

## Article

# Structural Smoothing Low-Rank Matrix Restoration Based on Sparse Coding and Dual-Weighted Model

Jiawei Wu and Hengyou Wang \* 

School of Science, Beijing University of Civil Engineering and Architecture, Beijing 100044, China; 2107010420009@stu.bucea.edu.cn

\* Correspondence: wanghengyou@bucea.edu.cn; Tel.: +86-135-0121-4858

**Abstract:** Group sparse coding (GSC) uses the non-local similarity of images as constraints, which can fully exploit the structure and group sparse features of images. However, it only imposes the sparsity on the group coefficients, which limits the effectiveness of reconstructing real images. Low-rank regularized group sparse coding (LR-GSC) reduces this gap by imposing low-rankness on the group sparse coefficients. However, due to the use of non-local similarity, the edges and details of the images are over-smoothed, resulting in the blocking artifact of the images. In this paper, we propose a low-rank matrix restoration model based on sparse coding and dual weighting. In addition, total variation (TV) regularization is integrated into the proposed model to maintain local structure smoothness and edge features. Finally, to solve the problem of the proposed optimization, an optimization method is developed based on the alternating direction method. Extensive experimental results show that the proposed SDWLR-GSC algorithm outperforms state-of-the-art algorithms for image restoration when the images have large and sparse noise, such as salt and pepper noise.

**Keywords:** group sparse coding; low-rank regularized group sparse coding; TV norm; dual-weighted; image restoration



**Citation:** Wu, J.; Wang, H. Structural Smoothing Low-Rank Matrix Restoration Based on Sparse Coding and Dual-Weighted Model. *Entropy* **2022**, *24*, 946. <https://doi.org/10.3390/e24070946>

Academic Editor: Chun-Hung Liu

Received: 16 June 2022

Accepted: 5 July 2022

Published: 7 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Sparse representation theory [1–3] has always been a very interesting research field due to its good performance in the fields of computer vision and image processing. It can be well used to represent or extract the main features of images. Among existing works, the image denoising methods based on patch sparse coding [4] divide the image into each patch equally, and use the patch's structure to encode the image into a combination of dictionary and sparse coefficient. By constraining the  $L_0$  norm of coding coefficients, the real coding can be approximately estimated on the noisy image, so as to remove the noise.

On this topic, the dictionary is very important for the performance of these methods and should be learned firstly. Thus, the popular technologies such as PCA [5], K-SVD [6], and S-KSVD [7] are used to train dictionaries, to achieve higher expressive ability. However, they are all large-scale and highly non-convex problems, which often have high computational complexity. On the other hand, patches are units of sparse representation. Each patch is usually considered independently in dictionary learning and sparse coding, which leads to its focus on the local structure of the image and essentially ignores the relationship between similar patches—that is, the non-local self similarity (NSS) of the image [8–11].

For these non-local image denoising approaches, they provide a new research theory and direction for image denoising. Group sparse coding (GSC) [12], which uses groups instead of a single patch as the basic unit of sparse coding, combines the advantages of local sparsity and NSS of images, and shows great potential in various image processing tasks [13,14]. Similar to the sparse representation based on patches, each image patch group can also be accurately fitted by the sparse linear combination of dictionary atoms. In order to make full use of the similarity between groups, Zha et al. proposed low-rank regulated

group sparse coding (LR GSC) [15], which imposed a low-rank constraint on the sparse coefficients of each group. It is the first time that the low-rank property of group sparsity in GSC has been used to ensure that the dictionary domain coefficients are not only sparse but also low-rank. Although the above methods have achieved good results, they are mainly aimed at the removal of Gaussian white noise, and cannot be used to remove the outlier noise effectively. In addition, when the density of noise is large, the acquisition of similar patches will be affected since each image patch also contains these noisy pixels. Figure 1 shows the effect of the LR-GSC model on image restoration from Gaussian noise and outlier noise, respectively. It can be seen that the LR-GSC model is more suitable for removing Gaussian noise than outlier noise.



**Figure 1.** (a) The Lena image corrupted by Gaussian noise. (b) Restoration result of (a) based on LR-GSC. (c) The Lena image corrupted by outlier noise. (d) Restoration result of (c) based on LR-GSC.

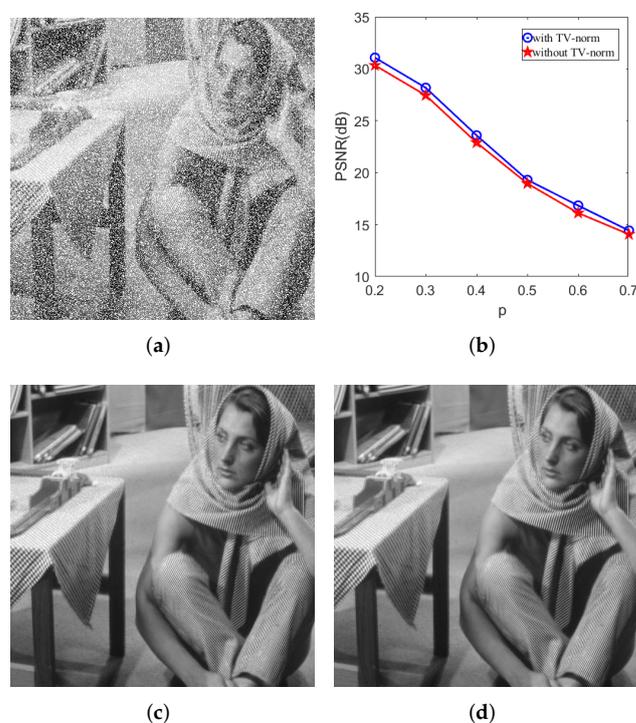
Because of the two-dimensional structure of a matrix, it can effectively protect the original details and structure information of the images. It is also more robust to data with outliers—for example, salt and pepper noise. Wright et al. [16] proposed the problem of low-rank restoration, which decomposes the original data as the sum of the low-rank matrix and sparse noise matrix. Candes et al. [17] accurately removed large low-rank matrices containing noise samples through kernel norm minimization. Using this theory to denoise the image, the original image structure with low-rank characteristics is separated from the observed noisy image, which has good performance on the removal of impulse and outlier noise.

When the rank or sparsity exceeds the threshold limit, the convex approximation model will not be able to accurately estimate the low-rank solution and sparse solution. For this reason, Candes et al. [18] proposed a weighted  $L_1$  norm method to allocate smaller weights to larger matrix elements, restrain the overshrinkage of the  $L_1$  norm, and improve the accuracy of the sparse solution. Gu et al. [19] proposed the weighted nuclear norm minimization (WNNM) model and the WNNM-based RPCA model (WNNM-RPCA) [20]. The model uses the weighted nuclear norm to relax the rank function. Considering the importance of different rank components, they assign weights according to the size of singular values to control the penalization of different rank components, so as to retain more important rank components and improve the accuracy of low-rank solutions. Peng et al. [21]

proposed a double-weighted model, which simultaneously weighted the sparse and low-rank terms in the RPCA model, and was combined with the reweighted method to allocate weights in the iterative algorithm. The model can improve the accuracy of the sparse solution and low-rank solution at the same time. Although the low-rank theory performs well in removing impulse or outlier noise, it is mainly based on local similarity—that is, only the overall low-rank structure is considered.

Therefore, in order to overcome the above shortcomings, we propose a smoothing dual-weighted low-rank group sparse coding (SDWLR-GSC) algorithm to remove impulse or outlier noise. The main contributions are as follows:

(1) In addition to using the non-local self-similarity prior to the image, as with LR-GSC, to maintain the low rank of similar patch coefficients when constructing similar patches, we also impose a low-rank constraint on the reconstructed similar patches as a whole. We not only consider the low-rank property between similar patches, but also consider local structural smoothness. Based on this, we combine the model with the TV norm to further strengthen the local structural smoothness of the matrix. Figure 2 shows an example of image recovery from dense noise based on using and not using the TV norm.



**Figure 2.** Comparison of the denoising results with TV norm and without; 30% of pixels of Barbara image are corrupted by large and sparse noise. (a) The corrupted image. (b) PSNR values for Barbara image corrupted by large and sparse noise with different density. (c) Restored by the model of dual-weighted low-rank group sparse coding without TV norm. PSNR = 30.34 dB. (d) Restored by the model of dual-weighted low-rank group sparse coding with TV norm. PSNR = 31.08 dB.

