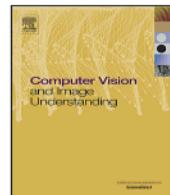


## ARTICLE IN PRESS



Contents lists available at ScienceDirect



## Computer Vision and Image Understanding

journal homepage: [www.elsevier.com/locate/cviu](http://www.elsevier.com/locate/cviu)

## Highlights

**Simultaneous image denoising and completion through convolutional sparse representation and nonlocal self-similarity***Computer Vision and Image Understanding xxx (xxxx) xxx*

Weimin Yuan, Yuanyuan Wang, Ruirui Fan, Yuxuan Zhang, Guangmei Wei, Cai Meng\*, Xiangzhi Bai

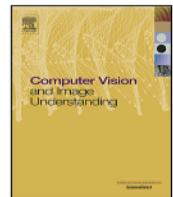
- Focusing on the task of simultaneous image denoising and completion, we propose a novel Low rank matrix approximation (LRMA) based method. The proposed method full leverages convolutional analysis sparse representation (ASR) and nonlocal statistical modeling (NLSM).
- By exploiting the alternating direction method of multipliers (ADMM), the proposed method, which embeds three regularization terms into a unified framework, is solved efficiently.
- Extensive experimental results demonstrate the superiority and generalizability of the proposed method by comparing it with seventeen state-of-the-art methods on four widely-used benchmarks and several challenging degradation scenarios.

**Graphical abstract and Research highlights will be displayed in online search result lists, the online contents list and the online article, but will not appear in the article PDF file or print unless it is mentioned in the journal specific style requirement. They are displayed in the proof pdf for review purpose only.**



Contents lists available at ScienceDirect

## Computer Vision and Image Understanding

journal homepage: [www.elsevier.com/locate/cviu](http://www.elsevier.com/locate/cviu)

# Simultaneous image denoising and completion through convolutional sparse representation and nonlocal self-similarity

Weimin Yuan<sup>a</sup>, Yuanyuan Wang<sup>a</sup>, Ruirui Fan<sup>c</sup>, Yuxuan Zhang<sup>a</sup>, Guangmei Wei<sup>c</sup>, Cai Meng<sup>a,b,\*</sup>, Xiangzhi Bai<sup>a,b</sup>

<sup>a</sup> Image Processing Center, Beihang University, Beijing, 100191, China

<sup>b</sup> Beijing Advanced Innovation Center for Biomedical Engineering, Beihang University, Beijing, 100191, China

<sup>c</sup> School of Mathematics and Systems Science, Beihang University, Beijing, 100191, China

## ARTICLE INFO

Communicated by Si Liu

**Keywords:**

Simultaneous image denoising and completion  
Low rank matrix approximation  
Convolutional analysis sparse representation

## ABSTRACT

Low rank matrix approximation (LRMA) has been widely studied due to its capability of approximating original image from the degraded image. According to the characteristics of degraded images, image denoising and image completion have become research objects. Existing methods are usually designed for a single task. In this paper, focusing on the task of simultaneous image denoising and completion, we propose a weighted low rank sparse representation model and the corresponding efficient algorithm based on LRMA. The proposed method integrates convolutional analysis sparse representation (ASR) and nonlocal statistical modeling to maintain local smoothness and nonlocal self-similarity (NLSM) of natural images. More importantly, we explore the alternating direction method of multipliers (ADMM) to solve the above inverse problem efficiently due to the complexity of simultaneous image denoising and completion. We conduct experiments on image completion for partial random samples and mask removal with different noise levels. Extensive experiments on four datasets, i.e., Set12, Kodak, McMaster, and BSD68, show that the proposed method prevents the transmission of noise while completing images and has achieved better quantitative results and human visual quality compared to 17 methods. The proposed method achieves (1.9%, 1.8%, 4.2%, and 3.7%) gains in average PSNR and (4.2%, 2.9%, 6.7%, and 6.6%) gains in average SSIM over the sub-optimal method across the four datasets, respectively. We also demonstrate that our method can handle the challenging scenarios well. Source code is available at [https://github.com/weimin581/demo\\_CSRNS](https://github.com/weimin581/demo_CSRNS).

## 1. Introduction

Image completion aims to recover missing entries from degraded images. It has evolved from the original restoration of old pictures to photo editing and film post-production with the development of digital image processing (Jam et al., 2021; Qin et al., 2021; Bhavsar and Rajagopalan, 2012; He and Wang, 2014; Yuan et al., 2021; Zhang et al., 2014b,c; Zhou et al., 2012; Nie et al., 2019; Liu et al., 2019; Yuan et al., 2024a; Zha et al., 2020a, 2021a,b, 2020b; Tirer and Giryes, 2019; Dmitry et al., 2020; Shi et al., 2022). Conventional image completion methods are typically designed for a single task, such as Xie et al. (2016), Nie et al. (2012), Zha et al. (2020a, 2021a,b, 2020b), Tirer and Giryes (2019), Dmitry et al. (2020), Shi et al. (2022) and Zha et al. (2018). When dealing with image completion in the presence of noise, the performance tends to degrade. Different from normal image completion, simultaneous image denoising and completion needs to

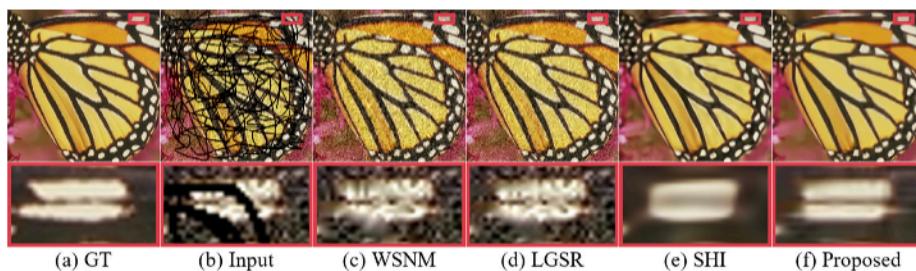
resist the interference of noise signal. At the same time, such methods should prevent noise from being transferred to the recovered entries. An intuitive approach is to perform denoising and completion sequentially (Zhang et al., 2014c,a). However, such methods result in redundant computations and increase the complexity of the algorithm. Therefore, methods that consider simultaneous denoising and completion have emerged (He and Wang, 2014). Generally, the typical linear inverse problem of reconstructing a high quality image from its degraded observed version can be expressed as:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{n} \quad (1)$$

where  $\mathbf{X}$ ,  $\mathbf{Y}$  denote the ground truth and the degraded image, respectively, and  $\mathbf{H}$  represents a non-invertible linear degradation operator.  $\mathbf{n}$  is the additive noise, which is usually assumed to obey Gaussian distribution. For simultaneous image denoising and completion problem,

\* Corresponding author.

E-mail addresses: [yuanweimin@buaa.edu.cn](mailto:yuanweimin@buaa.edu.cn) (W. Yuan), [wangyuanyuan@buaa.edu.cn](mailto:wangyuanyuan@buaa.edu.cn) (Y. Wang), [fanrr@buaa.edu.cn](mailto:fanrr@buaa.edu.cn) (R. Fan), [zhangyuxuan1996@buaa.edu.cn](mailto:zhangyuxuan1996@buaa.edu.cn) (Y. Zhang), [gmwei@buaa.edu.cn](mailto:gmwei@buaa.edu.cn) (G. Wei), [tsai@buaa.edu.cn](mailto:tsai@buaa.edu.cn) (C. Meng), [jackybxz@buaa.edu.cn](mailto:jackybxz@buaa.edu.cn) (X. Bai).



**Fig. 1.** Visual comparison of *Butterfly* mixed corrupted by additive white Gaussian noise and irregular scratches. (a) Ground Truth (PSNR/SSIM). (b) Input corrupted image (9.83/0.209). (c) WSNM (18.75/0.440). (d) LGSR (Zha et al., 2022) (19.32/0.468). (e) SHI (Shi et al., 2022) (24.52/0.835). (d) **Proposed** (24.96/0.848). Through comparison, our method achieves the performance closest to GT. The proposed method not only produces the most natural-looking effect but also effectively denoises the image and restores missing information simultaneously. Zoom-in is recommended to see the detailed comparisons.

$H$  is an indicator matrix. The diagonal elements of  $H$  are 0 or 1 and the remaining position elements are 0. The 1 and 0 in diagonal of  $H$  indicates the remaining and missing pixels, respectively.

To handle the above ill-posed problem (Zhang and Zuo, 2017), prior information is typically exploited as regularization. The latent result  $\hat{X}$  can be obtained by solving the following minimization problem:

$$\hat{X} = \arg \min_X \frac{1}{2} \|Y - HX\|_2^2 + \lambda \Psi(X) \quad (2)$$

here,  $\|Y - HX\|_2^2$  represents the  $l_2$ -norm data fidelity term, and  $\Psi(X)$  denotes the prior term.  $\lambda$  denotes the trade-off between the fidelity term and the regularization term. Many priors have been explored as regularizers to enhance image completion performance (Zhang et al., 2014c; He and Wang, 2014; Zha et al., 2020a, 2021a,b, 2022). Low rank matrix approximation (LRMA) has achieved great success as a representative regularization method recently. Various LRMA models (Nie et al., 2012; Dong et al., 2013; Yu et al., 2017; Zhang et al., 2020, 2019) impose different constraints on the original rank minimization problem.

Among those methods, the nuclear norm has been proved to be the tightest convex relaxation (Fazel, 2002). Nuclear norm minimization (NNM) has been well investigated and has achieved success in various image restoration problems. The singular value thresholding (SVT) has been proposed for solving the NNM problem, which provides the theoretical guarantee (Cai et al., 2010). However, NNM shrinks all singular values equally, which may damage important information in image. The weighted nuclear norm minimization (WNNM) (Gu et al., 2014, 2017) was proposed, assuming that larger singular values encode more information and should be shrunk less. Method proposed by Nie et al. (2012) uses the Schatten  $p$ -norm to enforce low-rank regularization. The weighted Schatten  $p$ -norm minimization (WSNM) proposed by Xie et al. (2016) has further improved the performance of WNNM. WSNM was applied to image completion by Zha et al. (2018) and achieved good performance under noisy-free conditions. Despite the success of WSNM in completing missing entries, it cannot meet the requirements of simultaneous image denoising and completion, as shown in Fig. 1(c). Lately, Liu et al. exploited smooth low tensor tree rank minimization for image completion (STTC), which could maintain global data structure (Liu et al., 2019). Nie et al. (2019) proposed the robust matrix completion based on non-convex relaxation (RMC-NC), which is specialized for noisy data. RMC-NC achieves good completion performance under low-level noise conditions.

To our knowledge, methods designed for image restoration typically perform image denoising and image completion separately (Zhang et al., 2014c; Zha et al., 2020a, 2021a,b, 2020b; Tirer and Giryes, 2019; Dmitry et al., 2020; Shi et al., 2022; Xie et al., 2016; Nie et al., 2012; Zha et al., 2018, 2022). When a method designed for one task is applied to the other task, its effectiveness is significantly reduced. In this paper, focusing on the task of simultaneous image denoising and completion, we propose a novel LRMA based method. The proposed method full leverages convolutional analysis sparse representation (ASR) and non-local statistical modeling (NLSM). Specifically, ASR is employed to ensure local smoothness, while the nonlocal self-similarity is constrained

by NLSM and WSNM is powerful for completing the missing entries. The main contributions of this paper are summarized as follows:

- The ASR model is introduced to suppress noise and smooth the local regions of the image. Instead of directly multiplying the image patch by the analysis operator, a convolutional implementation is adopted for the ASR regularization term, which makes the problem much more simple with good performance. NLSM is considered to further identify noise and preserve the accurate structures of the image. It is more effective for keeping nonlocal self-similarity of natural image than WSNM especially under the condition of noise.
- By exploiting the alternating direction method of multipliers (ADMM) (Boyd et al., 2011), the proposed method, which embeds three regularization terms into a unified framework, is solved efficiently.
- Extensive experimental results demonstrate the superiority and generalizability of the proposed method by comparing it with seventeen state-of-the-art methods (SoTAs) on four widely-used benchmarks and several challenging degradation scenarios (hyper-spectral image, remote sensing image, and historic mural). Fig. 1 shows an example of simultaneous image denoising and completion. We can see that the proposed method contains fewer visual artifacts and preserves better details compared to the competing methods.

The rest of this paper is organized as follows: Section 2 introduces related works. Section 3 presents the preliminaries. Section 4 elaborates on the proposed method. Section 5 analyzes the optimization scheme used to solve the proposed method and describes each module in detail. Extensive experiments and discussion are presented in Section 6. Finally, Section 7 concludes the paper.

## 2. Related works

### 2.1. LRMA-based methods

There has been a growing interest in exploiting low rank priors of images to address image completion. For instance, as a convex relaxation for rank minimization, nuclear norm minimization (NNM) (Fazel, 2002; Cai et al., 2010) is widely exploited for image completion. However, NNM treats different singular values equally, which can damage important information. To improve the performance, extensive research has focused on NNM variants such as truncated version (Hu et al., 2013), weighted version (Zhang et al., 2019; Gu et al., 2014, 2017), and Schatten  $p$ -norm version (Zhang et al., 2023a; Zha et al., 2018; Nie et al., 2012; Xie et al., 2016; Zhang et al., 2020, 2024c).

Specifically, Hu et al. (2013) proposed Truncated Nuclear Norm Regularization, which penalizes only the smaller singular values. Gu et al. (2014, 2017) proposed a weighted nuclear norm regularizer, which shows better performance than NNM (Fazel, 2002; Cai et al., 2010). Liu et al. (2019) exploited the low tensor tree rank minimization

for image completion. Nie et al. (2019) proposed RMC-NC, which is specialized for noisy data. RMC-NC can achieve a good completion results under low-level noise conditions. Zha et al. (2021a) proposed a NLSM based low-rank regularized group sparse coding model. Another joint low-rank and sparsity work proposed by Zha et al. (2022) is a novel low-rankness guided group sparse representation model for effective image completion. Methods (Zha et al., 2021a, 2022) exploited low-rank and sparsity as a joint prior and solved in a unified framework, can effectively recover the missing texture and structure information in noise free conditions. Li et al. (2022) introduced a novel non-convex surrogate and adaptive correlation learning approach. Motivated by the Schatten p-norm (Nie et al., 2012) and WNNM (Gu et al., 2014), Xie et al. (2016) proposed the weighted Schatten p-norm minimization (WSNM). Compared to WNNM (Gu et al., 2014), WSNM shows more flexibility in handing different ranks and achieves better approximation. Zhang et al. (2020) devised a modified Schatten p-norm, denoted as  $MS_pNM$  by incorporating Nesterov strategy for reducing complexity.  $MS_pNM$  does not have a matrix decomposable formulation like the nuclear norm or Schatten p-norm, which can reduce the computational cost. Moreover, the optimization method termed ISVTA and its fast version AISVTA were proposed for LRMA. Zhang et al. (2019) also proposed to use the double singular values function, a continuous relaxation of the rank function, in matrix rank minimization and achieved satisfactory performance. These algorithms have been applied to solve the low-rank matrix recovery problems, and shown good results on single degradation image completion task.

