Pattern Recognition 125 (2022) 108537

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/patcog

Phase retrieval from incomplete data via weighted nuclear norm minimization

Zhi Li[a], Ming Yan[b], Tieyong Zeng[c,*], Guixu Zhang[a,*]

[a] Department of Computer Science and Technology, Shanghai Key Laboratory of Multidimensional Information Processing, East China Normal University, Shanghai 200241, China

[b] Department of Computational Mathematics, Science and Engineering and Department of Mathematics, Michigan State University, East Lansing, MI, 48824, USA

[c] Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong

ARTICLE INFO

Article history:
Received 14 November 2020
Revised 1 November 2021
Accepted 12 January 2022
Available online 14 January 2022

Keywords:
Phase retrieval
Partial magnitudes
Nuclear norm minimization
Impulse noise

ABSTRACT

Recovering an unknown object from the magnitude of its Fourier transform is a phase retrieval problem. Here, we consider a much difficult case, where those observed intensity values are incomplete and contaminated by both salt-and-pepper and random-valued impulse noise. To take advantage of the low-rank property within the image of the object, we use a regularization term which penalizes high weighted nuclear norm values of image patch groups. For outliers (impulse noise) in the observation, the ℓ_{1-2} metric is adopted as the data fidelity term. Then we break down the resulting optimization problem into smaller ones, for example, weighted nuclear norm proximal mapping and ℓ_{1-2} minimization, because the nonconvex and nonsmooth subproblems have available closed-form solutions. The convergence results are also presented, and numerical experiments are provided to demonstrate the superior reconstruction quality of the proposed method.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

In this paper, we consider the phase retrieval problem in which we try to recover the target image from the intensity measurements of its Fourier transform.

Many fields, such as astronomical imaging [1], crystallography [2], and optical imaging [3], etc., benefit from the development of phase retrieval algorithms. The error reduction algorithm [4], as a popular method for phase retrieval, works on both the spatial and Fourier domains and can be seen as an alternating projection method. Fienup pointed out that the error reduction algorithm may fail to find the optimal solution of the original problem [5], he improved it and proposed a basic input-output algorithm, which can be linked to the Dykstra's algorithm [6], and a hybrid input-output algorithm [5], which is an application of the Douglas-Rachford algorithm [6]. Projection-based approaches for phase retrieval also include the hybrid projection-reflection algorithm [7], the relaxed averaged alternation reflections method [8], and the augmented Lagrangian alternating direction method [9].

Due to the difficulty of non-convexity of the phase retrieval problem, semi-definite relaxation methods are used to transfer the non-convex problem into convex ones at the expense of squaring the number of variables [10]. Typical methods include PhaseLift [11] and PhaseCut [12]. PhaseLift optimizes a cost function defined on a convex set of complex Hermitian positive semidefinite matrices, and relaxes the rank minimization problem, while PhaseCut also drops the rank constraint. More recently, PhaseMax relaxes the non-convex equality constraint to an inequality constraint and changes the problem into a convex one without lifting [10]. However, in some situations, its numerical results are not as good as its counterparts with lifting [13]. To handle noisy measurements, Chang et al. [14] proposed to use the total variation regularization together with structured illuminated patterns in holography to build a model, later he applied a global convergent algorithm to solve phase retrieval problems with observed measurements polluted by Poisson or Gaussian noise [15].

1.1. Phase retrieval problem

Our goal is to recover the image $\mathbf{u}(t_1, t_2) : \mathbf{\Omega} \to \mathbb{R}^{n_1 \times n_2}$ defined on the lattice $\mathbf{\Omega}$ from its incomplete noisy phaseless measurements, where $t_1 \in \{0, 1, \ldots, n_1 - 1\}$, $t_2 \in \{0, 1, \ldots, n_2 - 1\}$. These measurements are obtained by illuminating the object with three

* Corresponding authors.
E-mail addresses: zeng@math.cuhk.edu.hk (T. Zeng), gxzhang@cs.ecnu.edu.cn (G. Zhang).