(2) Because the low rank of the coefficients of similar patches and the global low rank of the reconstruction matrix of similar patches are both considered in the objective function, the optimization problem becomes a challenging non-convex optimization problem. In order to solve this problem effectively, we develop a new numerical solution based on the inexact augmented Lagrange multiplier (IALM) and non-uniform singular value threshold (NSVT). The experimental results show that the proposed method can improve the quality of the restored image significantly.

The remainder of this paper is organized as follows. Section 2 reviews the related work on low-rank regularized group sparse coding (LR-GSC). Section 3 presents our

proposed SDWLR-GSC model, and Section 4 shows the solution of our proposed method. The experimental results are presented in Section 5. Finally, Section 6 summarizes this paper.

## 2. Related Work

In this section, we will review the work related to our proposed method.

### 2.1. Low-Rank Regularized Group Sparse Coding

In GSC [13,14], the method is mainly divided into two steps: grouping similar patches, and then sparsely representing the grouped patches, which can be formulated as the following problem,

$$\hat{A}_i = \arg \min_{A_i} \left( \frac{1}{2} \|X_i - D_i A_i\|_F^2 + \lambda \|A_i\|_1 \right) \quad \forall i \tag{1}$$

where  $X_i \in R^{b \times m}$  is a group composed of patches with similar structure, and  $D_i$  is the dictionary learned from each group of  $X_i$ .  $\|\cdot\|_F$  and  $\|\cdot\|_1$  denote the Frobenius norm and  $L_1$  norm, respectively. Thus, the optimal  $\hat{A}_i$  is the sparse codes of  $X_i$  with  $\lambda \geq 0$ .

Zha et al. [15] proposed LR-GSC, which combines the low-rank characteristics of group coefficients in GSC, to further utilize the similarity between groups; the dictionary domain coefficients of each group are constrained as not only sparse but also low-rank. Taking the sparsity penalty and low-rank penalty as dual regularizers, the following problem is obtained:

$$\{\hat{A}_i, \hat{B}_i\} = \arg \min_{A_i, B_i} \frac{1}{2} \|X_i - D_i A_i\|_F^2 + \lambda \|A_i\|_1 + \frac{1}{\eta} \|A_i - B_i\|_F^2 + \tau \|B_i\|_* \quad \forall i, \tag{2}$$

where  $\|\cdot\|_1$  is applied for the sparsity penalty, and the nuclear norm  $\|\cdot\|_*$  is applied for the low-rank penalty.  $\tau$  is a non-negative constant, and  $\eta$  is a balancing factor in making (2) more feasible. A low-rank approximation  $B_i$  is jointly estimated for each group sparse matrix  $A_i$ . Similar to GSC, the optimal sparse codes  $\{\hat{A}_i\}_{i=1}^n$  are used to restore the latent image.

### 2.2. Dual-Weighted Low-Rank Matrix Recovery

The authors in [22] proved that the RPCA model can be transformed into solving the following convex optimization problem by minimizing the combined decomposition matrix of the  $L_1$  norm and nuclear norm, i.e.,

$$\begin{aligned} \min_{X,S} & \|X\|_* + \lambda \|S\|_1, \\ \text{s.t.} & X + S = Y, \end{aligned} \tag{3}$$

where  $Y \in M^{m \times n}$  is the observed matrix.  $\|X\|_*$  is the nuclear norm of matrix  $X$ , which is defined as the sum of matrix singular values, i.e.,  $\|X\|_* = \sum_{i=1}^r \sigma_i(X)$ ,  $r = \min(m, n)$ ,  $\sigma_i(X)$  is the  $i$ -th singular value of matrix  $X$ .  $\|S\|_1$  is the  $L_1$  norm of matrix  $S$ , which is defined as the sum of matrix absolute elements—that is,  $\|S\|_1 = \sum_{i=1}^m \sum_{j=1}^n |s_{i,j}|$ ,  $s_{i,j}$  represents the elements in matrix  $S$ .  $\lambda > 0$  is the regularization parameter.

In the process of solving, the model can easily lead to overshrinkage, which affects the accuracy of the solution. In [21], the sparse term and low-rank term in the RPCA model are weighted at the same time, and a double-weighted model is proposed, i.e.,

$$\begin{aligned} \min_{X,S} & \|X\|_{\Omega,*} + \lambda \|W \odot S\|_1, \\ \text{s.t.} & X + S = Y, \end{aligned} \tag{4}$$

where  $\|X\|_{\Omega,*} = \sum_{i=1}^r w_i \sigma_i(X)$ ,  $\|W \odot S\|_1 = \sum_{i=1}^m \sum_{j=1}^n w_{i,j} |s_{i,j}|$ ,  $s_{i,j}$  represents the elements in matrix  $S$ ,  $W \in R^{m \times n}$  is the weight matrix. The weight of  $s_{i,j}$  is defined as  $w_{i,j} = 1/(|s_{i,j}| + \varepsilon)$ .

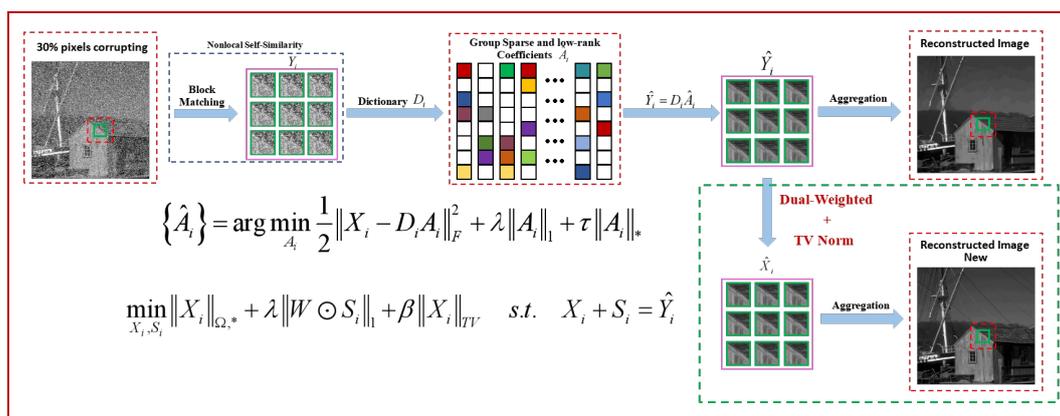
### 3. Formulation of the Proposed Method

Low-rank regularized group sparse coding (LR-GSC) [15] ensures that the dictionary domain coefficients of each group are not only sparse but also low-rank. However, when the density of outlier noise becomes higher, the method cannot recover the image well. In this section, inspired by the low-rank matrix restoration algorithm in [21,23], we propose a new group sparse coding model, called SDWLR-GSC, which introduces total variation (TV) regularization into low-rank group sparse coding to realize image structure smoothing. At the same time, the dual-weighted model is used to recover the clean image better from the large and sparse noise. Our new method can be formulated as follows:

$$\begin{aligned} \min_{X_i, S_i, A_i} & \sum_{j=1}^q w_{X_i, j} \cdot \tilde{\sigma}_j + \theta \|W_{S_i} \odot S_i\|_1 + \beta \|X_i\|_{TV} \\ & + \frac{1}{2\rho} \sum_{i=1}^n \|Y_i - D_i A_i\|_F^2 + \lambda \sum_{i=1}^n \|A_i\|_1 + \tau \sum_{i=1}^n \|A_i\|_* \\ \text{s.t.} & \quad X_i \in B_{l, \mu} \equiv \{x_{i, j}, l \leq x_{i, j} \leq u\}, \\ & \quad \hat{Y}_i = X_i + S_i. \end{aligned} \tag{5}$$

where  $Y_i \in R^{b \times m}$  represents the matrix constructed by the batches of similar patches for each basic patch  $y_i$ .  $\hat{Y}_i$  is calculated by multiplying  $\hat{D}_i$  and  $\hat{A}_i$ .  $\hat{D}_i$  and  $\hat{A}_i$  are the dictionary learned from each group  $Y_i$  and the sparse representation of each group  $X_i$  under a given dictionary, respectively.  $W_{X_i} = \{w_{X_i, j}\}$  and  $W_{S_i} \in R^{m \times n}$  are weights of  $\{\tilde{\sigma}_j\}$  and  $S_i$ .  $\{\tilde{\sigma}_j\}$  are singular values of matrix  $X_i$ . The constraint  $B_{l, \mu}$  states that pixel values are bounded—for example,  $[0, 255]$ .  $\|\cdot\|_{TV}$  denotes the TV norm. In [24], total variation is proposed as a penalty term to deal with image restoration. The advantage of TV regularization is that it can well restore the edge and eliminate the noise.