However, these methods struggle to cope with mixed degradation scenarios, i.e. image completion and denoising simultaneously (as shown in the following experimental section). In contrast, we introduce ASR and NLSM into the existing low-rank matrix minimization to form our proposed model, and solved efficiently using ADMM in an optimization framework. Our method organically integrates the respective strengths of ASR and NLSM to deliver promising results. Our model has the advantage of simultaneously solving challenging mixed degradation problems that are closer to real scenes. More recently, Xie et al. (2024) proposed a nonlocal self-similarity based weighted tensor low-rank decomposition model, termed NSWTL, for image completion with mixture noise. A pixel weighting strategy based on pixel credibility is employed in NSWTL to characterize the combined effects of missing pixels and noise corruption. Compared to NSWTL, our proposed model additionally considers sparse priors, that is, our model simultaneously jointly considers three different prior models, namely low-rank regularization, non-local self-similarity regularization and sparsity regularization in a unified manner, to characterize the degradation model of complex scenes.

## 2.2. Deep learning-based methods

Significant progress has been made in image completion by the emergence of deep feature learning (Zhang et al., 2017; Tirer and Giryes, 2019; Dmitry et al., 2020; Shi et al., 2022; Wei et al., 2020; Zhang et al., 2022; Yuan et al., 2024b). For example, Zhang et al. (2017) proposed a deep CNN denoiser model, IRCNN, which is solved using the half quadratic splitting technique. Tirer and Giryes (2019) presented a prior information driven by deep network for image completion using off-the-shelf denoiser. Dmitry et al. (2020) introduced deep image prior (DIP) for image completion. DIP lacks clarity on how to regulate the prior information and does not provide a clear stopping criterion, which were solved by Shi et al. (2022) through exploiting frequency band correspondence. Wei et al. (2020) proposed the tune-free plug-and-play proximal method, which determines the internal factors automatically.

More lately, there emerges a handful of methods have been designed to multiple degradations (Yu et al., 2018; Saganuma et al., 2019; Li et al., 2020; He et al., 2022; Shin et al., 2022). Specifically, Yu et al. (2018) presented a model based on reinforcement learning by training

multiple light-weight networks to deal with multiple degradations simultaneously, including remove blur, noise and JPEG. Saganuma et al. (2019) utilized the attention mechanism to choose the best selection of different strategies in parallel for unknown combined distortions. Based on the intrinsic interaction among mixed distortions, Li et al. (2020) introduced the disentangled feature learning. He et al. (2022) proposed an efficient framework by dynamic controllable residual learning.

However, methods (Yu et al., 2018; Saganuma et al., 2019; Li et al., 2020; He et al., 2022; Shin et al., 2022) completely disregard the structured distortion modes, such as holes, scratches and missing regions, focusing only on combinations of unstructured degradations like noise, blur, and JPEG artifacts. Therefore, Yu et al. (2018), Saganuma et al. (2019), Li et al. (2020), He et al. (2022) and Shin et al. (2022) may not be suitable for the task of image completion and denoising simultaneously. Furthermore, the above deep learning-based methods are appealing when aiming for high-quality performance, however the disadvantage is also generally acknowledged. These methods rely too heavily on the quantity of training data and are susceptible to multiple degradation mode. In contrast, our proposed algorithm does not require training with large-scale data. Our method is efficient and flexible enough to handle the multiple degradation in complex real scenarios.

## 3. Preliminaries

LRMA is based on the idea that matrix generated by stacking similar vectorized patches of natural image into its columns actually exhibits low rank property (Zha et al., 2018). The nonlocal self-similarity has achieved great success as a low rank prior. Specifically, dividing image  $\mathbf{X}$  into  $n$  overlapped patches  $\{\mathbf{x}_i\}(i = 1, \dots, n)$ , for each patch  $\mathbf{x}_i$  with size  $\sqrt{d} \times \sqrt{d}$ , computing the  $m$  nearest patches by K-Nearest neighbor (KNN) (Keller et al., 1985) in a  $L \times L$  searching window to constitute a set  $\mathbf{P}_i = \{\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \dots, \mathbf{x}_{i,m}\}$ . Then stacking each element of  $\mathbf{P}_i$  as a column to form a matrix denoted by  $\mathbf{X}_i$ . Patches in  $\mathbf{P}_i$  naturally have similar geometric textures and structures, which means  $\mathbf{X}_i$  has a low rank characteristic. Given a matrix  $\mathbf{R}$ , LRMA intends to recover a low-rank matrix  $\mathbf{X}$  which should be as close to  $\mathbf{R}$  as possible by solving:

$$\arg \min_{\mathbf{X}} \frac{1}{2} \|\mathbf{X} - \mathbf{R}\|_2^2 + \lambda \sum_{i=1}^n \text{rank}(\mathbf{X}_i) \quad (3)$$

here,  $\lambda$  balances the relatively importance between the low rank regularization term and data fidelity term. The weighted Schatten  $p$ -norm ( $0 < p \leq 1$ ) of matrix  $\mathbf{X}$  is defined as:

$$\|\mathbf{X}\|_{w,S_p} = \left( \sum_{i=1}^{\min\{n,m\}} w_i \delta_i^p \right)^{\frac{1}{p}} \quad (4)$$

where the weight vector  $\mathbf{w} = [w_1, \dots, w_{\min\{n,m\}}]$  is non-negative, and  $\delta_i$  is the  $i$ th singular value of  $\mathbf{X}$ . Thus the model of WSNM can be reformulated as:

$$\arg \min_{\mathbf{X}} \frac{1}{2} \|\mathbf{X} - \mathbf{R}\|_2^2 + \lambda \|\mathbf{X}\|_{w,S_p}^p \quad (5)$$

let SVD of  $\mathbf{R}$  is  $U \Sigma V^T$  with  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_r)$ . Under the hypothesis that all the singular values  $\sigma_i$  are in non-ascending order, if the weights  $w_i$  are in non-descending order,  $\delta_i$  also satisfy the non-ascending order constraint  $\delta_1 \geq \delta_2 \geq \dots \geq \delta_r$ , Eq. (5) can be converted into:

$$\min_{\delta_1, \dots, \delta_r} \sum_{i=1}^r [(\delta_i - \sigma_i)^2 + w_i \delta_i^p], i = 1, \dots, r \quad (6)$$

where the SVD of  $\mathbf{X}$  is  $U \Delta V^T$  with  $\Delta = \text{diag}(\delta_1, \dots, \delta_r)$ . Each subproblem in Eq. (6) could be efficiently solved by GST (Zuo et al., 2013), then the solution of Eq. (5) can be obtained. When it comes to image completion and the WSNM regulation being applied to image nonlocal similar patches, the problem is expressed as:

$$\arg \min_{\mathbf{X}} \frac{1}{2} \|\mathbf{Y} - \mathbf{H}\mathbf{X}\|_2^2 + \lambda \|\mathbf{X}_i\|_{w,S_p}^p \quad (7)$$

## 4. The proposed method

### 4.1. Motivation

WSNM performs well for noisy-free image completion. However, when completing entries of noisy image, the image recovered by WSNM has much noise and the local smoothness cannot be ensured, as shown in Fig. 1.(c). For simultaneous image denoising and completion, the challenges is how to make the method effective for denoising and completion simultaneously, especially when there is a large proportion of missing pixels or a high level of noise. Different from normal image completion, it is difficult to prevent noise from being transferred to the missing entries and this has a negative effect on the completion result. The completion process must eliminate the interference of noise and preserve fine structures of image at the same time.

To tackle the problem that most methods only work well in single degradation mode, we consider the two properties of natural image, namely local smoothness and nonlocal self-similarity, for simultaneous image denoising and completion. The proposed method consists of three parts: (a) the local smoothness regularization term constrained by convolutional ASR model, (b) the nonlocal self-similarity term constrained by NLSM and (c) the WSNM-based low-rank regularization term. Specifically, the convolutional ASR term can not only eliminate the effects of noise during the completion process but also effectively preserve the structures of image. The NLSM term helps to preserve nonlocal consistency, maintaining the sharpness and small structures, while WSNM is effective for completing missing entries. Our method produces a naturally looking recovered result with no distracting visual artifacts and sharp structures, as depicted in Fig. 1.(d). The details of the proposed method are provided below.

### 4.2. Convolutional ASR for local smoothness

Due to severe degradation, it is not practical to rely solely on the information from the degraded image for simultaneous image denoising and completion. We integrate the ASR (Hawe et al., 2013; Sandbichler and Schnass, 2019) model to suppress noise and preserve the local smoothness of the image. The ASR model employs an analysis operator  $\Omega$  for original signal  $X$  to generate the analysis coefficients  $\alpha = \Omega X$ . The model can be formulated under the constraint  $\|Y - HX\|_2^2 \leq \epsilon$  as:

$$X^* = \arg \min_X g(\Omega X) \quad (8)$$

here,  $X^*$  denotes the estimate signal of original signal  $X$ .  $Y$ , which is corrupted by degradation operator  $H$ , is the measured signal of  $X$ .  $g(\cdot)$  measures the sparsity.  $\epsilon \in \mathbb{R}^+$  estimates the upper bound on the noise power. The zero elements of the analysis vector  $\Omega X$  describe the subspace containing signal  $X$ .

We employ the patch-based analysis operator Geometric Analysis operator Learning (GOAL) (Hawe et al., 2013), which has shown competitive performance in image reconstruction. To obtain the dictionary  $\Omega$ , we first pick five undegraded natural images and divide them into  $\sqrt{d} \times \sqrt{d}$  image blocks. Then, we randomly select 200,000 image blocks to form the training samples, representing each block as a column vector and perform normalization. Finally, the geometric conjugate gradient method is used to solve the optimization problem, yielding the learned analysis operator  $\Omega$ . The size of the concrete operator is  $128 \times 64$ . As shown in Fig. 3, the input image exhibits very sparse responses to GOAL, which not only preserves large textures but also ensures local smoothness. We execute a convolution implementation on  $X$  and each column of the operator, and the response is assumed to be sparse.

Since non-convex  $l_q$ -norm minimization often yields more sparse results than convex  $l_1$ -norm minimization and may produce better results in many sparse coding-based image restoration problems, we adopt the  $l_q$ -norm minimization with power  $q = 0.4$  on sum of the

convolution of each column  $\Omega_j$  in the GOAL operator and the whole image  $X$ . Note that this convolutional implementation can also achieve a global effect during the completion process. More importantly, it helps to solve the whole proposed method efficiently while keeping a good performance at the same time. Therefore, the convolutional ASR term is obtained as:

$$\Psi_{ASR} = \sum_{j=1}^J \|\Omega_j \otimes \tilde{X}\|_q^q \quad (9)$$

where  $J$  is the total number of column in analysis operator  $\Omega$ ,  $\otimes$  indicates the convolution operator.  $\tilde{X} \in \mathbb{R}^{h \times w}$  represents the original signal and  $X_v \in \mathbb{R}^N$  is the vectorization of  $\tilde{X}$  with  $N = h \times w$ , i.e.,  $X_v = \text{vec}(\tilde{X})$ .

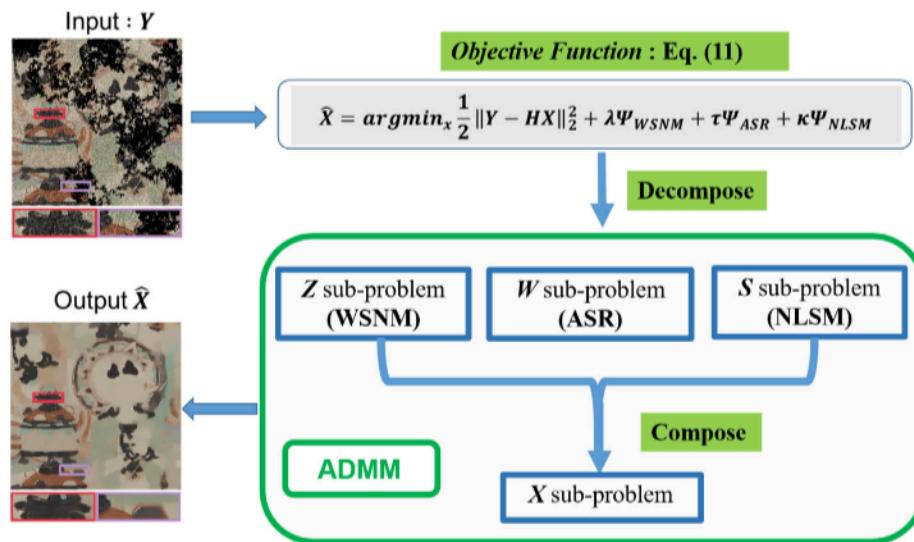
Fig. 4 visually illustrates the effect of only adopting the convolutional ASR with GOAL operator during image completion process. Fig. 4(a) is the input image of *House* corrupted by 50% random pixel missing with noise level  $\sigma_n = 30$ . From Fig. 4(b)–(e), we can observe that as the iteration number  $k$  increases, the noise is progressively suppressed and the local smoothness of image is gradually recovered. The sparsity of analysis coefficients implies that the signal should be orthogonal to most of atoms in the analysis operator. However, ASR alone may struggle to preserve fine textures in images with complex patterns. Therefore, an additional regularization term is required to maintain the non-local self-similarity of the image.

