] L. Zhang, L. Zhang, X. Mou, and D. Zhang, “FSIM: A Feature Similarity Index for Image Quality Assessment,” *IEEE Transactions Image Process.*, vol. 20, no. 8, pp. 2378–2386, 2011.

[38] Sungkwang Mun and J. E. Fowler, “Block compressed sensing of images using directional transforms,” in *Proceedings of the International Conference on Image Processing*, 2009, pp. 3021–3024.

[39] J. Zhang, D. Zhao, F. Jiang, and W. Gao, “Structural group sparse representation for image Compressive Sensing recovery,” in *Data Compression Conference Proceedings*, 2013, pp. 331–340.

[40] N. Eslahi and A. Aghagolzadeh, “Compressive Sensing Image Restoration Using Adaptive Curvelet Thresholding and Nonlocal Sparse Regularization,” *IEEE Transactions Image Process.*, vol. 25, no. 7, pp. 3126–3140, 2016.

[41] M. Zhou, H. Chen, L. Ren, G. Sapiro, and L. Carin, “Non-parametric Bayesian dictionary learning for sparse image representations,” in *Neural Information Processing Systems (NIPS2009), Vancouver, Canada, Dec. 2009*, 2009, pp. 1–9.

[42] I. Ram, M. Elad, and I. Cohen, “Image processing using smooth ordering of its patches,” *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2764–2774, 2013.

[43] K. H. Jin and J. C. Ye, “Annihilating Filter-Based Low-Rank Hankel Matrix Approach for Image Inpainting,” *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3498–3511, 2015.

[44] J. Zhang, D. Zhao, R. Xiong, S. Ma, and W. Gao, “Image restoration using joint statistical modeling in a space-Transform domain,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 6, pp. 915–928, 2014.

[45] V. Papyan and M. Elad, “Multi-scale patch-based image restoration,” *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 249–261, 2016.

[46] Jielin Jiang, Lei Zhang, and Jian Yang, “Mixed Noise Removal by Weighted Encoding With Sparse Nonlocal Regularization,” *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2651–2662, 2014.

[47] C. Lung, P. Chen, L. Liu, L. Chen, and Y. Y. Tang, “Weighted Couple Sparse Representation With Classified Regularization for Impulse Noise Removal,” *IEEE Transactions Image Process.*, vol. 24, no. 11, pp. 4014–4026, 2015.

[48] H. Hwang and R. A. Haddad, “Adaptive Median Filters: New Algorithms and Results,” *IEEE Trans. Image Process.*, vol. 4, no. 4, pp. 499–502, 1995.