Apparently, the proposed SDWLR-GSC, based on LR-GSC, exploits the sparsity and low-rankness of the dictionary domain coefficients of each group at the same time. For the image patches reconstructed by the given dictionary and the sparse coding, we further introduce the dual-weighted model to weight low-rank terms and sparse terms, and we add the TV norm to maintain the smoothing structure of images. Therefore, the proposed model can better separate the clean image from the sparse noise and achieve better reconstruction results. Figure 3 shows the flowchart of the proposed SDWLR-GSC by taking the simple image denoising as an example.



**Figure 3.** Flowchart of the proposed SDWLR-GSC model for image denoising. On the basis of LR-GSC [15], we applied the dual-weighted model to the reconstructed similar patches, and introduced the TV norm to maintain the smoothness of the image structure. Finally, the optimal sparse codes are used to estimate the clean patch groups for constructing the restored image. (Firstly, the corrupted image is extracted from non-local similar patches through a block matching operator. Secondly, patches with similar structures are grouped to perform dictionary learning to obtain group coefficients. At the same time, the group coefficients remain sparse and low-rank concurrently.)

### 4. The Proposed Solution

In this section, to solve the proposed SDWLR-GSC problem, we will present our solution method based on the alternating direction method of multipliers, which separates the multi-variable optimal problem into several single-variable problems, so that these variables can be updated one by one in each iteration.

First, to address the problem more easily, we introduce an auxiliary variable as follows:

$$\begin{aligned} \min_{H_i, X_i, S_i, A_i, B_i} & \sum_{j=1}^q w_{H_i, j} \cdot \sigma_j + \theta \|W_{S_i} \odot S_i\|_1 + \beta \|X_i\|_{TV} \\ & + \frac{1}{2\rho} \sum_{i=1}^n \|Y_i - D_i A_i\|_F^2 + \lambda \sum_{i=1}^n \|A_i\|_1 + \tau \sum_{i=1}^n \|B_i\|_* \\ \text{s.t.} & X_i \in B_{l, u} \equiv \{x_{i, j}, l \leq x_{i, j} \leq u\}, \\ & \hat{Y}_i = X_i + S_i, H_i = X_i, B_i = A_i. \end{aligned} \tag{6}$$

where  $W_{H_i} = \{w_{H_i, j}\}$  are weights for  $\{\tilde{\sigma}_j\}$ ,  $\{\tilde{\sigma}_j\}$  are singular values of matrix  $H_i$ , and  $w_{H_i, j} = w_{X_i, j}, j = 1, \dots, n$ .

In this way, the augmented Lagrange function of (7) can be obtained as follows:

$$\begin{aligned} & f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2) \\ = & \sum_{j=1}^n w_{H_i, j} \cdot \sigma_j + \theta \|W_{S_i} \odot S_i\|_1 + \beta \|X_i\|_{TV} \\ & + \frac{1}{2\rho} \sum_{i=1}^m \|Y_i - D_i A_i\|_F^2 + \lambda \sum_{i=1}^n \|A_i\|_1 \\ & + \frac{1}{2\eta} \sum_{i=1}^m \|A_i - B_i\|_F^2 + \tau \sum_{i=1}^n \|B_i\|_* \\ & + \langle Y_1, Y_i - H_i - S_i \rangle + \langle Y_2, X_i - H_i \rangle \\ & + \frac{\mu}{2} (\|Y_i - H_i - S_i\|_F^2 + \|X_i - H_i\|_F^2) \\ \text{s.t.} & X_i \in B_{l, u} \equiv \{x_{i, j}, l \leq x_{i, j} \leq u\}. \end{aligned} \tag{7}$$

Then, we use the iterative alternating direction method, which optimizes one variable and fixes the remaining optimization variables in an iterative way. In this way, the original complex multi-variable optimization problem can be simplified to a single-variable optimization problem, and its solution can be obtained analytically. In our problem, there are seven sub-problems, i.e.,  $H_i, X_i, S_i, A_i, B_i, Y_1, Y_2$  sub-problems. Next, we will give a detailed implementation for each of them.

#### 4.1. $H_i$ Sub-Problem

If we fix the variables  $(X_i, S_i, A_i, B_i, Y_1, Y_2)$ , the variable  $H_i$  can be solved by minimizing  $f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2)$ . Specifically,

$$\begin{aligned} & \arg \min_{H_i} f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2) \\ & = \arg \min_{H_i} \sum_{j=1}^n w_{H_i, j} \cdot \sigma_j \\ & \quad + \langle Y_1, Y_i - H_i - S_i \rangle + \langle Y_2, X_i - H_i \rangle \\ & \quad + \frac{\mu}{2} (\|Y_i - H_i - S_i\|_F^2 + \|X_i - H_i\|_F^2) \\ & = \arg \min_{H_i} \sum_{j=1}^n w_{H_i, j} \cdot \sigma_j \\ & \quad + \mu \|H_i - \frac{1}{2}(Y_i + X_i - S_i + Y_1/\mu + Y_2/\mu)\|_F^2 \\ & = \arg \min_{H_i} \sum_{j=1}^n w_{H_i, j} \cdot \sigma_j + \mu \|H_i - L\|_F^2, \end{aligned} \tag{8}$$

where  $L = \frac{1}{2}(Y_i + X_i - S_i + Y_1/\mu + Y_2/\mu)$ . This minimization problem (8) can be solved by non-uniform singular value thresholding (NSVT) [23] as follows:

$$H = D_{(2\mu)^{-1}W_H}(L). \tag{9}$$

where  $D_\epsilon(Q) = US_\epsilon(\Sigma)V^T$ ,  $U\Sigma V^T$  is the singular value decomposition of matrix  $Q$ .  $S_\epsilon(\cdot)$  denotes non-uniform soft thresholding operator [25], and its  $(i, j)$ -th element is  $\max(|q_{ij}| - \epsilon, 0)\text{sgn}(q_{ij})$ , where the parameter  $\epsilon > 0$ .

4.2.  $X_i$  Sub-Problem

If the variables  $(H_i, S_i, A_i, B_i, Y_1, Y_2)$  are fixed, we can optimize  $X_i$  by minimizing  $f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2)$ . Specifically,

$$\begin{aligned} & \arg \min_{X_i} f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2) \\ &= \arg \min_{X_i} \beta \|X_i\|_{TV} + \langle Y_2, X_i - H_i \rangle + \frac{\mu}{2} \|X_i - H_i\|_F^2 \\ &= \arg \min_{X_i} \beta \|X_i\|_{TV} + \frac{\mu}{2} \|X_i - (H_i - Y_2/\mu)\|_F^2 \\ &= \arg \min_{X_i} \beta \|X_i\|_{TV} + \frac{\mu}{2} \|X_i - R\|_F^2, \end{aligned} \tag{10}$$

where  $R = (H_i - Y_2/\mu)$ ,  $X_i \in B_{l,u} \equiv \{x_{i,j}, l \leq x_{i,j} \leq u\}$ . Similar to the constrained convex problem represented by the image denoising problem using the TV norm in [23], if  $\mu$  and  $\beta$  are given,  $\gamma$  can be obtained by  $\beta/\mu$ . Then, the solution of problem (10) is given by:

$$X_i = P_{B_{l,u}}(R - \gamma L(p, q)). \tag{11}$$

where  $(p, q)$ ,  $L, P$  are the matrix pairs, linear operator, and orthogonal projection operator, respectively. In addition, the bounds of the constraint are set to be  $[l, u] = [0, 255]$ .