### 4.3. NLSM for nonlocal self-similarity

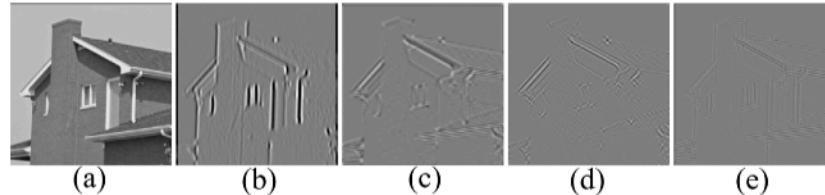
Besides local smoothness, NLSM is also significant for natural images because it describes the repetitive structures within the nonlocal areas of an image. Therefore, NLSM can maintain fine textures and details to preserve the nonlocal consistency of an image. As shown in Fig. 5, block regions with the same color have repetitive textures and can be regarded as similar patches. By imposing constraints on those similar patches, we can measure the NLSM of the entire image. Based on the idea of searching few patches with the most similarity, stacking them into a 3D tensor and shrinking the transform coefficients (Zha et al., 2020a), NLSM aims to characterize the nonlocal self-similarity of the image, which is designed by the following steps. First, divide the whole image  $X$  with size  $N$  into  $n$  overlapped patches  $\{x_i\}_{i=1 \dots n}$ . Second, for each patch  $x_i$  with the size of  $\sqrt{d} \times \sqrt{d}$ , computing the  $c$  nearest patches in a  $L \times L$  local region of the image to constitute a set  $S_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,c}\}$ . Third, stacking patches in  $S_i$  into a 3D array, which is denoted by  $Z_i$ . Fourth, the coefficients of orthogonal 3D transform for each  $Z_i$  could be obtained. Note that the orthogonal 3D transform could be decomposed into a 2D discrete cosine transform and an 1D Haar transform. Denote the  $T_i(\cdot)$  as the operator of an orthogonal 3D transform, and  $T_i$  as the transform coefficients for  $Z_i$ . Finally, denote  $\Theta_X$  as the vector constructed from all the transform coefficients of  $T_i$  arranged in lexicographic order. The NLSM imposes constraint on means of the distributions of the transform coefficients, most of the coefficients are near zero (Zhang et al., 2014c). The formulation of NLSM in 3D transform domain is obtained by:

$$\Psi_{NLSM} = \|\Theta_X\|_1 = \sum_{i=1}^n \|T_i\|_1 \quad (10)$$

After that, there is an inverse operator  $\Xi_{NLSM}$  corresponding to  $\Psi_{NLSM}$  for generating estimate of patches and further getting the final estimate image  $X = \Xi_{NLSM}(\Theta_X)$ . Since patches are estimated in a manner of 3D array, the final image is generated through putting the estimate patches back to their original positions and averaging them. As shown in Fig. 6, NLSM is effective for preserving image details. It imposes constraints on the transform coefficient distributions to ensure nonlocal self-similarity, which is crucial for preserving finer textures and more accurate structures in an image while further reducing noise interference.



**Fig. 2.** The flowchart illustrates our proposed method for simultaneous image denoising and completion. The degraded image is reconstructed using the proposed model by jointly incorporating the ASR and NLSM into existing WSNM to form the objective function Eq. (11), which is iteratively solved through the ADMM framework. Specifically, the objective function is divided into  $Z$ ,  $W$ , and  $S$  sub-problems and solved separately, then the reconstruction result  $X$  is obtained by combining these three modules.



**Fig. 3.** (a) Input image. (b)–(e) Responses of the input image to four columns of the learned  $128 \times 64$  GOAL operator.

#### 4.4. The objective function

By incorporating the convolutional ASR model with GOAL as the analysis operator Eq. (9) and the NLSM term Eq. (10) into the WSNM framework Eq. (7), the proposed simultaneous image denoising and completion method can be formulated as:

$$\hat{X} = \arg \min_X \frac{1}{2} \|Y - HX\|_2^2 + \lambda \Psi_{WSNM} + \tau \Psi_{ASR} + \kappa \Psi_{NLSM} \quad (11)$$

here  $\lambda$ ,  $\tau$  and  $\kappa$  are three trade-off parameters. The first two terms represent the WSNM baseline for completing missing entries, while the third and fourth terms ensure the local smoothness and edge preservation, respectively. Moreover, ASR term is also the major part for suppressing noise. The large structures and fine textures can both be preserved after making a trade-off between the above constraints. Fig. 2 presents the flowchart of the proposed method.

#### 5. Solution for the proposed method

Tackling our objective function in Eq. (11) presents significant challenges due to its nature as a large-scale non-convex optimization problem. To enhance the tractability and robustness of our solution, we adopt the alternating direction method of multipliers (ADMM) (Boyd et al., 2011) for solving Eq. (11) in this paper. ADMM is an effective approach for addressing various large-scale optimization problems. Its core concept involves transforming an unconstrained minimization problem into several constrained sub-problems through variable splitting. Numerical simulations have demonstrated that ADMM can achieve convergence with minimal memory usage, making it particularly appealing for a wide range of large-scale optimization tasks (Zhang et al., 2024b,a, 2023b).

Specifically, auxiliary variables  $Z$ ,  $W$  and  $S$  are introduced, Eq. (11) is converted into:

$$\begin{aligned} \hat{X} = \arg \min_X & \frac{1}{2} \|Y - HX\|_2^2 + \lambda \sum_{i=1}^n \|Z_i\|_{w,S_p}^p + \frac{\mu_1}{2} \|X - Z - B\|_2^2 + \\ & \tau \sum_{j=1}^J \|\Omega_j \otimes W\|_q^q + \frac{\mu_2}{2} \|\tilde{X} - W - C\|_2^2 + \kappa \|\Theta_S\|_1 + \frac{\mu_3}{2} \|X - S - D\|_2^2 \end{aligned} \quad (12)$$

here  $B$ ,  $C$  and  $D$  are three Lagrange multipliers.  $\mu_1$ ,  $\mu_2$  and  $\mu_3$  are the corresponding parameters. By applying the ADMM framework, we can iteratively solve Eq. (12) by dividing it into the following four sub-problems:

$$\begin{aligned} X^{k+1} = & \min_X \frac{1}{2} \|Y - HX\|_2^2 + \frac{\mu_1}{2} \|X - Z^k - B^k\|_2^2 \\ & + \frac{\mu_2}{2} \|X - \text{vec}(W^k + C^k)\|_2^2 + \frac{\mu_3}{2} \|X - S^k - D^k\|_2^2 \end{aligned} \quad (13)$$

$$Z^{k+1} = \min_Z \frac{\mu_1}{2} \|X^{k+1} - Z - B^k\|_2^2 + \lambda \sum_{i=1}^n \|Z_i\|_{w,S_p}^p \quad (14)$$

$$W^{k+1} = \min_W \tau \sum_{j=1}^J \|\Omega_j \otimes W\|_q^q + \frac{\mu_2}{2} \|\tilde{X}^{k+1} - W - C^k\|_2^2 \quad (15)$$

$$S^{k+1} = \min_S \kappa \|\Theta_S\|_1 + \frac{\mu_3}{2} \|X^{k+1} - S - D^k\|_2^2 \quad (16)$$

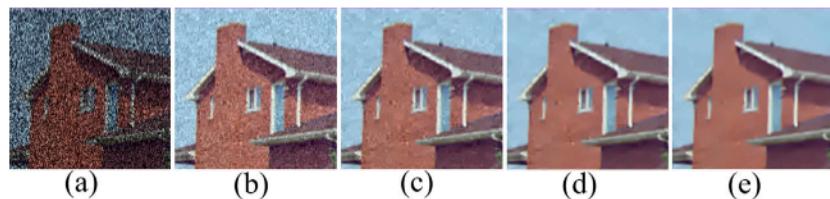
with

$$B^{k+1} = B^k - (X^{k+1} - Z^{k+1}) \quad (17)$$

$$C^{k+1} = C^k - (\tilde{X}^{k+1} - W^{k+1}) \quad (18)$$

$$D^{k+1} = D^k - (X^{k+1} - S^{k+1}) \quad (19)$$

here,  $k$  denotes the iteration number. Then we present that each sub-problem can be efficiently solved.



**Fig. 4.** Simultaneous image denoising and completion effects adopting only the convolutional ASR with GOAL for *House* as the iteration number  $k$  increases. (a) Image mixed corrupted by 50% random pixels missing and noise level  $\sigma_n = 30$ . (b)  $k = 5$ . (c)  $k = 10$ . (d)  $k = 15$ . (e)  $k = 20$ .



**Fig. 5.** Illustrations by *Barbara* for the nonlocal self-similarity. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 6.** Simultaneous image denoising and completion effects adopting only the NLSM for *Barbara*. (a) Corrupted image by text mask covering with noise level  $\sigma_n = 30$ . (b) Recovered result by NLSM.

### 5.1. X sub-problem

Fixing all other parameters,  $X$  sub-problem Eq. (13) can be written as:

$$\begin{aligned} \min_X \frac{1}{2} \|Y - HX\|_2^2 + \frac{\mu_1}{2} \|X - Z - B\|_2^2 \\ + \frac{\mu_2}{2} \|X - \text{vec}(W + C)\|_2^2 + \frac{\mu_3}{2} \|X - S - D\|_2^2 \end{aligned} \quad (20)$$

due to the quadratic optimization, Eq. (20) has the closed-form solution as:

$$X = (H^T H + (\mu_1 + \mu_2 + \mu_3)I)^{-1} \cdot \{H^T Y \\ + \mu_1(Z + B) + \mu_2(\text{vec}(W + C)) + \mu_3(S + D)\} \quad (21)$$

where  $I$  denotes the identity matrix with desired dimensions.

### 5.2. Z sub-problem

The  $Z$  sub-problem Eq. (14) can be solved by:

$$\min_Z \frac{\mu_1}{2} \|Z - R\|_2^2 + \lambda \sum_{i=1}^n \|Z_i\|_{w, S_p}^p \quad (22)$$

where  $R = X - B$ . By making a general hypothesis, Eq. (22) can have a closed-form solution. With this assumption we have the theorem below.

**Theorem 1.** Define  $Z, R \in \mathcal{R}^N$ , and  $Z_i, R_i \in \mathcal{R}^{d \times m}$ ,  $e(j)$  as each element of  $e$  with  $e = Z - R$ ,  $j = 1, \dots, N$ . Assume that  $e(j)$  follows an independent

distribution with zero mean and variance  $\delta^2$ . For any  $\epsilon > 0$ , the relationship between  $\frac{1}{N} \|Z - R\|_2^2$  and  $\frac{1}{K} \sum_{j=1}^J \|Z_j - R_j\|_F^2$  is

$$\lim_{N \rightarrow \infty, K \rightarrow \infty} P\left\{ \left| \frac{1}{N} \|Z - R\|_2^2 - \frac{1}{K} \sum_{j=1}^J \|Z_j - R_j\|_F^2 \right| < \epsilon \right\} = 1 \quad (23)$$

where  $P(\cdot)$  denotes the probability and  $K = d \times m \times n$ .

According to Theorem 1, Eq. (22) can be converted into:

$$\min_{Z_i} \sum_{i=1}^n \left( \frac{1}{2} \|Z_i - R_i\|_2^2 + \eta \|Z_i\|_{w, S_p}^p \right) \quad (24)$$

here  $\eta = \lambda K / \mu_1 N$ .  $Z$  can be obtained by efficiently solving the above  $n$  sub-problems for each  $Z_i$  that is formed by the  $m$  similar patches of  $Z$ .

The SVD of  $R_i$  in Eq. (24) is  $U_i \Sigma_i V_i^T$  with  $\Sigma_i = \text{diag}(\sigma_{i1}, \sigma_{i2}, \dots, \sigma_{ir})$ . To reconstruct  $X_i$ , the higher singular values conveyed major edge and texture information should be kept, and the smaller singular values should be decreased. Therefore, regarding the shrinking weight  $\omega_{i,j}$ , we have:

$$\omega_{i,j} = \frac{1}{|\sigma_{i,j}| + \zeta} \quad (25)$$

here  $\zeta$  is a small constant. Under the hypothesis that all the singular values are in non-ascending order, we can calculate the solution of Eq. (24) by  $Z_i = U_i \Delta_i V_i^T$  where  $\Delta_i = \text{diag}(\delta_{i1}, \delta_{i2}, \dots, \delta_{ir})$ .  $\delta_{ij}$  is obtained by solving the following  $r$  independent sub-problems:

$$\min_{\delta_{ij} \geq 0} f(\delta_{ij}) = (\delta_{ij} - \sigma_{ij})^2 + \eta w_{ij} \delta_{ij}^p, j = 1, \dots, r \quad (26)$$

each sub-problem in Eq. (26) can be succinctly solved by GST (Zuo et al., 2013), presented in Algorithm 1. Therefore the closed-form of Eq. (24) is:

$$\delta_{ij} = GST(\sigma_{ij}, \eta w_{ij}, p, K) \quad (27)$$

where  $K$  denotes the iteration number. The regularization parameter  $\lambda$  is set as:

$$\lambda = \frac{2\sqrt{2}\sigma_n}{\Delta_i + \varrho} \quad (28)$$

here,  $\varrho$  is a small constant,  $Z$  is updated by aggregating all groups of  $Z_i$ .

### 5.3. W sub-problem

The  $W$  sub-problem Eq. (15) can be formulated as:

$$\min_W \tau \sum_{j=1}^J \|\Omega_j \otimes W\|_q^q + \frac{\mu_2}{2} \|\tilde{X} - W - C\|_2^2 \quad (29)$$

updating  $W$  is actually a sparse gradient-based deconvolution problem, we employ shrinkage function and associated generalized iterated shrinkage algorithm (GISA) (Zuo et al., 2013) to solve Eq. (29). We introduce the group of auxiliary variables  $\{\Omega_j = \Omega_j \otimes W\}_{j=1, \dots, J}$ , and rewrite the Eq. (29) as:

$$\min_W \frac{1}{2} \|\tilde{X} - W - C\|_2^2 + \frac{\xi \tau}{2\mu_2} \sum_{j=1}^J \|\Omega_j \otimes W - \Omega_j\|_2^2 + \frac{\tau}{\mu_2} \sum_{j=1}^J \|\Omega_j\|_q^q \quad (30)$$

**Algorithm 1**  $GST : T_p^{GST}(y; \lambda) = GST(y, \lambda, p, J)$ 

```

Require:  $y, \lambda, p, J;$ 
 $\tau_p^{GST}(\lambda) = (2\lambda(1-p))^{\frac{1}{2-p}} + \lambda p(2\lambda(1-p))^{\frac{p-1}{2-p}}$ 
if  $|y| \leq \tau_p^{GST}(\lambda)$  then
     $T_p^{GST}(y; \lambda) = 0$ 
else
     $k = 0, x^{(k)} = |y|$ 
    Iterate on  $k = 0, 1, \dots, J$ 
     $x^{(k+1)} = |y| - \lambda p(x^{(k)})^{p-1}$ 
     $k = k + 1$ 
     $T_p^{GST}(y; \lambda) = sgn(y)x^{(k)}$ 
end if
return  $T_p^{GST}(y; \lambda).$ 

```

we can observe that when  $\xi \rightarrow \infty$ , Eq. (30) would have the same solution as Eq. (29). We use alternating minimization to solve Eq. (30).