https://doi.org/10.1016/j.patcog.2022.108537
0031-3203/© 2022 Elsevier Ltd. All rights reserved.

light fields [16], and the resulting $3n_1n_2$ measurements have the form

$$\{\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2\} := \boldsymbol{J}\{|\boldsymbol{F}\mathbf{u}|, |\boldsymbol{F}(\mathbf{u} + \mathbf{D}^{\mathbf{s}}\mathbf{u})|, |\boldsymbol{F}(\mathbf{u} - i\mathbf{D}^{\mathbf{s}}\mathbf{u})|\}, \tag{1}$$

where $\boldsymbol{F}$ is the Fourier transform given by

$$(\boldsymbol{F}\mathbf{u})(\omega_1, \omega_2) := \frac{1}{\sqrt{n_1 n_2}} \sum_{(t_1,t_2)\in\boldsymbol{\Omega}} \mathbf{u}(t_1, t_2) \exp\left(-2\pi i\left(\frac{\omega_1 t_1}{n_1} + \frac{\omega_2 t_2}{n_2}\right)\right),$$

with $\boldsymbol{\omega} = (\omega_1, \omega_2) \in \boldsymbol{\Omega}$, $\mathbf{D}^{\mathbf{s}}$ is the modulation defined as

$$(\mathbf{D}^{\mathbf{s}}\mathbf{u})(t_1, t_2) = \exp\left(2\pi i\left(\frac{s_1 t_1}{n_1} + \frac{s_2 t_2}{n_2}\right)\right)\mathbf{u}(t_1, t_2), \tag{2}$$

with $\mathbf{s} = (s_1, s_2) = (0.5, 0.5)$ to yield an exact recovery [14], the $|*|$ is to take the element-wise absolute value of the matrix (the result of the Fourier transform), similarly, + and - are the element-wise addition and subtraction. $\boldsymbol{J}$ acts on all three magnitudes of the measurement in the curly brackets separately to apply both data loss and noise, details are given later in (4) in Section 2.

1.2. Weighted nuclear norm

To build the image recovery model and reconstruct $\mathbf{u}$, we need to adopt a properly chosen regularization term that reflects sparse properties of the underlying solutions. The low-rank approximation [17] is a fundamental tool in image processing, and weights [18,19] are incorporated into the sparsity promoting term to treat components adaptively. Sparse coding is a powerful strategy that is commonly implemented in a group-based way [20–22], which groups similar image patches together, then enforce the low-rank property over each group, the application includes, denoising [21,22], inpainting and deblocking[20], etc.

Sparse coding can be achieved by minimizing the weighted nuclear norm (WNN), which can alleviate the bias of the nuclear norm. The WNN of a matrix $\mathbf{X}$ is defined as

$$\|\mathbf{X}\|_{\mathbf{w},*} = \sum_i w_i \sigma_i(\mathbf{X}), \tag{3}$$

where weights $\{w_i\}_{i=1}^m$ are non-negative, $\sigma_i(\mathbf{X})$ is the i^{th} singular value of matrix $\mathbf{X}$, in this paper, $\mathbf{X}$ is a patch group (see Section 3.3) in an image. For more details, please check [21,23], the WNN has attracted researchers' attention recently, especially in image processing [24–28]. In [21,25,27], authors have demonstrated that WNN models can better recover images polluted by Gaussian noise, and can outperform many state-of-the-art methods in terms of the visual quality. WNN models are also designed to handle other image recovery problems such as, Cauchy noise removal [26] and deblurring [28], and for recovery of special types of images, such as hyperspectral images [24]. Considering the fact that the weighted ℓ_1 norm can enhance sparsity [23], Gu et al. [21] suggested to choose the weights based on the singular values, which works better than the nuclear norm [21,29]. They also provide a closed-form shrinkage operator to obtain a low-rank approximation of a matrix. However, the nonconvex penalty function corresponding to the approximate solution is not given.