4.3.  $S_i$  Sub-Problem

Similar to other variables, if we fix the variables  $(H_i, X_i, A_i, B_i, Y_1, Y_2)$ , the variable  $S_i$  can be updated by the optimal solution of minimizing  $f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2)$ .

$$\begin{aligned} & \arg \min_{S_i} f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2) \\ &= \arg \min_{S_i} \theta \|W_{S_i} \odot S_i\|_1 + \langle Y_1, Y_i - H_i - S_i \rangle \\ & \quad + \frac{\mu}{2} \|Y_i - H_i - S_i\|_F^2 \\ &= \arg \min_{S_i} \theta \|W_{S_i} \odot S_i\|_1 \\ & \quad + \frac{\mu}{2} \|S_i - (Y_i - H_i + Y_2/\mu)\|_F^2, \end{aligned} \tag{12}$$

This minimization problem (12) can be solved by non-uniform soft thresholding (NST) [25] as follows:

$$S = S_{\theta\mu^{-1}W_S}(Y - H + Y_1/\mu). \tag{13}$$

4.4.  $A_i$  Sub-Problem

If we fix the variables  $(H_i, S_i, X_i, B_i, Y_1, Y_2)$ , the  $A_i$  sub-problem can be reduced to the following optimal problem.

$$\begin{aligned} & \arg \min_{A_i} f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2) \\ &= \frac{1}{2p} \sum_{i=1}^n \|Y_i - D_i A_i\|_F^2 \\ & \quad + \frac{1}{2\eta} \sum_{i=1}^n \|A_i - B_i\|_F^2 + \lambda \sum_{i=1}^n \|A_i\|_1 \end{aligned} \tag{14}$$

We use the principal component analysis (PCA) of each group to learn the grouping sub-dictionary  $D_i$  [26]. Since the learned PCA dictionary is singular, Equation (14) is equivalent to the following problem.

$$\begin{aligned} & \arg \min_{A_i} f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2) \\ &= \arg \min_{A_i} \frac{1}{2\rho} \sum_{i=1}^n \|G_i - A_i\|_F^2 \\ &+ \frac{1}{2\eta} \sum_{i=1}^n \|A_i - B_i\|_F^2 + \lambda \sum_{i=1}^n \|A_i\|_1 \\ &= \arg \min_{A_i} \frac{1}{2} \sum_{i=1}^n \|P_i - A_i\|_F^2 + v \sum_{i=1}^n \|A_i\|_1, \end{aligned} \quad (15)$$

where  $Y_i \triangleq D_i G_i$ ,  $P_i \triangleq \frac{\eta G_i + \rho B_i}{\eta + \rho}$  and  $v = \rho \eta \lambda$ . Then, this minimization problem (15) can be solved by non-uniform soft thresholding (NST) [25] as follows:

$$S = S_v(P_i) \quad \forall i. \quad (16)$$

#### 4.5. $B_i$ Sub-Problem

The same as the above methods, if we want to find the updating formula of  $B_i$ , the variables  $(H_i, S_i, X_i, A_i, Y_1, Y_2)$  should also be fixed firstly, and then the solution of the variable  $B_i$  can be obtained by minimizing  $f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2)$ . Specifically,

$$\begin{aligned} & \arg \min_{B_i} f(H_i, X_i, S_i, A_i, B_i, Y_1, Y_2) \\ &= \arg \min_{B_i} \frac{1}{2\eta} \|A_i - B_i\|_F^2 + \tau \|B_i\|_* \end{aligned} \quad (17)$$

This minimization problem (17) can be solved by non-uniform singular value thresholding (NSVT) [23] as follows:

$$B_i = D_{\eta\tau}(A_i). \quad (18)$$

#### 4.6. $Y_1$ and $Y_2$ Sub-Problems

Finally,  $Y_1$  and  $Y_2$  are Lagrange multiplier matrices of the original optimization problem. They should be updated after other variables. If  $Y_1$  is unknown and other variables are fixed,  $Y_1$  can be updated as follows:

$$Y_1 = Y_1 + \mu(Y - H - S). \quad (19)$$

If  $Y_2$  is unknown and other variables are fixed,  $Y_2$  can be updated as follows:

$$Y_2 = Y_2 + \mu(X - H). \quad (20)$$

In Algorithm 1, we summarize the complete algorithm of SDWLR-GSC for image restoration.

---

**Algorithm 1** Smoothed and Dual-Weighted Low-Rank Group Sparse Coding (SDWLR-GSC).

---

**Require:** The degraded image  $y \in R^{m \times n}$  (assuming  $m \geq n$ );

**Output:** The final restored image  $\hat{X}^{(K)}$ ;

- 1: **Initialization:**  $\hat{x}^{(0)} = y, A_i^{(0)} = 0, B_i^{(0)} = 0$ ;
  - 2: Set parameters  $b, m, \rho, \eta, \lambda, \tau$  and Max-Iter;
  - 3: **for**  $k = 1$  **to** Max-Iter **do**
  - 4:   Iterative regularization  $y^{(k)} = \hat{x}^{(k-1)} + \delta(y - \hat{y}^{(k)})$
  - 5:   Divide  $y^{(k)}$  into a set of overlapping patches with size  $b \times b$
  - 6:   **for** each patch  $y_i$  in  $y^{(k)}$  **do**
  - 7:     Find non-local similar patches to form a group  $Y_i$
  - 8:     Construct dictionary  $D_i$  by  $Y_i$  using PCA
  - 9:     Update  $A_i$  by computing  $A_i = D_i^T Y_i$
  - 10:    Perform  $[U_i, \Delta_i, V_i] = SVD(A_i)$
  - 11:    Estimate  $\hat{B}_i$  by computing Equation (18).
  - 12:    Estimate  $\hat{A}_i$  by computing Equation (16).
  - 13:    Update  $\hat{Y}_i$  by computing  $\hat{Y}_i = D_i A_i$
  - 14:    Use **Algorithm 2** to  $\hat{Y}_i$  to estimate  $X_i$
  - 15:   **end for**
  - 16:   Aggregate  $X_i$  to form the clean image  $\hat{X}^{(k)}$
  - 17: **end for**
- 

**Algorithm 2** Smoothed and Dual-Weighted Model for Image Denoising.

---

**Require:** non-local similar patched group  $Y_i \in R^{p \times q}$  (assuming  $p \geq q$ );

**Output:** solutions  $X_i^{(k)} = X_i^{(t+1)}$ ;

- 1: **Initialization:**  $W_{X_i}^{(0)} = \mathbf{1} \in R^q, W_{S_i}^{(0)} = \mathbf{1} \cdot \mathbf{1}^T \in R^{p \times q}, (X_i^{(0)}, S_i^{(0)}) \in R^{p \times q}, H^{(0)} \in R^{p \times q}, Y_1^{(0)} \in R^{p \times q}, Y_2^{(0)} \in R^{p \times q}$ ;
- 2: Set parameters  $\mu_0 > 0, \zeta = 10^{-7}, k = 0, \lambda, \delta$  and inneriter;
- 3: **while**  $\|Y_i - X_i - S_i\|_F / \|Y_i\|_F > \zeta$  and  $t < inneriter$  **do**
- 4:   Let  $L^{(t+1)} = Y_i + X_i^{(t)} - S_i^{(t)} + Y_1^{(t)} / \mu^{(t)}, \rho = \eta / \mu^{(t)}, H_i^{(t+1)} = D_{\mu^{(t)-1} W_{X_i}}(L^{(t+1)})$
- 5:   Let  $R^{(t+1)} = H^{(t+1)} - Y_2^{(t)}$  using the **FGP Algorithm [23]** to compute  $X_i^{(t+1)} = P_{B_i, \mu}(R^{(t+1)} - \rho L(p, q))$
- 6:    $S^{(t+1)} = S_{\lambda \mu^{(t)-1} W_{S_i}}(Y_i - H^{(t+1)} + Y_1^{(t)} / \mu^{(t)})$
- 7:    $Y_1^{(t+1)} = Y_1^{(t)} + \mu^{(t)}(Y_i - H^{(t+1)} - S_i^{(t+1)})$
- 8:    $Y_2^{(t+1)} = Y_2^{(t)}(X_i^{(t+1)} - H^{(t+1)})$
- 9:    $\mu^{(t+1)} = \delta \mu^{(t)}, t \rightarrow t + 1$
- 10: Update weights [27]: The weights for each  $i = 1, \dots, p$  and  $j = 1, \dots, q$  are updated by

$$w_{X_i, j}^{(k+1)} = \frac{1}{\sigma_j^{(k)} + \epsilon_{X_i}}, w_{S_i, ij}^{(k+1)} = \frac{1}{|S_{ij}^{(k)}| + \epsilon_{S_i}}.$$

where  $\epsilon_{X_i}$  and  $\epsilon_{S_i}$  are predetermined positive constants, and the singular value matrix

$$\Sigma^{(k)} = \text{diag}\left(\left[\sigma_1^{(k)}, \dots, \sigma_n^{(k)}\right]\right) \in R^{n \times n}$$

with  $[U^{(k)}, \Sigma^{(k)}, V^{(k)}] = \text{svd}(X_i^{(k)})$

11: **end while**

---

## 5. Results

To verify the performance of the SDWLR-GSC algorithm in image restoration, we carried out a large number of experiments and compared its performance with the most novel methods based on the low-rank matrix. We also selected 16 classical images (Lena, Barbara, Couple, House, Monarch are from the Set12 dataset. Frame, Road, Bridge, Elaine, Pentagon, Lin are from Google Images. Flower is from the Kodak24 dataset. Monkey, Tank are from the USC-SIPI image dataset <http://sipi.usc.edu/database/> (accessed on 28 March 2022)) with size  $512 \times 512$  as the test dataset. The images used in all experiments are shown in Figure 4. This experiment mainly aimed at outlier noise with high density. Outlier noise includes two types: salt and pepper noise and random-valued sparse additive noise. Thus, we introduced two classes of numerical experiments. First, the test image was only destroyed by different levels of large and sparse additive noise. Second, the test image was only destroyed by different levels of salt and pepper noise. The peak signal to noise ratio (PSNR) [28] and structural similarity index metric (SSIM) [27] were used as quality evaluation indicators to evaluate performance. The larger the value of PSNR and SSIM, the better the image quality. All experiments were carried out on MATLAB R2020a running on Windows 10. The CPU was Intel Core i7-2600, the memory was 8 GB, and the frequency was 3.40 GHz.