**5.3.1. Updating  $\mathbf{W}$** 

Given fixed  $\{\mathbf{Q}_j\}_{j=1,\dots,J}$ , we consider:

$$\min_{\mathbf{W}} \frac{1}{2} \|\tilde{\mathbf{X}} - \mathbf{W} - \mathbf{C}\|_2^2 + \frac{\xi\tau}{2\mu_2} \sum_{j=1}^J \|\mathbf{Q}_j \otimes \mathbf{W} - \mathbf{Q}_j\|_2^2 \quad (31)$$

fast Fourier transform (FFT) is exploited, thus the closed-form solution of  $\mathbf{W}$  can be acquired as:

$$\mathbf{W} = \mathcal{F}^{-1} \left( \frac{\mu_2 \mathcal{F}(\tilde{\mathbf{X}} - \mathbf{C}) + \xi\tau \sum_{j=1}^J \overline{\mathcal{F}(\mathbf{Q}_j)} \circ \mathcal{F}(\mathbf{Q}_j)}{\mu_2 + \xi\tau \sum_{j=1}^J \overline{\mathcal{F}(\mathbf{Q}_j)} \circ \mathcal{F}(\mathbf{Q}_j)} \right) \quad (32)$$

$\mathcal{F}$  denotes the 2D Fourier transform and  $\mathcal{F}^{-1}$  indicates the inverse of  $\mathcal{F}$ .  $\bar{\cdot}$  is the complex conjugate operator,  $\circ$  is the component-wise multiplication.

**5.3.2. Updating  $\mathbf{Q}_j$** 

Given  $\mathbf{W}$ ,  $\mathbf{Q}_j$  can be obtained by:

$$\min_{\mathbf{Q}_j} \frac{\xi}{2} \sum_{j=1}^J \|\mathbf{Q}_j - \mathbf{Q}_j^{ref}\|_2^2 + \sum_{j=1}^J \|\mathbf{Q}_j\|_q^q \quad (33)$$

$$\mathbf{Q}_j^{ref} = \mathbf{Q}_j \otimes \mathbf{W} \quad (34)$$

the solution of  $\mathbf{Q}_j$  can be obtained by using GST:

$$\mathbf{Q}_j = T_q^{GST} \left( \mathbf{Q}_j^{ref}; \frac{1}{\xi} \right) \quad (35)$$

The complete algorithm for  $\mathbf{W}$  sub-problem is presented in **Algorithm 2**.

**Algorithm 2** Alternating minimization for  $\mathbf{W}$  sub-problem

```

Require:  $\tilde{\mathbf{X}} - \mathbf{C}, \mathbf{Q}, \mu_2, \tau, q,$ 
Initialize  $k = 0, \rho, T, \xi^{(0)}, \mathbf{W}^{(0)}, \mathbf{Q}^{(0)}$ .
for  $k = 1$  to  $T$  do
    Update  $\mathbf{W}^{(k+1)}$  by Eq. (30)
     $\mathbf{Q}^{ref} = \mathbf{Q} \otimes \mathbf{W}^{(k+1)}$  by Eq. (34)
     $\mathbf{Q}^{k+1} = GST(\mathbf{Q}^{ref}, \frac{1}{\xi})$  by Eq. (35)
     $\xi^{(k+1)} = \rho\xi^{(k)}$ 
     $k = k + 1$ 
end for
W =  $\mathbf{W}^k$ 
return  $\mathbf{W}$ .

```

**5.4. S sub-problem**

Fixing other parameters,  $S$  sub-problem Eq. (16) can be written as:

$$\min_S \kappa \|\Theta_S\|_1 + \frac{\mu_3}{2} \|S - (\mathbf{X} - \mathbf{D})\|_2^2 \quad (36)$$

the solution of Eq. (36) is depended on the same account as Eq. (22), the closed-form solution of Eq. (36) could be inferred via a soft-thresholding as:

$$S = \Theta^{-1} \left( soft(\Theta_{X-D}, \sqrt{2\kappa dmn/\mu_3 N}) \right) \quad (37)$$

where  $\Theta^{-1}$  denotes the reverse operator of  $\Theta_{(\cdot)}$ .

We summarize the proposed method in **Algorithm 3**.

**Algorithm 3** The Proposed Algorithm.

**Require:** The observed image  $\mathbf{Y}$ , the indicator matrix  $\mathbf{H}$ ;

Initialize  $\mu_1, \mu_2, \mu_3, \tau, \kappa, \mathbf{B} = \mathbf{0}, \mathbf{C} = \mathbf{0}, \mathbf{D} = \mathbf{0}$ ;

**while** not converged **do**

    Compute  $\Delta_i$  by Eq. (26) and calculate  $Z_i = U_i \Delta_i V_i^T$ , then aggregate  $Z_i$  to estimate  $Z$ .

    Compute  $\mathbf{W}$  sub-problem by Eq. (30).

    Compute  $S$  sub-problem by Eq. (36).

    Compute  $X$  sub-problem by Eq. (21).

    Update  $\mathbf{B}$  by Eq. (17).

    Update  $\mathbf{C}$  by Eq. (18).

    Update  $\mathbf{D}$  by Eq. (19).

**end while**

**return** The final reconstructed image  $\hat{\mathbf{X}}$ .

**6. Experiment**

In this section, we evaluate the performance of the proposed method. Specifically, the implementation details are provided in Section 6.1, and the performance comparison is detailed in Section 6.2. We further analyze the proposed method including: parameter selection in Section 6.3, ablation study in Section 6.4, computational efficiency and convergence in Section 6.5, and exploring on challenging scenarios in Section 6.6.

**6.1. Implementation details**

**Setting:** We apply the proposed method to 6 types of mixed degradation modes in the simulation. Specifically, the input image is either random missing pixels or covered by a binary mask, and is corrupted by additive white Gaussian noise (AWGN). As illustrated in Fig. 7, the degradation cases are defined as follows: **Case 1:** 50% random pixel missing with noise level  $\sigma_n = 30$ , **Case 2:** 40% random pixel missing with noise level  $\sigma_n = 40$ , **Case 3:** 30% random pixel missing with noise level  $\sigma_n = 50$ , **Case 4:** line mask covering with noise level  $\sigma_n = 30$ , **Case 5:** grid mask covering with noise level  $\sigma_n = 40$ , **Case 6:** text mask covering with noise level  $\sigma_n = 50$ . Here  $\sigma_n$  is the variance of AWGN.

**Datasets:** We perform evaluation on four widely used benchmarks: Set12 (Zhang et al., 2018), Kodak (Franzen, 1999), McMaster (Zhang et al., 2011), and CBSD68 (Martin et al., 2001) dataset. More specifically, Set12 contains 12 classic images with size of  $256 \times 256$  or  $512 \times 512$ ; Kodak contains 24 color images with size of  $512 \times 512$ ; McMaster contains 18 color images with size of  $500 \times 500$ ; CBSD68 contains 68 color images with size of  $321 \times 481$ .

**Compared Methods:** We compare our method with 17 advanced image completion methods. 12 classic methods: ISDSB (He and Wang, 2014), WLP (Zha et al., 2018), GSR (Zhang et al., 2014b), JSM (Zhang et al., 2014c), BPFA (Zhou et al., 2012), RMC-NC (Nie et al., 2019), STTC-A (Liu et al., 2019), JPG-SR (Zha et al., 2020a), LR-GSC (Zha et al., 2021a), HSSE (Zha et al., 2021b), LGSR (Zha et al., 2022), and

## ARTICLE IN PRESS

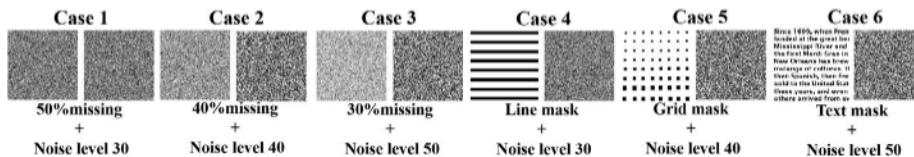


Fig. 7. 6 cases of degradation mode employed in image inpainting with noise intervention.

Table 1

Comparisons of average PSNR/SSIM to 17 methods on Set12 (Zhang et al., 2018) dataset. The best result is highlighted in boldface.

Method	50% missing + $\sigma_n = 30$	40% missing + $\sigma_n = 40$	30% missing + $\sigma_n = 50$	Line mask + $\sigma_n = 30$	Grid mask + $\sigma_n = 40$	Text mask + $\sigma_n = 50$
Input	8.28/0.127	8.92/0.143	9.49/0.154	11.37/0.287	13.79/0.300	10.86/0.182
ISDSB (He and Wang, 2014)	20.98/0.491	18.28/0.389	16.07/0.317	18.69/0.432	16.77/0.366	15.49/0.301
WLP (Zha et al., 2018)	19.67/0.399	19.65/0.314	15.04/0.239	18.99/0.371	16.48/0.296	14.57/0.228
GSR (Zhang et al., 2014b)	19.51/0.419	17.05/0.343	14.90/0.279	18.69/0.448	16.24/0.246	14.57/0.279
JSM (Zhang et al., 2014c)	20.61/0.477	17.78/0.380	15.53/0.294	18.68/0.449	16.50/0.357	14.86/0.293
RMC-NC (Nie et al., 2019)	21.83/0.701	20.30/0.667	18.74/0.621	19.28/0.639	20.37/0.686	18.61/0.619
STTC-A (Liu et al., 2019)	20.23/0.380	17.51/0.284	15.22/0.218	18.10/0.336	16.24/0.256	14.78/0.209
BPFA (Zhou et al., 2012)	26.48/0.789	25.74/0.756	25.07/0.727	23.00/0.652	25.50/0.764	23.78/0.654
JPG-SR (Zha et al., 2020a)	19.83/0.341	17.30/0.253	15.18/0.197	18.77/0.330	16.32/0.236	14.84/0.191
LR-GSC (Zha et al., 2021a)	19.94/0.341	17.44/0.313	15.17/0.241	18.74/0.333	16.49/0.298	14.60/0.229
HSSE (Zha et al., 2021b)	20.44/0.358	17.77/0.318	15.44/0.238	18.76/0.334	16.28/0.232	14.65/0.229
LGSR (Zha et al., 2022)	20.07/0.347	17.60/0.257	15.38/0.190	18.71/0.337	16.27/0.236	14.84/0.192
GSRC-NLP (Zha et al., 2020b)	20.72/0.368	17.84/0.266	15.29/0.201	18.79/0.337	16.32/0.243	14.76/0.191
IDBP (Tirer and Giryes, 2019)	26.64/0.810	25.87/0.780	25.34/0.750	23.17/0.714	25.58/0.783	24.11/0.711
DIP (Dmitry et al., 2020)	25.37/0.724	23.54/0.708	23.13/0.685	23.17/0.705	23.71/0.723	23.01/0.688
DD (Heckel and Hand, 2019)	24.54/0.738	24.86/0.749	24.47/0.730	24.43/0.764	25.40/0.771	24.22/0.716
NAS-DIP (Chen et al., 2020)	26.61/0.789	25.28/0.701	24.59/0.736	24.41/0.715	25.59/0.741	24.77/0.717
SHI (Shi et al., 2022)	26.29/0.783	25.53/0.746	25.35/0.762	24.10/0.771	26.18/0.779	24.91/0.725
Ours	<b>26.89/0.819</b>	<b>26.22/0.795</b>	<b>25.67/0.777</b>	<b>24.67/0.787</b>	<b>26.41/0.811</b>	<b>25.38/0.769</b>

Table 2

Comparisons of average PSNR/SSIM to 17 methods on Kodak (Franzen, 1999) dataset. The best result is highlighted in boldface.