Moreover, the WNN is implemented patch-wisely, and in a group-based manner, that is to say, similar patches are grouped together as a matrix, on which the WNN is computed. As far as we know, for those group-based methods, there was no convergence proof. In summary, the main contributions of this paper are listed as follows:

- We propose a new scheme for solving the phase retrieval problem with noisy measurements using group sparse representation via weighted nuclear norm minimization.
- We suggest to use the ℓ_{1-2} metric fidelity term for reconstruction of images form incomplete magnitude measurements corrupted by both salt-and-pepper and random-valued impulse noise.
- We present closed-form solutions for sub-problems after decomposing the main optimization problem. We give the analytical justification of the closed-form solution for the weighted nuclear norm proximal mapping, so that we can prove the convergence within the group-based sparse representation framework.

Organization: This paper is structured as follows: In Section 2, we introduce the proposed model. Then in Section 3, we revisit the approximate solution of the weighted nuclear norm proximal mapping suggested in [21], and give the analytical justification of the closed-form solution, so that, in the group-based sparse representation framework, we can provide the convergence proof of the penalty method which we adopt to solve the model. Numerical experiments are conducted on natural images to demonstrate the superior performance of the proposed model in Section 4. In the conclusion section, we also discuss future works.

2. The proposed model

The observed measurements $\mathbf{b} := \{\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2\}$ in (1) can be incomplete, e.g., from the perspective of compressive sensing, one can reconstruct the object from just its sub-sampled diffraction pattern [30–32]. We introduce three subsets: $\boldsymbol{\Omega}_i \subset \boldsymbol{\Omega}$, for $i \in \{0, 1, 2\}$ to define the observed measurements. We assume that $\mathbf{b}_i(j,k)$ is given when $(j,k) \in \boldsymbol{\Omega}_i$. In this paper, we consider the situation when the pollution is caused by both salt-and-pepper and random-valued impulse noise [33]. We assume that these two types of contamination are mutually exclusive, since salt-and-pepper noise overshadows other noise. We assume that the noise ratios of salt-and-pepper and random-valued impulse noise are r_1 and r_2, respectively, and $r_1 + r_2 \in (0, 1)$. Similar to the definition of the noise in [33], the intensity value of the corrupted measurement $\mathbf{b}_i(j,k)$, at location $(j,k) \in \boldsymbol{\Omega}_i$ is given by

$$\mathbf{b}_i(j,k) = \begin{cases} \max(\mathbf{b}_i) & \text{with probability } r_1/2; \\ \min(\mathbf{b}_i) & \text{with probability } r_1/2; \\ \boldsymbol{\eta}_i(j,k) & \text{with probability } r_2; \\ \mathbf{b}_i(j,k) & \text{with probability } 1 - r_1 - r_2, \end{cases} \tag{4}$$

where $\boldsymbol{\eta}_i(j,k)$ is a uniformly-distributed random value in $[\min(\mathbf{b}_i), \max(\mathbf{b}_i)]$. Note that the direct detection of noisy measurement is extremely difficult, especially for the random-valued noise, because observed measurements suffer from a major data loss at random places. Therefore, we propose to use a data loss term with the ℓ_{1-2} minimization [34,35]. The loss function is

$$\begin{aligned} &\||\boldsymbol{F}\mathbf{u}| - \mathbf{b}_0\|_{\ell_{1-2},\boldsymbol{\Omega}_0} + \||\boldsymbol{F}(\mathbf{u} + \mathbf{D}^{\mathbf{s}}\mathbf{u})| - \mathbf{b}_1\|_{\ell_{1-2},\boldsymbol{\Omega}_1} \\ &\quad + \||\boldsymbol{F}(\mathbf{u} - i\mathbf{D}^{\mathbf{s}}\mathbf{u})| - \mathbf{b}_2\|_{\ell_{1-2},\boldsymbol{\Omega}_2}, \end{aligned} \tag{5}$$

where $\|\mathbf{v}\|_{\ell_{1-2},\boldsymbol{\Omega}_i} := \sum_{j\in\boldsymbol{\Omega}_i} |\mathbf{v}(j)| - \alpha\big(\sum_{j\in\boldsymbol{\Omega}_i} |\mathbf{v}(j)|^2\big)^{\frac{1}{2}}$. We fix $\alpha = 0.5$ in this paper [34].