**Figure 4.** Test images in our experiments. First row: Lena, Barbara, Goldhill, Frame, Road, Bridge, Couple, Monkey. Second row: Boat, Elaine, Flower, House, Lin, Monarch, Pentagon, Tank.

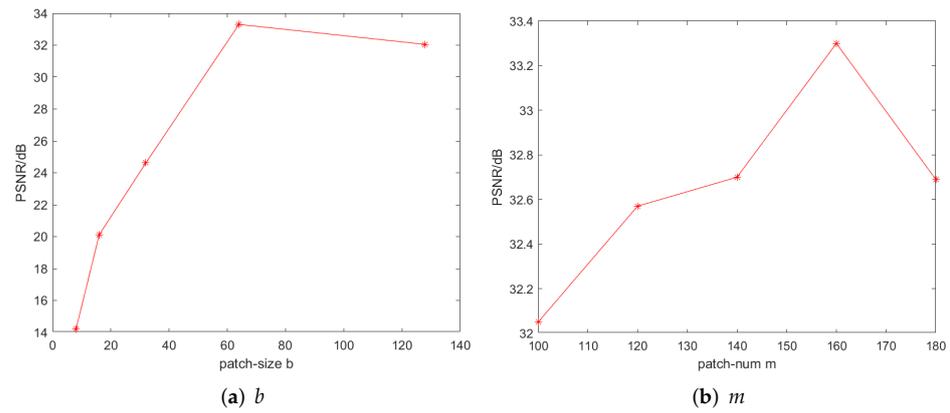
### 5.1. Parameter Setting

The parameter settings involved in this algorithm are as follows. The size of the input matrix is  $m \times n$ . In Algorithm 1, we set  $A_i^{(0)} = 0$ ,  $B_i^{(0)} = 0$ . In Algorithm 2, we set  $X_i^{(0)} = 0$ ,  $S_i^{(0)} = 0$ ,  $H_i^{(0)} = 0$ . Following the practice in [29] and our tests, we set the Lagrange multiplier matrices  $Y_1^{(0)} = M / \max(\|M\|, \lambda^{-1}\|M\|_\infty)$ , and  $Y_2^{(0)} = 0$ , where  $\|M\|$  is the spectral norm of matrix  $M$  and  $\|M\|_\infty$  is the maximum absolute value of the entries in matrix  $M$ . In addition, the constants  $\epsilon_{X_i}$  and  $\epsilon_{S_i}$  in Algorithm 2 are set to 0.1, the same as in [21].

In addition, we set  $\theta$  in (13), which controls the sensitivity of the model to coefficient errors to be  $1.25 / \sqrt{\max(m, n)}$ , and we set  $\beta$  in (11), which controls the sensitivity of the model to the TV norm to be  $10^{-8} / \sqrt{\max(m, n)}$ . Finally, in Algorithm 2,  $\gamma$  can be computed by  $\gamma = \eta / \beta$ , and  $\mu$  corresponds to parameter  $\mu^{(t)}$  in Algorithm 2, and we set  $\delta = 0.1$ .

The above parameters can be set according to experience. However, in Algorithm 1, the size of patch  $b$ , the number of non-local similar patches  $m$ , balancing factors  $\rho$  and  $\eta$  in (15), regularization parameters  $\lambda$  in (15) and  $\tau$  in (17), and other parameters are determined by parameter experiments.

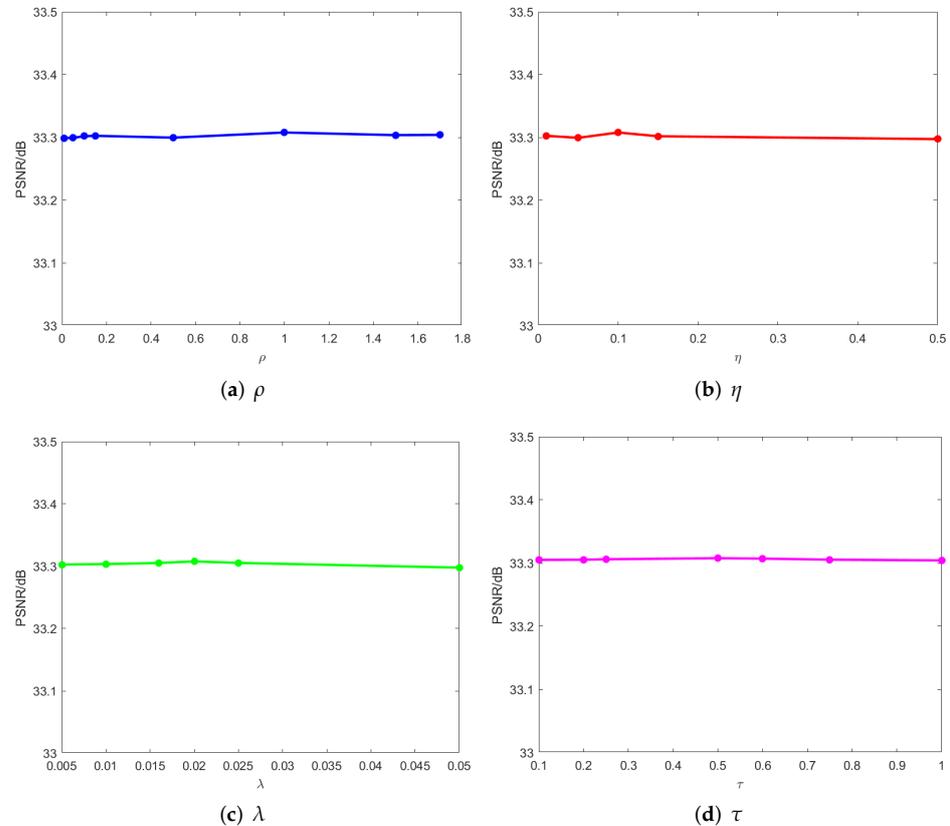
In the parameter experiments of patch size  $b$  and patch number  $m$ , the value range of  $b$  is multiples of 2 from 8 to 64, the value range of  $m$  is 100~180, and the sampling interval is 20. Figure 5 shows the influence of the values of  $b$  and  $m$  on the restoration results when the noise density  $p = 0.3$ . Therefore, according to the experiment, when the noise density is  $p = 0.3$ , we set the size of image blocks to  $64 \times 64$  and look for 160 similar image patches. In addition, through experiments, when  $p = 0.2$ ,  $p = 0.4$ , we set patch sizes to  $60 \times 60$  and  $70 \times 70$ , and the number of non-local similar patches selected is 160 and 190.



**Figure 5.** The influence of the values of  $b$  and  $m$  on the restoration results when the noise density  $p = 0.3$ .

It can be seen that when the noise density increases, the size of the image patch also increases. When the noise density increases, the original similar patches may be different from each other, so expanding the size of similar patches and increasing the number of similar patches can improve the accuracy of the algorithm.

In addition, we analyze the parameter experiments of balancing factors  $\rho$  and  $\eta$ , and regularization parameters  $\lambda$  and  $\tau$ . Figure 6 shows that when the noise density  $p = 0.3$ , the influence of the value of this parameter on the restoration result when other parameters are fixed is noticeable. We set  $\rho = 1, \eta = 0.1, \lambda = 0.02, \tau = 0.5$  when the noise density  $p = 0.3$ .



**Figure 6.** The influence of the values of  $\rho, \eta, \lambda,$  and  $\tau$  on the restoration results when the noise density  $p = 0.3$ .