Method	50% missing + $\sigma_n = 30$	40% missing + $\sigma_n = 40$	30% missing + $\sigma_n = 50$	Line mask + $\sigma_n = 30$	Grid mask + $\sigma_n = 40$	Text mask + $\sigma_n = 50$
Input	9.43/0.167	9.96/0.186	10.35/0.197	12.37/0.336	<b>14.21/0.350</b>	11.55/0.209
ISDSB (He and Wang, 2014)	21.19/0.574	18.36/0.465	16.10/0.372	19.31/0.521	<b>16.77/0.427</b>	15.46/0.357
WLP (Zha et al., 2018)	20.04/0.472	17.37/0.391	15.22/0.334	19.40/0.551	16.88/0.482	14.38/0.401
GSR (Zhang et al., 2014b)	19.77/0.451	17.18/0.366	14.94/0.310	19.11/0.517	16.35/0.419	14.59/0.317
JSM (Zhang et al., 2014c)	20.77/0.541	17.79/0.407	15.44/0.334	19.11/0.517	16.38/0.414	14.84/0.343
RMC-NC (Nie et al., 2019)	22.32/0.802	20.51/0.766	18.87/0.720	20.98/0.750	20.67/0.787	18.87/0.720
STTC-A (Liu et al., 2019)	20.27/0.531	17.60/0.434	15.30/0.367	17.99/0.507	16.33/0.458	14.80/0.377
BPFA (Zhou et al., 2012)	27.47/0.801	26.71/0.754	26.22/0.730	24.24/0.689	25.34/0.738	23.49/0.611
JPG-SR (Zha et al., 2020a)	19.99/0.280	17.41/0.198	15.32/0.145	19.12/0.289	16.39/0.193	14.89/0.152
LR-GSC (Zha et al., 2021a)	20.18/0.294	17.87/0.186	15.55/0.151	19.43/0.310	16.67/0.203	14.23/0.167
HSSE (Zha et al., 2021b)	20.61/0.294	17.94/0.211	15.54/0.153	19.64/0.315	16.14/0.183	14.47/0.149
LGSR (Zha et al., 2022)	20.31/0.296	17.94/0.179	15.34/0.157	19.12/0.307	16.34/0.207	14.84/0.167
GSRC-NLP (Zha et al., 2020b)	20.90/0.301	17.97/0.212	15.46/0.146	19.77/0.363	16.42/0.149	14.84/0.130
IDBP (Tirer and Giryes, 2019)	27.51/0.805	26.86/0.774	26.37/0.746	24.32/0.721	26.31/0.768	24.56/0.668
DIP (Dmitry et al., 2020)	26.57/0.784	26.39/0.779	24.83/0.735	23.12/0.455	23.53/0.471	23.69/0.494
DD (Heckel and Hand, 2019)	27.13/0.804	26.43/0.772	25.88/0.761	24.81/0.770	27.17/0.819	26.07/0.768
NAS-DIP (Chen et al., 2020)	27.39/0.799	26.41/0.760	26.32/0.771	26.39/0.790	27.58/0.820	26.19/0.730
SHI (Shi et al., 2022)	26.48/0.766	25.52/0.709	24.60/0.663	24.99/0.743	25.93/0.756	24.71/0.696
Ours	<b>27.91/0.826</b>	<b>27.30/0.802</b>	<b>26.79/0.780</b>	<b>26.51/0.796</b>	<b>28.02/0.820</b>	<b>26.55/0.773</b>

GSRC-NLP (Zha et al., 2020b). 5 recently proposed deep learning-based methods: IDBP (Tirer and Giryes, 2019), DIP (Dmitry et al., 2020), DD (Heckel and Hand, 2019), NAS-DIP (Chen et al., 2020), and SHI (Shi et al., 2022). PSNR and SSIM (Wang et al., 2004) on Y channel of YCrCb are exploited as metrics in the YCbCr color space in a way similar to prior works (Zha et al., 2020a, 2021a,b, 2022). The results for all algorithms are obtained by carefully tuning the parameters of both the proposed method and the compared methods, based on the default values from the released code or the suggestions in the published papers, to best suit the task of simultaneous image denoising and completion.

## 6.2. Performance on datasets

**Quantitative Comparison.** Table 1, Table 2, Tables 3 and 4 depict the average PSNR/SSIM results of compared methods on Set12 (Zhang et al., 2018), Kodak (Franzen, 1999), McMaster (Zhang et al., 2011) and CBSD68 (Martin et al., 2001), respectively. It can be observed that the proposed method performs the best in all the 6 degradation cases among the 17 comparison methods, demonstrating its effectiveness for simultaneous image denoising and completion. Taking Table 1 as an example, under 6 degradation cases, our method obtains (46.3%, 49.2%, 53.9%, 49.9%, 30.3%, 52.1%, 3.8%, 51.9%, 51.7%, 50.5%, 51.2%, 50.1%, 3.0%, 9.5%, 5.6%, 3.8%, 1.9%) gains in average PSNR and (108.5%, 160.4%, 138.8%, 114.1%, 20.9%, 182.2%, 9.5%, 209.7%, 173.4%, 181.2%, 207.4%, 202.7%, 4.8%, 12.6%, 6.4%, 13.6%, 4.2%) gains in average SSIM.

## ARTICLE IN PRESS

Table 3

Comparisons of average PSNR/SSIM to 17 methods on McMaster (Zhang et al., 2011) dataset. The best result is highlighted in boldface.

Method	50% missing + $\sigma_n = 30$	40% missing + $\sigma_n = 40$	30% missing + $\sigma_n = 50$	Line mask + $\sigma_n = 30$	Grid mask + $\sigma_n = 40$	Text mask + $\sigma_n = 50$
Input	10.28/0.218	10.70/0.231	10.93/0.236	13.20/0.370	14.52/0.388	11.95/0.247
ISDSB (He and Wang, 2014)	21.51/0.623	18.66/0.511	16.37/0.419	19.28/0.551	16.99/0.471	15.77/0.381
WLP (Zha et al., 2018)	20.14/0.521	17.45/0.439	15.28/0.367	19.84/0.555	16.18/0.512	14.14/0.380
GSR (Zhang et al., 2014b)	19.88/0.494	17.16/0.412	15.04/0.348	18.98/0.546	16.31/0.442	14.61/0.368
JSM (Zhang et al., 2014c)	20.78/0.569	17.91/0.456	15.47/0.371	18.96/0.542	16.31/0.444	14.88/0.366
RMC-NC (Nie et al., 2019)	22.61/0.788	20.64/0.745	18.89/0.697	20.41/0.690	20.57/0.754	18.78/0.691
STTC-A (Liu et al., 2019)	20.44/0.581	17.64/0.484	15.30/0.423	17.62/0.534	16.32/0.502	14.87/0.429
BPFA (Zhou et al., 2012)	28.17/0.824	27.32/0.789	26.67/0.760	24.21/0.718	26.80/0.797	24.81/0.671
JPG-SR (Zha et al., 2020a)	20.13/0.289	17.45/0.208	15.34/0.151	19.05/0.291	16.37/0.214	14.92/0.164
LR-GSC (Zha et al., 2021a)	20.09/0.348	17.68/0.259	15.39/0.193	18.24/0.325	16.43/0.294	14.96/0.176
HSSE (Zha et al., 2021b)	20.71/0.279	17.86/0.238	15.27/0.112	18.85/0.332	16.97/0.239	14.89/0.176
LGSR (Zha et al., 2022)	20.13/0.353	17.71/0.262	15.43/0.199	18.94/0.310	16.34/0.219	14.77/0.154
GSRC-NLP (Zha et al., 2020b)	20.88/0.346	17.92/0.269	15.41/0.190	18.42/0.387	16.79/0.267	14.88/0.134
IDBP (Tirer and Giryes, 2019)	28.36/0.843	27.52/0.810	26.93/0.783	24.35/0.749	26.71/0.808	25.01/0.702
DIP (Dmitry et al., 2020)	26.18/0.779	26.21/0.771	25.07/0.729	24.11/0.532	25.30/0.749	23.30/0.666
DD (Heckel and Hand, 2019)	27.58/0.847	26.49/0.817	25.68/0.799	23.74/0.792	27.62/0.809	26.25/0.715
NAS-DIP (Chen et al., 2020)	27.68/0.854	26.60/0.818	25.81/0.809	24.60/0.817	27.28/0.827	26.57/0.724
SHI (Shi et al., 2022)	26.89/0.818	26.04/0.784	24.62/0.741	24.92/0.771	25.79/0.788	24.65/0.748
Ours	<b>28.66/0.859</b>	<b>27.95/0.838</b>	<b>27.37/0.819</b>	<b>26.43/0.821</b>	<b>28.02/0.845</b>	<b>26.96/0.808</b>

Table 4

Comparisons of average PSNR/SSIM to 17 methods on CBSD68 (Martin et al., 2001) dataset. The best result is highlighted in boldface.

Method	50% missing + $\sigma_n = 30$	40% missing + $\sigma_n = 40$	30% missing + $\sigma_n = 50$	Line mask + $\sigma_n = 30$	Grid mask + $\sigma_n = 40$	Text mask + $\sigma_n = 50$
Input	9.42/0.068	9.95/0.074	10.34/0.078	12.29/0.206	14.15/0.187	13.09/0.148
ISDSB (He and Wang, 2014)	20.97/0.332	18.29/0.237	16.06/0.178	18.68/0.431	16.74/0.360	15.48/0.302
WLP (Zha et al., 2018)	20.27/0.318	17.65/0.229	15.47/0.172	18.81/0.372	16.55/0.287	15.01/0.241
GSR (Zhang et al., 2014b)	19.58/0.304	17.02/0.221	14.88/0.165	18.94/0.314	16.26/0.214	14.53/0.158
JSM (Zhang et al., 2014c)	20.54/0.329	17.74/0.235	15.36/0.170	18.94/0.313	16.27/0.215	14.80/0.167
RMC-NC (Nie et al., 2019)	21.82/0.572	20.22/0.528	18.66/0.473	20.46/0.594	20.52/0.575	18.56/0.468
STTC-A (Liu et al., 2019)	20.08/0.360	17.46/0.268	15.21/0.204	18.18/0.332	16.25/0.249	14.76/0.197
BPFA (Zhou et al., 2012)	26.40/0.691	25.78/0.660	25.24/0.634	23.81/0.577	25.59/0.675	23.97/0.575
JPG-SR (Zha et al., 2020a)	19.80/0.312	17.31/0.227	15.19/0.169	19.20/0.302	16.36/0.223	14.72/0.168
LR-GSC (Zha et al., 2021a)	19.96/0.315	17.44/0.282	15.20/0.211	19.14/0.304	16.53/0.272	14.50/0.202
HSSE (Zha et al., 2021b)	20.37/0.325	17.76/0.280	15.44/0.207	19.16/0.310	16.31/0.221	14.51/0.203
LGSR (Zha et al., 2022)	19.95/0.314	17.52/0.231	15.24/0.171	18.87/0.314	16.27/0.218	14.72/0.168
GSRC-NLP (Zha et al., 2020b)	20.66/0.333	17.75/0.235	15.16/0.167	18.88/0.314	16.25/0.214	14.63/0.165
IDBP (Tirer and Giryes, 2019)	26.41/0.695	25.85/0.670	25.32/0.649	23.89/0.623	25.47/0.680	24.14/0.608
DIP (Dmitry et al., 2020)	23.37/0.524	22.70/0.444	21.67/0.450	22.56/0.475	22.05/0.441	22.29/0.477
DD (Heckel and Hand, 2019)	22.74/0.534	23.98/0.469	22.92/0.479	23.77/0.515	23.66/0.477	23.58/0.469
NAS-DIP (Chen et al., 2020)	25.01/0.643	23.84/0.578	22.61/0.568	24.08/0.581	23.91/0.602	23.27/0.592
SHI (Shi et al., 2022)	24.72/0.639	24.09/0.616	23.31/0.589	23.80/0.630	24.48/0.638	23.45/0.600
Ours	<b>27.07/0.714</b>	<b>26.33/0.711</b>	<b>26.04/0.690</b>	<b>25.18/0.703</b>	<b>26.41/0.714</b>	<b>25.38/0.649</b>

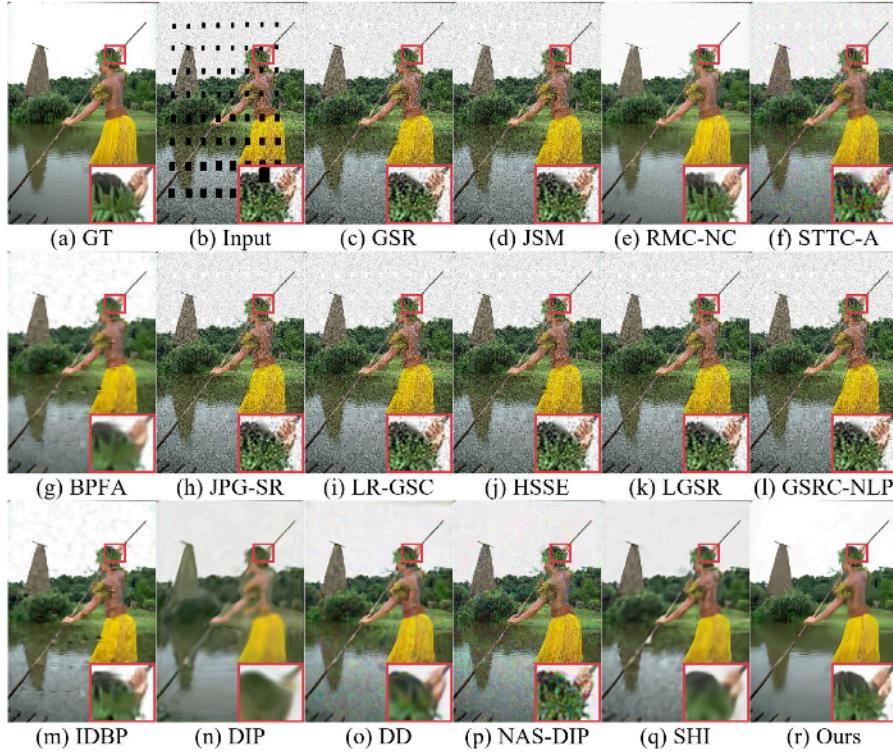
gains in average SSIM over 17 compared methods in order. The superior performance of our method stems from two aspects, the first is the introduction of ASR model, which can effectively remove noise and ensure local smoothness. The other is NLSM, which can recover the structure and texture of the whole image from the image patch perspective. We incorporate these two modules into the existing WSNM and combine the advantages of all three modules by constructing a unified optimization objective. This approach allows us to effectively achieve both denoising and completion simultaneously, resulting in satisfactory performance.

Besides, since handcrafted prior alone is insufficient to describe the complex image structures in challenging mixed degradation modes, advanced model-based methods such as JPG-SR (Zha et al., 2020a), LR-GSC (Zha et al., 2021a), HSSE (Zha et al., 2021b), LGSR (Zha et al., 2022), and GSRC-NLP (Zha et al., 2020b), which performed well on image completion without noise interference (or with very light noise), exhibit degraded performance under the mixed degradation mode (moderate or severe noise intervention). Based on the generalization ability of network, deep learning-based IDBP (Tirer and Giryes, 2019), DIP (Dmitry et al., 2020), NAS-DIP (Chen et al., 2020) and SHI (Shi et al., 2022) can achieve good results, but there is still a gap compared to the performance of our method.