In Algorithm 1, the maximum number of iterations  $maxiter$ , and in Algorithm 2, the internal maximum number of iterations  $inneriter$ , are parameters related to the convergence of the algorithm. Because SDWLR-GSC is a non-convex model, it is difficult to prove the global convergence of the algorithm in theory. Therefore, this paper analyzes the influence of iteration times on the restoration results through experiments. Through experiments,  $maxiter = 2$  and  $inneriter = 100$  are selected to avoid unnecessary iterative calculation.

### 5.2. Large Sparse Additive Noise Removal

We first test the experimental results of the proposed SDWLR-GSC for removing large and sparse additive noise. We suppose that the ratio of damaged pixels to all pixels in the image is  $p$ , and the values of these noisy pixels are 255, and a noisy observation image is generated. In this test, we give the denoising results of three noise levels, namely  $p = 0.2$ ,  $p = 0.3$ , and  $p = 0.4$ . There are several other parameters in the algorithm, which are set as follows. In Algorithm 1, the number of iterations  $k$  and the patch size are set according to the noise level. For higher noise levels, we choose larger patches and run more iterations. According to experiments, let the ratio of the number of damaged pixels to the number of all pixels be  $p$ , and set patch sizes to  $60 \times 60$ ,  $64 \times 64$ , and  $70 \times 70$  for  $p = 0.2$ ,  $p = 0.3$ , and  $p = 0.4$ , respectively. At these noise levels, the number of non-local similar patches selected is 160, 160, and 190, respectively.

We compare the proposed SDWLR-GSC with the most advanced existing methods, including the PCP algorithm [30], reweighted  $L_1$  algorithm [31], NSVT method [21], and SRLRMR algorithm [29]. We also compare it with the LR-GSC algorithm [15]. It can be seen from Table 1 that under all noise levels, compared with other competitive methods, the SDWLR-GSC proposed by us achieves higher peak signal-to-noise ratio results. It can be seen from Table 1 that under all noise levels, compared with other competitive methods, the SDWLR-GSC proposed by us achieves higher PSNR results. In addition, the LR-GSC is more suitable for removing Gaussian noise than large and sparse noise.

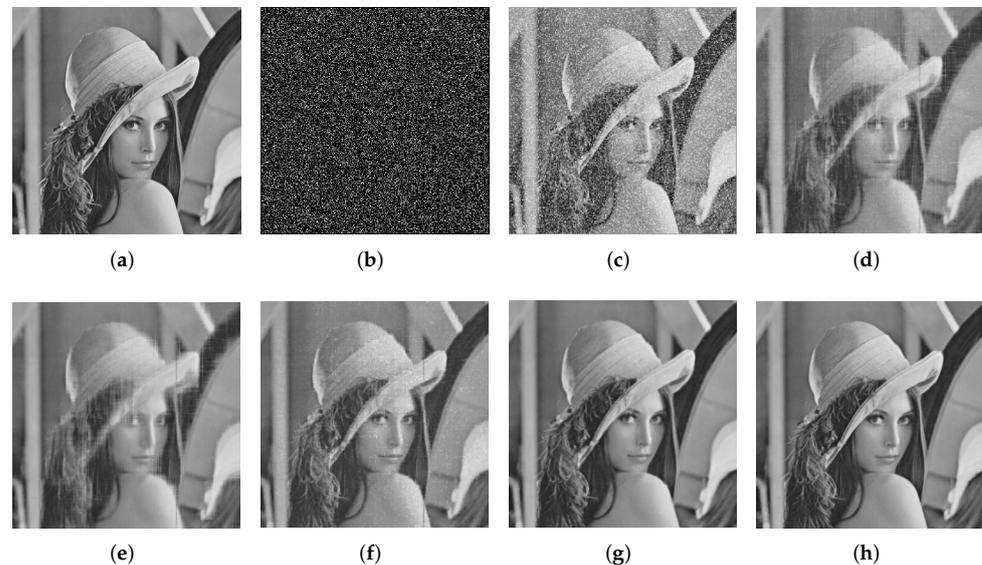
**Table 1.** Comparison of different denoising methods under different large sparse noise probabilities in terms of PSNR.

$p$ Value	Image No.	PSNR (dB)					
		LR-GSC	PCP	Reweight $l_1$	NSVT	SRLRMR	SDWLR-GSC (Ours)
$p = 0.2$	Lena	12.25	26.75	23.84	32.15	32.47	<b>36.76</b>
	Barbara	11.74	24.23	22.33	26.28	26.36	<b>34.56</b>
	Goldhill	11.56	28.41	26.75	31.99	33.71	<b>34.48</b>
	Frame	9.63	25.90	22.79	28.10	29.64	<b>30.95</b>
	Road	10.01	29.77	26.43	34.63	<b>34.88</b>	34.50
	Bridge	10.17	29.98	27.96	34.01	33.73	<b>34.69</b>
	Couple	12.19	25.71	23.49	28.78	29.31	<b>31.18</b>
	Monkey	12.67	21.72	20.43	22.12	23.71	<b>25.48</b>
	Boat	12.86	25.22	23.01	27.65	28.35	<b>32.23</b>
	<b>Average</b>	11.45	26.41	24.11	29.52	30.24	<b>32.76</b>
$p = 0.3$	Lena	10.48	21.27	22.17	23.84	30.67	<b>33.31</b>
	Barbara	9.97	19.10	20.85	20.77	24.35	<b>31.08</b>
	Goldhill	9.80	23.00	25.42	25.56	31.15	<b>31.91</b>
	Frame	7.86	22.52	23.04	26.61	27.36	<b>29.31</b>
	Road	8.24	25.20	26.69	28.73	<b>33.31</b>	33.18
	Bridge	8.41	27.29	26.57	30.57	33.01	<b>33.09</b>
	Couple	10.44	20.92	22.01	22.93	27.60	<b>28.18</b>
	Monkey	10.91	18.09	19.21	18.12	21.90	<b>22.93</b>
	Boat	11.12	21.25	21.62	22.01	27.66	<b>28.74</b>
	<b>Average</b>	9.69	22.07	23.06	24.35	28.56	<b>30.19</b>

Table 1. Cont.

$p$ Value	Image No.	PSNR (dB)					
		LR-GSC	PCP	Reweight $l_1$	NSVT	SRLRMR	SDWLR-GSC (Ours)
$p = 0.4$	Lena	9.22	13.18	16.53	13.57	20.53	<b>27.44</b>
	Barbara	8.71	12.25	15.10	12.50	17.88	<b>24.43</b>
	Goldhill	8.54	13.04	19.39	13.80	21.66	<b>27.48</b>
	Frame	7.61	12.32	19.61	15.69	22.58	<b>26.26</b>
	Road	6.98	11.39	18.65	13.21	20.04	<b>29.87</b>
	Bridge	7.14	13.94	24.49	18.04	28.12	<b>32.15</b>
	Couple	9.18	13.28	17.24	13.50	20.55	<b>23.60</b>
	Monkey	9.65	12.65	15.00	12.06	17.07	<b>18.68</b>
	Boat	9.86	13.82	17.03	13.84	19.64	<b>24.01</b>
	<b>Average</b>	<b>8.54</b>	<b>12.88</b>	<b>18.12</b>	<b>14.02</b>	<b>20.90</b>	<b>25.99</b>

When  $p = 0.4$  noise probability, on average, the PSNR performance of this method is 13.11 dB higher than that of PCP, 7.87 dB higher than that of specific gravity reweighted  $L_1$ , 11.97 dB higher than that of NSVT, and 5.09 dB higher than that of SRLRMR. Figure 7 shows the visual comparison results of an image (Lena) when the noise density is  $p = 0.2$ . It can be observed that the PCP, reweighted  $L_1$ , and NSVT methods cannot completely restore the damaged image, while the SRLRMR method can better complete the image denoising task, but it is slightly lacking in structure and the edge is still blurred. In contrast, our proposed SDWLR-GSC method not only effectively eliminates noise, but also retains sharp edges and fine details.



**Figure 7.** Restoration results comparison of the single image, Lena, of the PCP, reweighted  $l_1$ , NSVT, SRLRMR, and our method. Here, 20% image pixels are corrupted by large and sparse noise. (a) The original Lena image. (b) The large and sparse noise. (c) The corrupted image (12.24 dB). (d) Restoration result by PCP (26.75 dB). (e) Restoration result by reweighted  $l_1$  norm minimization (23.84 dB). (f) Restoration result by NSVT (32.15 dB). (g) Restoration result by SRLRMR (32.47 dB). (h) Restoration result by our method (36.76 dB).