**Qualitative Evaluation.** Fig. 8, Fig. 9, Figs. 10 and 11 depict the visual comparisons. Specifically, the results from GSR (Zhang et al., 2014b), JSM (Zhang et al., 2014c), and RMC-NC (Nie et al., 2019) still exhibit noticeable noise and visual artifacts. BPFA (Zhou et al., 2012) shows better noise removal by training a non-parametric Bayesian dictionary for simultaneous denoising and completion, but visible missing regions remain. STTC-A (Liu et al., 2019), LR-GSC (Zha et al., 2021a), and LGSR (Zha et al., 2022), which are based on the low-rank property of natural image matrices, also struggle to maintain the local smoothness and suppress noise. In contrast, we can see that the proposed model does better in eliminating noise compared to the above model-based methods. And our method is more effective in completing missing entries than JPG-SR (Zha et al., 2020a), GSRC-NLP (Zha et al., 2020b) and HSSE (Zha et al., 2021b), which are not efficient when there is much noise. When compared to advanced deep learning-based methods like IDBP (Tirer and Giryes, 2019), DIP (Dmitry et al., 2020), and SHI (Shi et al., 2022), the proposed model preserves more image details and achieves better local smoothness. Specifically, the local smoothness of IDBP (Tirer and Giryes, 2019) cannot be preserved and the textures in images completed by DIP (Dmitry et al., 2020), NAS-DIP (Chen et al., 2020) and SHI (Shi et al., 2022) are less clear. In contrast, the proposed method delivers satisfactory visual quality



**Fig. 8.** Image 2 corrupted by text mask covering with AWGN  $\sigma_n = 50$  from McMaster (Zhang et al., 2011) (a) GT, (b) Input. (c-r) Results by comparison methods.



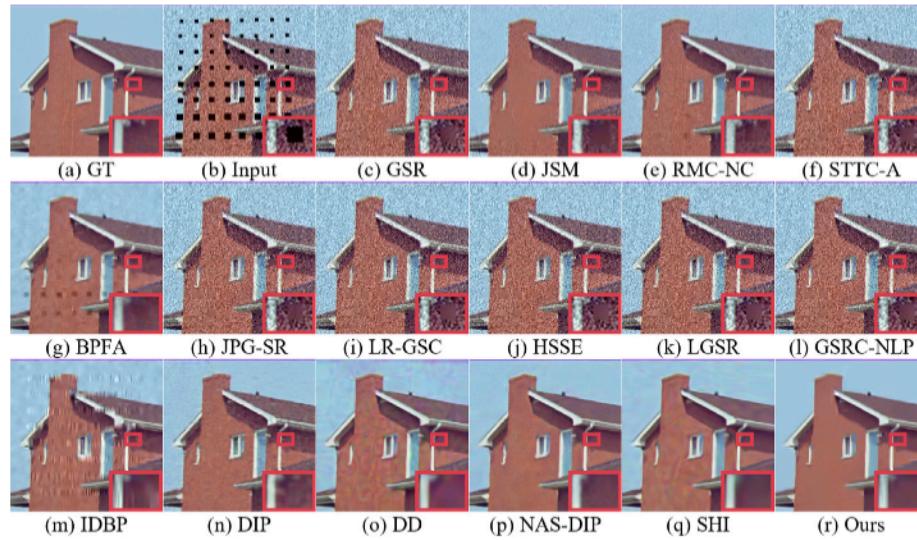
**Fig. 9.** Image 101087 corrupted by grid mask covering with AWGN  $\sigma_n = 40$  from CBSD68 (Martin et al., 2001). (a) GT, (b) Input. (c-r) Results by comparison methods.

and is much more effective at noise suppression than the competing methods.

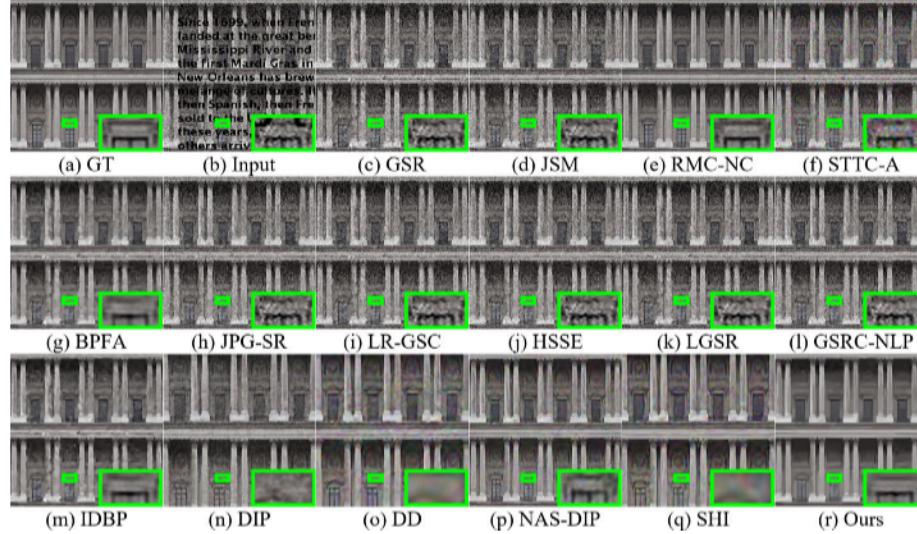
### 6.3. Selection of parameters

In our implementation, the patch size  $\sqrt{d} \times \sqrt{d}$  is set to  $8 \times 8$ . For the parameters in the  $Z$  sub-problem, following Zha et al. (2018), iteration number  $K$  is set as 2, and the power  $p$  is set to 0.95. The parameters  $m$ ,  $\rho$ , and  $\zeta$  are set to 60, 0.3, and 0.1, respectively. For the parameters in the  $W$  sub-problem,  $T$  and  $\tau$  are set to 5 and 1, respectively.  $\xi$  is initialized to 1 and increased by a factor of 1.05. For parameters in  $S$  sub-problem, trade off parameter  $\kappa$  is set to 10 according to Zhang et al. (2014c).

Moreover, in our method, we embed three regularization terms (namely WSNM, ASR and NLSM) into one unified framework, and effectively solved by ADMM (Boyd et al., 2011). Four hyper-parameters: step  $\rho$ , Lagrange multipliers  $\mu_1$ ,  $\mu_2$  and  $\mu_3$  are introduced correspondingly, which play important roles in the proposed method. In this section, we give a detailed discussion on selecting optimal  $\rho$ ,  $\mu_1$ ,  $\mu_2$  and  $\mu_3$ . Specifically, we use three common test images (*Butterfly*, *Girl* and *House*) and perform on the case of 50% random pixel missing with noise level  $\sigma_n = 30$ , with different step  $\rho$ , which ranges from 1 to 2 with interval 0.01. As depicted in Fig. 12.(a), we can see that the best image restoration performance is achieved with  $\rho$  in the range [1.01, 2], in our method  $\rho$  is empirically set to be 1.01. Based on the optimal  $\rho$ , we fix other factors,  $\mu_1$  varies from 0 to 1 with interval 0.1. Fig. 12.(b)



**Fig. 10.** Visual comparison of *House* corrupted by grid mask covering with AWGN  $\sigma_n = 30$ . (a) GT, (b) Input, (c-r) Results by comparison methods.



**Fig. 11.** Visual comparison of *re10* corrupted by text mask covering with AWGN  $\sigma_n = 30$  from Hu et al. (2013). (a) GT, (b) Input. (c-r) Results by comparison methods.

depicts that the best performance can be achieved with  $\mu_1$  in the range [0.1, 1],  $\mu_1$  is set to 0.1. Based on the same operation, we give the performances of  $\mu_2$  and  $\mu_3$  on the same test images, under different value ranges in Fig. 12.(c) and Fig. 12.(d), respectively. As marked by the black lines, the optimal  $\mu_2$  and  $\mu_3$  is set to 0.1 and 1.7, respectively. Hence, we fix  $\rho = 1.01$  and  $\mu_1 = 0.1$ ,  $\mu_2 = 0.1$ ,  $\mu_3 = 1.7$  in our method. The performance of the proposed model for different input images can reach a stable state and show a good robustness.

#### 6.4. Ablation study

Focusing on simultaneous image denoising and completion, we adopt three regularization terms (WSNM, ASR and NLSM) into one unified framework, and solved by ADMM. To prove the effectiveness of individual technical contributions, we conduct the ablation study. For the convenience of, we refer to WSNM, ASR, NLSM as R1, R2, R3, respectively. We sequentially add 3 regularization terms R1, R2,

and R3 to the proposed framework. We conduct experiments on 3 most commonly used benchmarks: Set12 (Zhang et al., 2018), Kodak (Franzen, 1999), and McMaster (Zhang et al., 2011), under 6 different degradation settings.

The quantitative results list in the Table 5, and visual qualitative results illustrated in Fig. 13, reveal the below: 1. Compared to the degraded input, When the restoration method based on R1 alone can improve the restoration performance to a certain extent. This indicates that the method based on low-rank restoration is effective for completing the missing entries. 2. When R1 and R2 are jointly employed, the performance, both quantitative and qualitative, surpasses that of model with individual employed R1. In our method, a convolutional implementation is adopted for the ASR regularization term (namely R2) which makes the problem much simpler with good performance for image smoothness. However, the performance is still fall short compared to the performance of our full model. 3. When WSNM, ASR and NLSM are combined together, i.e., R1+R2+R3, our full model

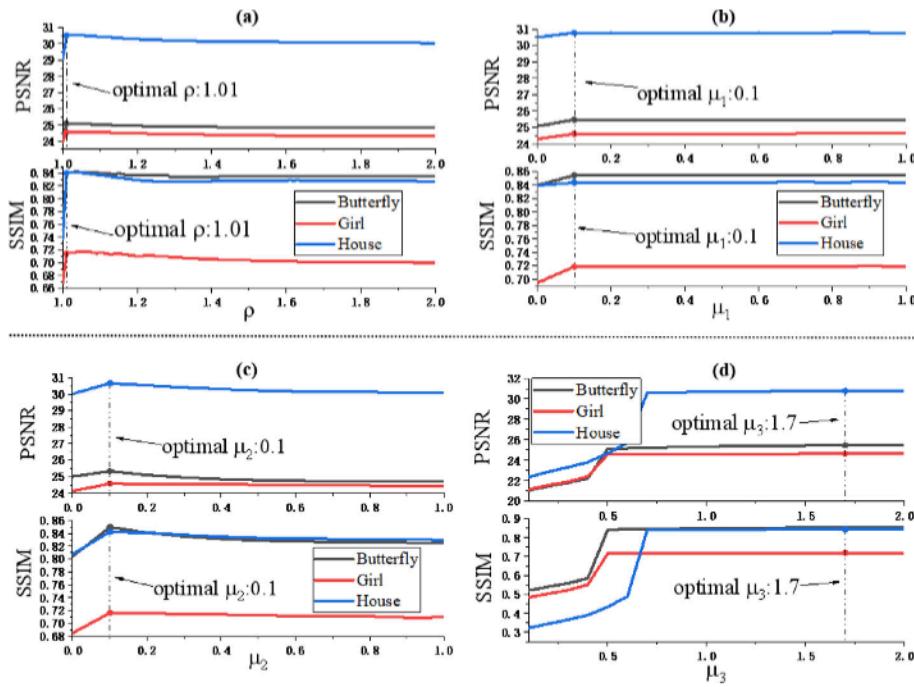
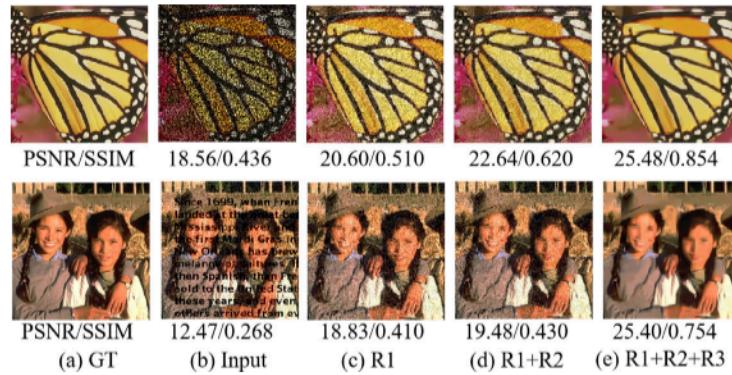
Fig. 12. PSNR and SSIM values with respect to parameters  $\rho$ ,  $\mu_1$ ,  $\mu_2$ ,  $\mu_3$ .Fig. 13. Visual comparison of adopting three different regularization terms (added one-by-one) for 50% random pixel missing with noise level  $\sigma_n = 30$  on *Butterfly* (top) and text mask covering with AWGN  $\sigma_n = 30$  on *Girl* (bottom).

Table 5

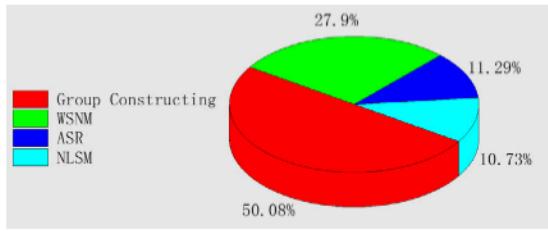
Ablation study. R1: WSNM, R2: ASR, R3: NLSM. Average PSNR/SSIM on Set12 (Zhang et al., 2018), Kodak (Franzen, 1999), and McMaster (Zhang et al., 2011), under six different settings. The best result is highlighted in **bold**.

Setting	Dataset	Input	R1	R1 + R2	R1+R2+R3 (Ours)
50% missing + $\sigma_n = 30$	Set12	8.28/0.127	19.10/0.488	21.04/0.536	<b>26.89/0.819</b>
	Kodak	9.43/0.167	19.82/0.492	21.83/0.541	<b>27.91/0.826</b>
	McMaster	10.28/0.218	20.36/0.512	22.42/0.562	<b>28.66/0.859</b>
40% missing + $\sigma_n = 40$	Set12	8.92/0.143	19.15/0.590	20.55/0.665	<b>26.22/0.795</b>
	Kodak	9.96/0.186	19.93/0.595	21.39/0.674	<b>27.30/0.802</b>
	McMaster	10.70/0.231	20.41/0.621	21.90/0.701	<b>27.95/0.838</b>
30% missing + $\sigma_n = 50$	Set12	9.49/0.154	20.37/0.635	21.86/0.706	<b>25.67/0.777</b>
	Kodak	10.35/0.197	21.25/0.637	22.81/0.718	<b>26.79/0.780</b>
	McMaster	10.93/0.236	21.71/0.668	23.03/0.753	<b>27.37/0.819</b>
Line mask + $\sigma_n = 30$	Set12	11.37/0.287	18.25/0.431	18.77/0.445	<b>24.67/0.787</b>
	Kodak	12.37/0.336	19.61/0.435	20.16/0.450	<b>26.51/0.796</b>
	McMaster	13.20/0.370	19.55/0.449	20.11/0.464	<b>26.43/0.821</b>
Grid mask + $\sigma_n = 40$	Set12	13.79/0.300	20.82/0.458	21.54/0.481	<b>26.41/0.811</b>
	Kodak	14.21/0.350	22.08/0.463	22.85/0.486	<b>28.02/0.820</b>
	McMaster	14.52/0.388	22.09/0.477	22.85/0.501	<b>28.02/0.845</b>
Text mask + $\sigma_n = 50$	Set12	10.86/0.182	17.05/0.347	17.76/0.385	<b>25.38/0.769</b>
	Kodak	11.55/0.209	17.84/0.348	18.57/0.387	<b>26.55/0.773</b>
	McMaster	11.95/0.247	18.11/0.364	18.86/0.405	<b>26.96/0.808</b>

**Table 6**

Time consumption of compared methods on three different image sizes.

Method	ISDSB	WLP	GSR	JSM	RMC-NC	STTC-A	BPFA	JPG-SR	LR-GSC
Time (s)	128 × 128	1.95	671.5	3.03	58.0	3.13	2.28	123.1	276.8
Method	HSSE	LGSR	GSRC-NLP	IDBP	DIP	DD	NAS-DIP	SHI	494.9 Ours
Time (s)	1378.0	456.1	651.1	10.55	80.33	5.08	11.4	653.4	32.15
Method	ISDSB	WLP	GSR	JSM	RMC-NC	STTC-A	BPFA	JPG-SR	LR-GSC
Time (s)	256 × 256	7.8	2797.9	12.1	241.7	12.5	9.1	512.5	1150.1
Method	HSSE	LGSR	GSRC-NLP	IDBP	DIP	DD	NAS-DIP	SHI	2061.9 Ours
Time (s)	5741.8	1900.2	2713.1	42.2	321.3	23.2	45.6	2722.5	128.6
Method	ISDSB	WLP	GSR	JSM	RMC-NC	STTC-A	BPFA	JPG-SR	LR-GSC
Time (s)	512 × 512	30.4	11471.4	47.2	991.0	48.8	35.5	2101.3	4715.4
Method	HSSE	LGSR	GSRC-NLP	IDBP	DIP	DD	NAS-DIP	SHI	8453.8 Ours
Time (s)	23540.2	7790.3	11123.3	164.6	1253.1	90.5	177.8	11162.3	501.5

**Fig. 14.** The proportion of time consumed by different modules of our method.

achieves the best performance both quantitatively (The last column of **Table 5**) and qualitatively (**Fig. 13(e)**). Specifically, NLSM is characterized by 3D transform domain, rather than by the traditional weighted graph. NLSM shows good performance on retaining the sharp edges and recovering the fine textures. In summary, the above results and analysis demonstrate the effectiveness of individual regularization term and are complementary to each other.

### 6.5. Computational time and numerical convergence

**Computation Efficiency.** We report the computing complexity of our method with other compared methods in this paper. Testing is conducted on images with sizes 128×128, 256×256, and 512 × 512, using an Intel (R) Core (TM) i9-9900K with 3.60GHZ CPU PC and a NVIDIA GeForce RTX 2070 GPU. All the compared deep learning-based algorithms are evaluated on the GPU, while the rest of competing methods are tested on the CPU. As listed in **Table 6**, the proposed method requires about 32 s, 130 s and 500 s for an image with size of 128 × 128, 256 × 256, and 512 × 512, respectively. Our method consumes less computational time than all competing methods except for ISDSB, RMC-NC, STTC-A, IDBP and DD. In particular, the proposed model requires only about 1/2 and 1/20 of the computational time compared to deep learning-based methods DIP and SHI, respectively, both of which are implemented on the GPU.

Briefly, in the proposed method, focusing on simultaneous image denoising and completion, we incorporate three regularization terms (WSNM, ASR and NLSM) into a unified framework and solved by ADMM. However, as mentioned above, both WSNM and NLSM implementations require searching for image patches and then constructing the corresponding image groups. Constructing the group involves dividing the image into  $n$  small patches. For each target patch,  $m$  similar patches within a certain neighborhood need to be searched by KNN to form a group. As shown in **Fig. 14**, the group construction process accounts for half of the total computational cost of our method. In the future, exploring more efficient searching algorithms than KNN and employing parallel patches extracting could be promising directions for accelerating the proposed algorithm to meet real-time requirements.

**Numerical Convergence.** We plot the curves of the PSNR/SSIM results versus the number of iterations for the proposed method applied to 50% random pixel missing with noise level  $\sigma_n = 30$ . As shown in **Fig. 15**, both PSNR and SSIM curves increase monotonically with the number of iteration and tend to flat. This indicates that the proposed method enjoys a good convergence property.

### 6.6. Exploring on challenging scenarios

**Hyper-spectral Image (HSI).** HSI is always corrupted or incomplete due to low light or sensors in real scenarios (Luo et al., 2022). Our method is evaluated on HSI completion and the result is presented in **Fig. 16**. The recovered image obtained by the proposed method **Fig. 16.(r)** can be much cleaner and fine textures and image details are preserved more accurately, as depicted in the enlarged part.

**Remote Sensing Image (RSI).** Due to water absorption and atmosphere flow, noise and stripes always coexist in RSI (Xia et al., 2017). As shown in **Fig. 17**, compared with other methods, our algorithm can effectively remove the noise and recover the information obscured by the stripes.

**Historic Mural.** We conduct experiments on semantic image completion task. The recovered results of historical mural (Yu et al., 2019) are shown in **Fig. 18**. We can observe that our method can effectively recover the missing information located in destroyed regions as shown in the close-up **Fig. 18.(r)** compared to other state-of-the-art methods. It demonstrates the effectiveness of the proposed model in handling complicated scenes in the real world. Overall, the results presented in above validate the effectiveness of our method in coping with challenging mixed degradation scenarios.

## 7. Conclusion

In this paper, we propose a new LRMA algorithm for high-quality simultaneous image denoising and completion, leveraging convolutional sparse representation and nonlocal self-similarity. This algorithm effectively utilizes the local smoothness, nonlocal self-similarity and low-rank characteristics of natural images. Additionally, we use ADMM to obtain an efficient solution for the proposed method. We perform two experiments on two kinds of degraded images: image completion for partial random samples and mask removal with different noise levels. Extensive experimental results demonstrate that the proposed method outperforms the comparison methods in both visual and quantitative measures. Future research will focus on developing a framework for image restoration that is both comprehensive and powerful. Specifically, image restoration encompasses image deblurring, completion, super-resolution, in accordance with the form of the degradation operator  $H$  in Eq. (1). To handle more difficult mixed degradation scenes, such as simultaneous image deblurring and denoising, simultaneous image super-resolution and denoising, etc. Moreover, the applications of high-dimensional data restoration tasks, such as video and medical imaging (Wang et al., 2024, 2023) are also explored in the future.

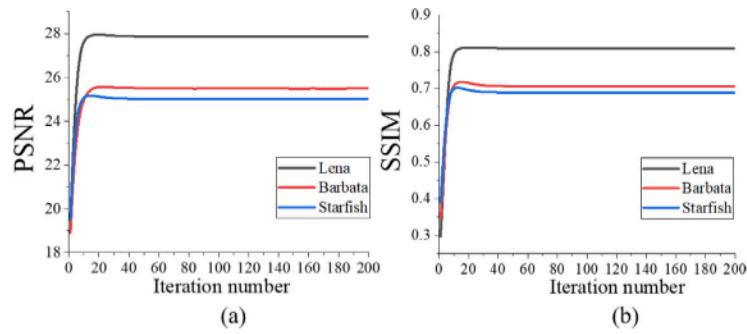


Fig. 15. Convergence behavior of the proposed method. PSNR/SSIM versus iteration.

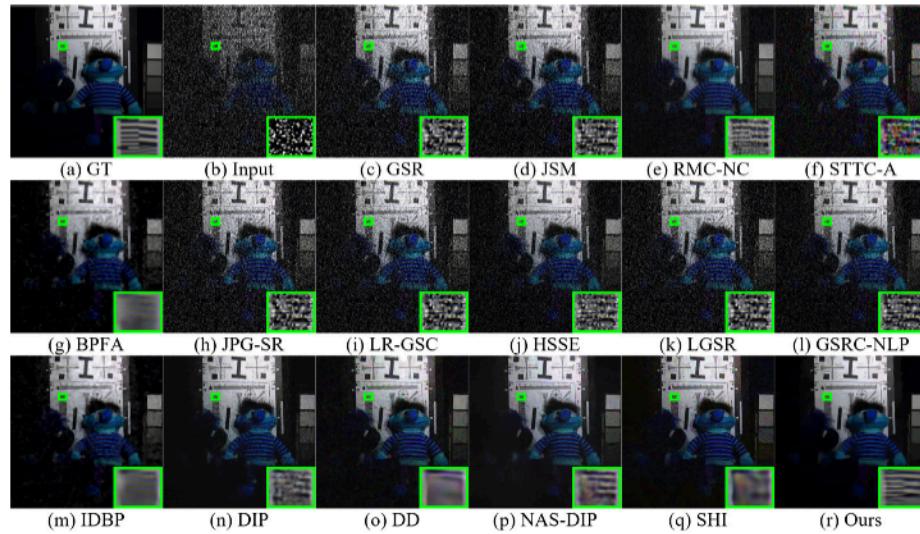


Fig. 16. Visual comparison of HSI Toy (composed of 10th, 20th and 30th bands) from Luo et al. (2022) corrupted by 50% random pixels missing with AWGN  $\sigma_n = 30$ . (a) GT, (b) Input. (c-r) Results by comparison methods.

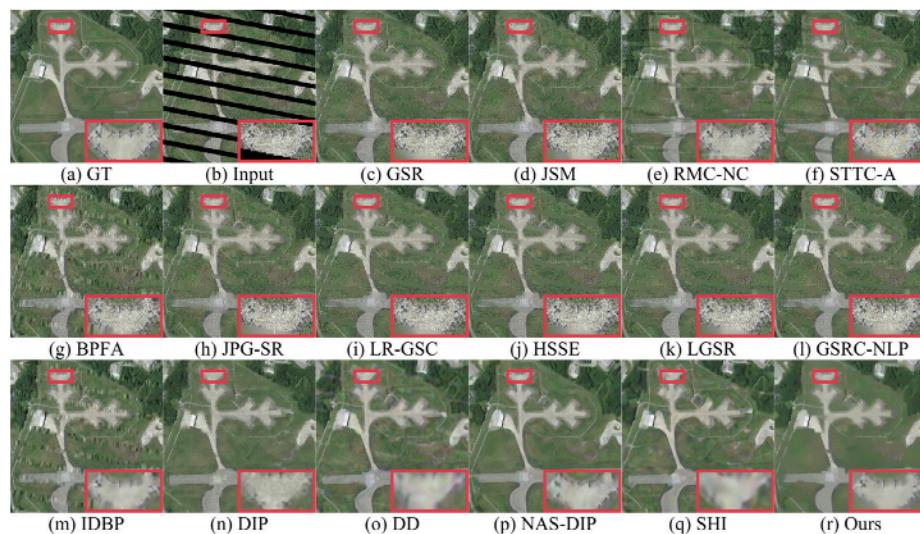
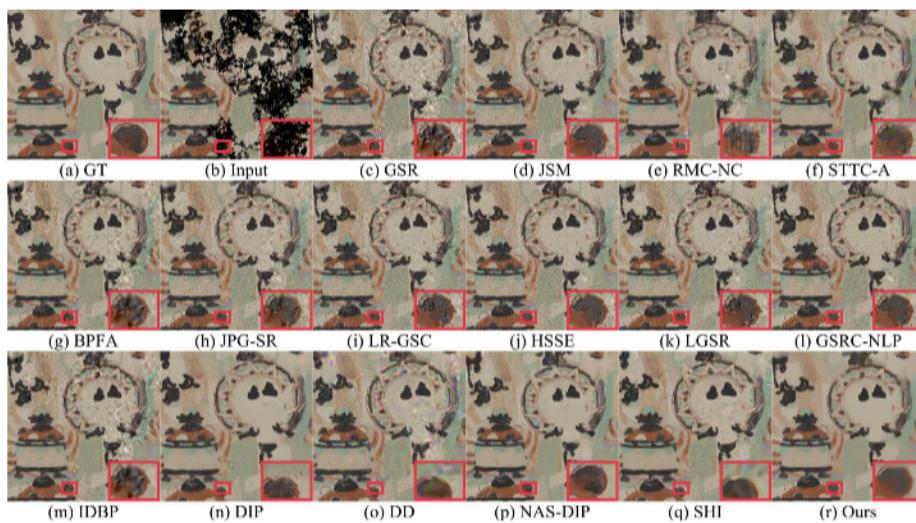


Fig. 17. Visual comparison of RSI Airport97 from AID (Xia et al., 2017) corrupted by stripe mask covering with noise level  $\sigma_n = 30$ . (a) GT, (b) Input. (c-r) Results by comparison methods.



**Fig. 18.** Visual comparison of historic mural 384 from Dunhuang (Yu et al., 2019) corrupted by irregular mask covering with AWGN  $\sigma_n = 30$ . (a) GT, (b) Input. (c-r) Results by comparison methods.

#### CRediT authorship contribution statement

Weimin Yuan: Methodology. Yuanyuan Wang: Methodology. Ruirui Fan: Methodology. Yuxuan Zhang: Methodology. Guangmei Wei: Investigation. Cai Meng: Supervision. Xiangzhi Bai: Writing – review & editing, Supervision.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

The authors would like to thank the editors and the anonymous reviewers for their constructive comments and suggestions. A preliminary version of this paper was presented in Fan et al. (2019). Compared with the conference version, we have made significant improvements. Specifically, we have provided a detailed introduction to related works and preliminaries. We have also emphasized the research motivation and offered a more comprehensive discussion of the algorithm details. We have added and compared several methods proposed in recent years. The superiority of our proposed method has been thoroughly validated on four newly added datasets. Moreover, we have demonstrated the algorithm's effectiveness and versatility by presenting recovered results from historical murals, hyperspectral images, and remote sensing images. To prove the effectiveness of each module of the algorithm, we have conducted a comprehensive ablation study. The time complexity and convergence performance of the algorithm, as well as parameter sensitivity analysis, have also been extensively discussed. This work was supported in part by the National Natural Science Foundation of China (Grants Nos. 62271016, 92148206) and the Beijing Natural Science Foundation under Grant 4222007.

#### Data availability

Data will be made available on request.