### 5.3. Salt and Pepper Noise Removal

In this subsection, we apply the proposed SDWLR-GSC to remove salt and pepper noise—that is, we add the noise with a pixel value of 0 on the basis of the noise in the previous section, and then the ratio of damaged pixels' total number with pixel values

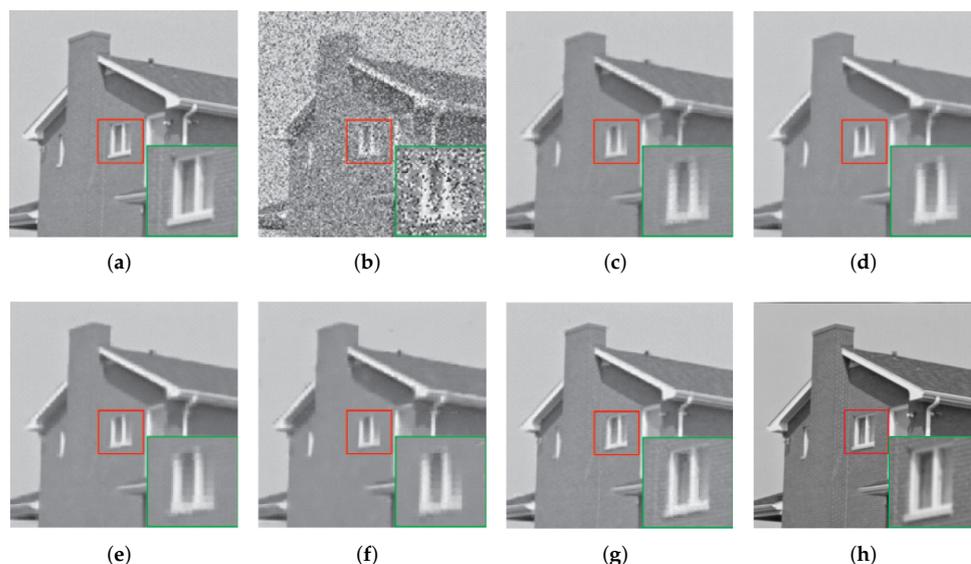
of 0 and 255 to all pixels' number is  $p$ . Similarly, in this test, we still give three different noise levels of  $p = 0.2$ ,  $p = 0.3$ , and  $p = 0.4$ . We compare this model with LR-GSC [15], PCP [30], the WNNM-RPCA model [20], the WSNM-RPCA model [32], WSNM- $L_1$  [33], and the DWLP model [34].

Table 2 quantitatively compares the denoising results of various methods under different salt and pepper noise probability  $p$ . It can be seen that when  $p = 0.2$ ,  $p = 0.3$ ,  $p = 0.4$ , the PSNR of our model is higher than that of LR-GSC, PCP, WNNM-RPCA, WSNM- $L_1$ , and DWLP. Moreover, Figure 8 shows the visual comparison results of an image (House) when the noise density is  $p = 0.3$ .

Through the experiments, we can draw several conclusions: firstly, the strong salt and pepper noise destroys the sparse prior and low-rank of the image, and the restoration performance of the PCP model is poor; secondly, the PSNR of our model is higher than the average level of other models, which shows that using the dual-weighted model in group sparse coding, processing low-rank components and sparse components at the same time, and introducing the TV norm to solve the problem of image structure smoothing can reconstruct the low-rank structure of images more accurately.

**Table 2.** Comparison of different denoising methods under different salt and pepper probabilities in terms of PSNR.

$p$ Value	Image	PSNR (dB)						
		LR-GSC	PCP	WNNM-RPCA	WSNM-RPCA	WSNM- $L_1$	DWLP	SDWLR-GS (Ours)
$p = 0.2$	Couple	12.60	26.34	26.03	26.29	25.08	31.57	<b>31.73</b>
	Elaine	12.58	30.35	30.24	30.48	28.64	36.38	<b>42.21</b>
	Flower	12.64	27.94	27.46	27.67	26.59	31.78	<b>41.30</b>
	Goldhill	12.46	28.64	28.00	28.16	26.79	32.97	<b>34.09</b>
	House	12.56	30.29	31.37	31.43	29.84	37.19	<b>46.21</b>
	Lin	12.01	29.61	29.51	29.54	26.72	33.04	<b>38.52</b>
	Monarch	12.15	24.51	26.38	26.55	25.16	29.53	<b>41.30</b>
	Pentagon	12.72	25.84	25.51	25.66	24.29	31.20	<b>31.83</b>
	Tank	12.91	31.67	30.12	30.23	29.49	<b>35.92</b>	34.57
	<b>Average</b>	12.51	28.35	28.29	28.45	26.96	32.29	<b>37.97</b>
$p = 0.3$	Couple	10.88	24.50	24.41	24.64	23.91	28.77	<b>29.68</b>
	Elaine	10.80	28.42	28.44	28.65	27.15	33.46	<b>38.77</b>
	Flower	10.91	27.11	26.28	26.43	25.38	29.29	<b>36.25</b>
	Goldhill	10.71	27.43	26.56	28.16	25.63	30.30	<b>31.65</b>
	House	10.76	28.88	29.51	29.77	28.15	33.89	<b>38.08</b>
	Lin	10.29	27.22	27.80	27.93	26.18	30.58	<b>36.03</b>
	Monarch	10.40	22.84	24.27	24.47	23.61	26.39	<b>36.82</b>
	Pentagon	10.95	24.15	23.97	24.18	23.16	28.93	<b>29.45</b>
	Tank	11.14	30.05	29.57	29.64	28.57	<b>33.21</b>	32.73
	<b>Average</b>	10.76	26.73	26.76	27.10	25.75	30.54	<b>34.97</b>
$p = 0.4$	Couple	9.60	23.43	23.77	23.81	22.62	26.57	<b>28.35</b>
	Elaine	9.55	25.54	26.60	27.03	25.18	30.71	<b>35.94</b>
	Flower	9.64	25.78	25.68	25.60	24.23	28.03	<b>33.43</b>
	Goldhill	9.44	24.79	25.40	25.92	24.12	28.68	<b>30.22</b>
	House	9.52	24.83	27.90	27.92	25.97	31.04	<b>34.60</b>
	Lin	9.02	25.43	26.11	26.15	23.71	27.23	<b>33.97</b>
	Monarch	9.14	21.34	22.98	23.22	22.22	24.32	<b>34.42</b>
	Pentagon	9.73	22.34	23.04	22.96	21.95	26.13	<b>27.97</b>
	Tank	9.90	27.57	29.12	29.19	27.29	<b>31.94</b>	31.70
	<b>Average</b>	9.50	24.56	25.62	25.76	24.14	28.29	<b>32.29</b>



**Figure 8.** Restoration results comparison of the image House of the PCP, WNNM-RPCA, WSNM-RPCA, WSNM- $l_1$ , DWLP, and our method. Here, 30% image pixels are corrupted by salt and pepper noise. (a) The original Lena image, and the image size is  $512 \times 512$ . (b) The input corrupted image (10.69 dB). (c) Restoration result by PCP (28.88dB). (d) Restoration result by WNNM-RPCA (29.51 dB). (e) Restoration result by WSNM-RPCA (29.77 dB). (f) Restoration result by WSNM-RPCA  $l_1$  norm minimization (28.15dB). (g) Restoration result by DWLP (33.89 dB). (h) Restoration result by our method (38.08 dB).

## 6. Conclusions

In this paper, we constructed a smooth dual-weighted low-rank group sparse coding model. It combined group sparse coding, the TV norm, and a dual-weighted model. We also proved the superior performance of the proposed method in image denoising. Experimental results show that this method is obviously superior to the original PCP optimization, reweighted  $L_1$  norm minimization, and NSVT and SRLRMR algorithms in removing large and sparse additive noise, and it is obviously superior to the PCP optimization, WNNM-RPCA, WSNM-RPCA, WSNM- $L_1$ , and DWLP methods in removing salt and pepper noise.

Although our proposed method has good performance, there is still room for further improvement. In this paper, the algorithm solver is based on the alternating direction method, due to the matrix–matrix multiplications and matrix inversions, and singular value decomposition is required for each iteration, which has high computational complexity for large matrices. In addition, when the expectation matrix becomes complex—for example, it has a high internal rank structure or the deletion becomes dense—satisfactory performance may not be obtained. Therefore, the question of how to reduce the computational complexity while maintaining high performance will be our research direction in the future. In addition, other applications of the proposed method are another important issue for future work.