#### References

- Bhavsar, A.V., Rajagopalan, A.N., 2012. Range map superresolution-inpainting, and reconstruction from sparse data. *Comput. Vis. Image Underst.* 116 (4), 572–591.
- Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* 3 (1), 1–122.
- Cai, J.-F., Candès, E.J., Shen, Z., 2010. A singular value thresholding algorithm for matrix completion. *SIAM J. Optim.* 20, 1956–1982.
- Chen, Y.-C., Gao, C., Robb, E., Huang, J.-B., 2020. NAS-dip: Learning deep image prior with neural architecture search. In: Proc. Eur. Conf. Comput. Vis.. ECCV, pp. 442–459.
- Dmitry, U., Andrea, V., Victor, L., 2020. Deep image prior. *Int. J. Comput. Vis.* 128 (7), 1867–1888.
- Dong, W., Shi, G., Li, X., 2013. Nonlocal image restoration with bilateral variance estimation: A low-rank approach. *IEEE Trans. Image Process.* 22 (2), 700–711.
- Fan, R., Wei, G., Zhang, Y., Bai, X., 2019. Weighted schatten P-norm minimization with local and nonlocal constraints for noisy image completion. In: 2019 IEEE International Conference on Image Processing. ICIP, pp. 2746–2750.
- Fazel, M., 2002. Matrix rank minimization with applications. Dissertation 54, 1–130.
- Franzen, R., 1999. Kodak lossless true color image suite, vol. 4.
- Gu, S., Xie, Q., Meng, D., Zuo, W., Feng, X., Zhang, L., 2017. Weighted nuclear norm minimization and its applications to low level vision. *Int. J. Comput. Vis.* 121, 183–208.
- Gu, S., Zhang, L., Zuo, W., Feng, X., 2014. Weighted nuclear norm minimization with application to image denoising. In: Proc. IEEE Comput. Vis. Pattern Recog.. CVPR, pp. 2862–2869.
- Hawe, S., Kleinstuber, M., Diepold, K., 2013. Analysis operator learning and its application to image reconstruction. *IEEE Trans. Image Process.* 22 (6), 2138–2150.
- He, J., Dong, C., Liu, Y., Qiao, Y., 2022. Interactive multi-dimension modulation for image restoration. *IEEE Trans. Pattern Anal. Mach. Intell.* 44 (12), 9363–9379.
- He, L., Wang, Y., 2014. Iterative support detection-based split bregman method for wavelet frame-based image inpainting. *IEEE Trans. Image Process.* 23 (12), 5470–5485.
- Heckel, R., Hand, P., 2019. Deep decoder: Concise image representations from untrained non-convolutional networks. In: Proc. Int. Conf. Learn. Represent.. ICLR.
- Hu, Y., Zhang, D., Ye, J., Li, X., He, X., 2013. Fast and accurate matrix completion via truncated nuclear norm regularization. *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (9), 2117–2130.
- Jam, J., Kendrick, C., Walker, K., Drouard, V., Hsu, J.G.-S., Yap, M.H., 2021. A comprehensive review of past and present image inpainting methods. *Comput. Vis. Image Underst.* 203, 103147.
- Keller, J.M., Gray, M.R., Givens, J.A., 1985. A fuzzy K-nearest neighbor algorithm. *IEEE Trans. Syst. Man Cybern.* (4), 580–585.
- Li, X., Jin, X., Lin, J., Liu, S., Wu, Y., Yu, T., Zhou, W., Chen, Z., 2020. Learning disentangled feature representation for hybrid-distorted image restoration. In: Eur. Conf. Comput. Vis.. ECCV, pp. 313–329.
- Li, X., Zhang, H., Zhang, R., 2022. Matrix completion via non-convex relaxation and adaptive correlation learning. *IEEE Trans. Pattern Anal. Mach. Intell.* 1.
- Liu, Y., Long, Z., Zhu, C., 2019. Image completion using low tensor tree rank and total variation minimization. *IEEE Trans. Multimedia* 21 (2), 338–350.

## ARTICLE IN PRESS

- Luo, Y.-S., Zhao, X.-L., Jiang, T.-X., Chang, Y., Ng, M.K., Li, C., 2022. Self-supervised nonlinear transform-based tensor nuclear norm for multi-dimensional image recovery. *IEEE Trans. Image Process.* 31, 3793–3808.
- Martin, D., Fowlkes, C., Tal, D., Malik, J., 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: *IEEE Int. Conf. Comput. Vision.. ICCV*, IEEE, pp. 416–423.
- Nie, F., Hu, Z., Li, X., 2019. Matrix completion based on non-convex low-rank approximation. *IEEE Trans. Image Process.* 28 (5), 2378–2388.
- Nie, F., Huang, H., Ding, C., 2012. Low-rank matrix recovery via efficient schatten p-norm minimization. In: *Proc. AAAI Conf. Artif. Intell.. AAAI*, Vol. 1, pp. 655–661.
- Qin, J., Bai, H., Zhao, Y., 2021. Multi-scale attention network for image inpainting. *Comput. Vis. Image Underst.* 204, 103155.
- Sandbichler, M., Schnass, K., 2019. Online and stable learning of analysis operators. *IEEE Trans. Signal Process.* 67 (1), 41–53.
- Shi, Z., Pascal, M., Subhransu, S., 2022. On measuring and controlling the spectral bias of the deep image prior. *Int. J. Comput. Vis.* 130 (4), 885–908.
- Shin, W., Ahn, N., Moon, J.-H., Sohn, K.-A., 2022. Exploiting distortion information for multi-degraded image restoration. In: *IEEE Comput. Vis. Pattern Recognit. Workshops. CVPRW*, pp. 536–545.
- Suganuma, M., Liu, X., Okatani, T., 2019. Attention-based adaptive selection of operations for image restoration in the presence of unknown combined distortions. In: *IEEE Comput. Vis. Pattern Recognit.. CVPR*, pp. 9031–9040.
- Tirer, T., Giryes, R., 2019. Image restoration by iterative denoising and backward projections. *IEEE Trans. Image Process.* 28 (3), 1220–1234.
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13 (4), 600–612.
- Wang, Y., Chen, K., Yuan, W., Tang, Z., Meng, C., Bai, X., 2024. SAMIHS: adaptation of segment anything model for intracranial hemorrhage segmentation. In: *2024 IEEE International Symposium on Biomedical Imaging. ISBI*, IEEE, pp. 1–5.
- Wang, Y., Yuan, W., Bai, X., 2023. CoAM-Net: coordinate asymmetric multi-scale fusion strategy for polyp segmentation. *Appl. Intell.* 53 (24), 30626–30641.
- Wei, K., Aviles-Rivero, A., Liang, J., Fu, Y., Schönlieb, C.-B., Huang, H., 2020. Tuning-free plug-and-play proximal algorithm for inverse imaging problems. In: *Proceedings of the 37th International Conference on Machine Learning. JMLR.org*.
- Xia, G.-S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., Zhang, L., Lu, X., 2017. AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Trans. Geosci. Remote Sens.* 55 (7), 3965–3981.
- Xie, Y., Gu, S., Liu, Y., Zuo, W., Zhang, W., Zhang, L., 2016. Weighted schatten  $p$ -norm minimization for image denoising and background subtraction. *IEEE Trans. Image Process.* 25 (10), 4842–4857.
- Xie, M., Liu, X., Yang, X., 2024. A nonlocal self-similarity-based weighted tensor low-rank decomposition for multichannel image completion with mixture noise. *IEEE Trans. Neural Netw. Learn. Syst.* 35 (1), 73–87.
- Yu, K., Dong, C., Lin, L., Loy, C.C., 2018. Crafting a toolchain for image restoration by deep reinforcement learning. In: *IEEE Comput. Vis. Pattern Recognit.. CVPR*, pp. 2443–2452.
- Yu, B., Liu, Y., Sun, Q., 2017. A content-adaptively sparse reconstruction method for abnormal events detection with low-rank property. *IEEE Trans. Syst. Man Cybern. Syst.* 47 (4), 704–716.
- Yu, T., Zhang, S., Lin, C., You, S., Wu, J., Zhang, J., Ding, X., An, H., 2019. Dunhuang grottoes painting dataset and benchmark. *arXiv preprint arXiv:1907.04589*.
- Yuan, W., Meng, C., Bai, X., 2024a. Weighted side-window based gradient guided image filtering. *Pattern Recognit.* 146, 110006.
- Yuan, W., Meng, C., Tong, X., Li, Z., 2021. Efficient local stereo matching algorithm based on fast gradient domain guided image filtering. *Signal Process., Image Commun.* 95, 116280.
- Yuan, W., Wang, Y., Li, N., Meng, C., Bai, X., 2024b. Mixed degradation image restoration via deep image prior empowered by deep denoising engine. In: *2024 International Joint Conference on Neural Networks. IJCNN*, IEEE, pp. 1–8.
- Zha, Z., Wen, B., Yuan, X., Zhou, J., Zhu, C., 2021a. Image restoration via reconciliation of group sparsity and low-rank models. *IEEE Trans. Image Process.* 30, 5223–5238.
- Zha, Z., Wen, B., Yuan, X., Zhou, J., Zhu, C., Kot, A.C., 2021b. A hybrid structural sparsification error model for image restoration. *IEEE Trans. Neural Netw. Learn. Syst.* 1–15.
- Zha, Z., Wen, B., Yuan, X., Zhou, J., Zhu, C., Kot, A.C., 2022. Low-rankness guided group sparse representation for image restoration. *IEEE Trans. Neural Netw. Learn. Syst.* 1–15.
- Zha, Z., Yuan, X., Wen, B., Zhang, J., Zhou, J., Zhu, C., 2020a. Image restoration using joint patch-group-based sparse representation. *IEEE Trans. Image Process.* 29, 7735–7750.
- Zha, Z., Yuan, X., Wen, B., Zhou, J., Zhu, C., 2020b. Group sparsity residual constraint with non-local priors for image restoration. *IEEE Trans. Image Process.* 29, 8960–8975.
- Zha, Z., Zhang, X., Wu, Y., Wang, Q., Tang, L., Yuan, X., 2018. Non-convex weighted  $l_p$  nuclear norm based ADMM framework for image restoration. *Neurocomputing* 311, 209–224.
- Zhang, Z., Ely, G., Aeron, S., Hao, N., Kilmer, M., 2014a. Novel methods for multilinear data completion and de-noising based on tensor-SVD. In: *Proc. IEEE Comput. Vis. Pattern Recog.. CVPR*, pp. 3842–3849.
- Zhang, H., Gao, J., Qian, J., Yang, J., Xu, C., Zhang, B., 2024a. Linear regression problem relaxations solved by nonconvex ADMM with convergence analysis. *IEEE Trans. Circuits Syst. Video Technol.* 34 (2), 828–838.
- Zhang, H., Gong, C., Qian, J., Zhang, B., Xu, C., Yang, J., 2019. Efficient recovery of low-rank matrix via double nonconvex nonsmooth rank minimization. *IEEE Trans. Neural Netw. Learn. Syst.* 30 (10), 2916–2925.
- Zhang, H., Li, S., Qiu, J., Tang, Y., Wen, J., Zha, Z., Wen, B., 2023a. Efficient and effective nonconvex low-rank subspace clustering via SVT-free operators. *IEEE Trans. Circuits Syst. Video Technol.* 33 (12), 7515–7529.
- Zhang, K., Li, Y., Zuo, W., Zhang, L., Van Gool, L., Timofte, R., 2022. Plug-and-play image restoration with deep denoiser prior. *IEEE Trans. Pattern Anal. Mach. Intell.* 44 (10), 6360–6376.
- Zhang, H., Qian, F., Shi, P., Du, W., Tang, Y., Qian, J., Gong, C., Yang, J., 2023b. Generalized nonconvex nonsmooth low-rank matrix recovery framework with feasible algorithm designs and convergence analysis. *IEEE Trans. Neural Netw. Learn. Syst.* 34 (9), 5342–5353.
- Zhang, H., Qian, J., Zhang, B., Yang, J., Gong, C., Wei, Y., 2020. Low-rank matrix recovery via modified schatten- $p$  norm minimization with convergence guarantees. *IEEE Trans. Image Process.* 29, 3132–3142.
- Zhang, L., Wu, X., Buades, A., Li, X., 2011. Color demosaicking by local directional interpolation and nonlocal adaptive thresholding. *J. Electron. Imaging* 20 (2).
- Zhang, H., Yang, J., Qian, J., Gao, G., Lan, X., Zha, Z., Wen, B., 2024b. Efficient image classification via structured low-rank matrix factorization regression. *IEEE Trans. Inf. Forensics Secur.* 19, 1496–1509.
- Zhang, J., Zhao, D., Gao, W., 2014b. Group-based sparse representation for image restoration. *IEEE Trans. Image Process.* 23 (8), 3336–3351.
- Zhang, J., Zhao, D., Xiong, R., Ma, S., Gao, W., 2014c. Image restoration using joint statistical modeling in a space-transform domain. *IEEE Trans. Circuits Syst. Video Technol.* 24 (6), 915–928.
- Zhang, H., Zhao, J., Zhang, B., Gong, C., Qian, J., Yang, J., 2024c. Unified framework for faster clustering via joint schatten  $p$ -norm factorization with optimal mean. *IEEE Trans. Neural Netw. Learn. Syst.* 35 (3), 3012–3026.
- Zhang, L., Zuo, W., 2017. Image restoration: From sparse and low-rank priors to deep priors. *IEEE Signal Process. Mag.* 34 (5), 172–179.
- Zhang, K., Zuo, W., Gu, S., Zhang, L., 2017. Learning deep CNN denoiser prior for image restoration. In: *IEEE Conf. Comput. Vis. Pattern Recognit.. CVPR*, pp. 2808–2817.
- Zhang, K., Zuo, W., Zhang, L., 2018. FFDNet: Toward a fast and flexible solution for CNN-based image denoising. *IEEE Trans. Image Process.* 27 (9), 4608–4622.
- Zhou, M., Chen, H., Paisley, J., 2012. Nonparametric Bayesian dictionary learning for analysis of noisy and incomplete images. *IEEE Trans. Image Process.* 21 (1), 130–144.
- Zuo, W., Meng, D., Zhang, L., 2013. A generalized iterated shrinkage algorithm for non-convex sparse coding. In: *Proc. IEEE Int. Conf. Comput. Vision.. ICCV*, pp. 217–224.