**Author Contributions:** Conceptualization, J.W. and H.W.; methodology, J.W. and H.W.; software, J.W.; validation, J.W. and H.W.; formal analysis, H.W.; investigation, H.W.; resources, H.W.; data curation, J.W.; writing—original draft preparation, J.W.; writing—review and editing, J.W. and H.W.; visualization, J.W.; supervision, H.W.; project administration, H.W.; funding acquisition, H.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China, grant number 41971396, 61971290; the Research Ability Enhancement Program for Young Teachers of Beijing University of Civil Engineering and Architecture, grant number X21024; the Outstanding Youth Program of Beijing University of Civil Engineering and Architecture, and the Beijing Municipal Education Commission Science and Technology General Project, grant number KM202110016001.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This work was supported in part by the National Natural Science Foundation of China (Nos. 41971396, 61971290), the Research Ability Enhancement Program for Young Teachers of Beijing University of Civil Engineering and Architecture (No. X21024), the Outstanding Youth Program of Beijing University of Civil Engineering and Architecture, and the Beijing Municipal Education Commission Science and Technology General Project (No. KM202110016001).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Elad, M.; Elad, M. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*; Springer: New York, NY, USA, 2010.
2. Wright, J.; Yi, M.; Mairal, J.; Sapiro, G.; Huang, T.S.; Yan, S. Sparse representation for computer vision and pattern recognition. *Proc. IEEE* **2010**, *98*, 1031–1044. [[CrossRef](#)]
3. Elad, M.; Figueiredo, M.A.; Ma, Y. On the role of sparse and redundant representations in image processing. *Proc. IEEE* **2010**, *98*, 972–982. [[CrossRef](#)]
4. Xu, J.; Lei, Z.; Zuo, W.; Zhang, D. Patch group based nonlocal self-similarity prior learning for image denoising. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 244–252.
5. Jolliffe, I.T. Principal component analysis. *J. Mark. Res.* **2002**, *87*, 513.
6. Aharon, M.; Elad, M.; Bruckstein, A. *rmK-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation*. *IEEE Trans. Signal Process.* **2006**, *54*, 4311–4322. [[CrossRef](#)]
7. Rubinstein, R.; Zibulevsky, M.; Elad, M. Double sparsity: Learning sparse dictionaries for sparse signal approximation. *IEEE Trans. Signal Process.* **2010**, *58*, 1553–1564. [[CrossRef](#)]
8. Mairal, J.; Bach, F.; Ponce, J.; Sapiro, G.; Zisserman, A. Non-local sparse models for image restoration. In Proceedings of the 2009 IEEE 12th International Conference on Computer Vision (ICCV), Kyoto, Japan, 27 September–4 October 2009.
9. Zhang, J.; Zhao, D.; Gao, W. Group-based sparse representation for image restoration. *IEEE Trans. Image Process. Publ. IEEE Signal Process. Soc.* **2014**, *23*, 3336. [[CrossRef](#)]
10. Dong, W.; Shi, G.; Ma, Y.; Li, X. Image restoration via simultaneous sparse coding: Where structured sparsity meets gaussian scale mixture. *Int. J. Comput. Vis.* **2015**, *114*, 217–232. [[CrossRef](#)]
11. Wang, Q.; Zhang, X.; Wu, Y.; Tang, L.; Zha, Z. Non-convex weighted  $l_p$  minimization based group sparse representation framework for image denoising. *IEEE Signal Process. Lett.* **2017**, *24*, 1686–1690. [[CrossRef](#)]
12. Liu, J.; Yang, W.; Zhang, X.; Guo, Z. Retrieval compensated group structured sparsity for image super-resolution. *IEEE Trans. Multimed.* **2016**, *19*, 302–316. [[CrossRef](#)]
13. Yu, H.; Gao, L.; Liao, W.; Zhang, B.; Zhuang, L.; Song, M.; Chanussot, J. Global spatial and local spectral similarity-based manifold learning group sparse representation for hyperspectral imagery classification. *IEEE Trans. Geosci. Remote. Sens.* **2020**, *58*, 3043–3056. [[CrossRef](#)]
14. Xu, J.; Qiao, Y.; Fu, Z.; Wen, Q. Image block compressive sensing reconstruction via group-based sparse representation and nonlocal total variation. *Circuits Syst. Signal Process.* **2019**, *38*, 304–328. [[CrossRef](#)]
15. Zha, Z.; Wen, B.; Yuan, X.; Zhou, J.; Zhu, C. Reconciliation Of Group Sparsity And Low-Rank Models For Image Restoration. In Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME), London, UK, 6–10 July 2020.
16. Wright, J.; Ganesh, A.; Rao, S.; Ma, Y. Robust principal component analysis: Exact recovery of corrupted low-rank matrices. In *NIPS'09: Proceedings of the 22nd International Conference on Neural Information Processing Systems Red Hook, NY, USA, 7–9 December 2009*; Curran Associates Inc.: Red Hook, NY, USA, 2009; pp. 2080–2088.
17. Candès, E.J.; Plan, Y. Matrix completion with noise. *Proc. IEEE* **2010**, *98*, 925–936.
18. Candès, E.J.; Wakin, M.B.; Boyd, S.P. Enhancing sparsity by reweighted  $l_1$  minimization. *J. Fourier Anal. Appl.* **2008**, *14*, 877–905.
19. Gu, S.; Lei, Z.; Zuo, W.; Feng, X. Weighted Nuclear Norm Minimization with Application to Image Denoising. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; pp. 2862–2869.
20. Gu, S.; Xie, Q.; Meng, D.; Zuo, W.; Feng, X.; Zhang, L. Weighted nuclear norm minimization and its applications to low level vision. *Int. J. Comput. Vis.* **2017**, *121*, 183–208. [[CrossRef](#)]
21. Peng, Y.; Suo, J.; Dai, Q.; Xu, W. Reweighted low-rank matrix recovery and its application in image restoration. *IEEE Trans. Cybern.* **2014**, *44*, 2418–2430. [[CrossRef](#)]
22. Candès, E.J.; Li, X.; Ma, Y.; Wright, J.A. Robust principal component analysis? *J. ACM* **2011**, *58*, 1–37. [[CrossRef](#)]
23. Beck, A.; Teboulle, M. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. Image Process.* **2009**, *18*, 2419–2434. [[CrossRef](#)]

24. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D Nonlinear Phenom.* **1992**, *60*, 259–268. [[CrossRef](#)]
25. Donoho, D.L. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory* **2002**, *41*, 613–627. [[CrossRef](#)]
26. Dong, W.; Lei, Z.; Shi, G. Nonlocally centralized sparse representation for image restoration. *IEEE Trans. Image Process. Publ. IEEE Signal Process. Soc.* **2013**, *22*, 1620–1630. [[CrossRef](#)]
27. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2014**, *13*, 600–612. [[CrossRef](#)]
28. Ni, Z.; Shi, Y.Q.; Ansari, N.; Wei, S. Reversible data hiding. *IEEE Trans. Circuits Syst. Video Technol.* **2006**, *16*, 354–362.
29. Wang, H.; Cen, Y.; He, Z.; He, Z.; Zhao, R.; Zhang, F. Reweighted low-rank matrix analysis with structural smoothness for image denoising. *IEEE Trans. Image Process.* **2018**, *27*, 1777–1792. [[CrossRef](#)] [[PubMed](#)]
30. Lin, Z.; Chen, M.; Ma, Y. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *arXiv* **2010**, arXiv:1009.5055.
31. Yue, D.; Dai, Q.; Liu, R.; Zhang, Z.; Hu, S. Low-rank structure learning via nonconvex heuristic recovery. *IEEE Trans. Neural Netw. Learn. Syst.* **2013**, *24*, 383–396.
32. Xie, Y.; Qu, Y.; Tao, D.; Wu, W.; Yuan, Q.; Zhang, W. Hyperspectral image restoration via iteratively regularized weighted Schatten  $p$ -norm minimization. *IEEE Trans. Geosci. Remote. Sens.* **2016**, *54*, 4642–4659. [[CrossRef](#)]
33. Chen, G.; Wang, J.; Zhang, F.; Wang, W. Image denoising in impulsive noise via weighted Schatten  $p$ -norm regularization. *IEEE Trans. Image Process.* **2019**, *28*, 1. [[CrossRef](#)]
34. Dong, H.; Yu, J.; Xiao, C. Dual reweighted lp-norm minimization for salt-and-pepper noise removal. *arXiv* **2018**, arXiv:1811.09173.