For simplicity, we introduce operators to rewrite (5) into a concise form: $\boldsymbol{A} := [\boldsymbol{I}, \boldsymbol{I} + \mathbf{D}^{\mathbf{s}}, \boldsymbol{I} - i\mathbf{D}^{\mathbf{s}}]$, where $\boldsymbol{I}$ is the identity transform, $\tilde{\boldsymbol{\Omega}} := (\boldsymbol{\Omega}_0, \boldsymbol{\Omega}_1, \boldsymbol{\Omega}_2)$, and $\boldsymbol{F}\mathbf{z} := (\boldsymbol{F}\mathbf{z}_0, \boldsymbol{F}\mathbf{z}_1, \boldsymbol{F}\mathbf{z}_2)$ for $\mathbf{z} = (\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_2)$. Then, the loss function in (5) is rewritten into a simple form

$$\mathscr{E} = \||\boldsymbol{F}\boldsymbol{A}\mathbf{u}| - \mathbf{b}\|_{\ell_{1-2},\tilde{\boldsymbol{\Omega}}}. \tag{6}$$

Taking the advantage of the sparsity at the image patch level, we incorporate the WNN [21] into (6) as a regularization term. The new model, which we named the WNN model, is

$$\min_{0\le\mathbf{u}\le 1} \mathscr{E}_{WNN}(\mathbf{u}) = \|\mathbf{u}\|_{WNN} + \tfrac{\lambda}{2}\||\boldsymbol{F}\boldsymbol{A}\mathbf{u}| - \mathbf{b}\|_{\ell_{1-2},\tilde{\boldsymbol{\Omega}}}, \tag{7}$$

where $\|\mathbf{u}\|_{WNN}$ is the regularization term, whose definition can be found in Section 3. The parameter λ balances the regularization term $\|\mathbf{u}\|_{WNN}$ and the data fitting term.

3. Numerical algorithms

In Section 3.1, we review the weighted nuclear norm proximal mapping (see Definition 1 and the weighted nuclear norm is defined in (3)), and its approximate solution suggested in [21], where the authors adaptively turn weights. Following [23], they choose the weight w_i in k^{th} iteration to be $\frac{C}{|x_i^k|+\varepsilon}$, where C is a non-zero constant. x_i is the i^{th} singular value calculated in the k^{th} iteration. ε is a small positive number to avoid divide overflow. Strategically, Gu et al. [21] came up with a very efficient closed-form solution to the non-convex problem, essentially it provides a relaxed soft-thresholding for the singular value, and the algorithm achieves superior performance in numerical experiments. However, because of the dynamic setting of weight values, without a fixed form of the objective function, it is hard to explore the convergence behavior. Therefore, we analyze the efficient closed-form approximate solution to the optimization problem involving the WNN, and prove the approximate solution proposed in [21] can be seen as an exact solution of a fixed functional. With this functional, after we review the way to generate patch groups in Section 3.3, and illustrate the algorithm to solve the proposed model in Section 3.4, we can give its convergence results in Section 3.5.

3.1. Weighted nuclear norm proximal mapping

In this paper, we follow the definition of the weighted nuclear norm proximal operator in [21].

Definition 1. For a matrix $\mathbf{Y} \in \mathbb{R}^{d\times m}$, the weighted nulcear norm proximal (WNNP) mapping with given weights $\{w_i\}$ is defined as

$$\begin{aligned}\mathbf{X}^* = \mathbf{prox}_{\|\cdot\|_{\mathbf{w},*}}(\mathbf{Y}) &= \arg\min_{\mathbf{X}} \|\mathbf{X}\|_{\mathbf{w},*} + \frac{1}{2}\|\mathbf{X}-\mathbf{Y}\|_F^2 \\ &= \arg\min_{\mathbf{X}} \sum_i w_i\sigma_i(\mathbf{X}) + \frac{1}{2}\|\mathbf{X}-\mathbf{Y}\|_F^2. \end{aligned} \tag{8}$$

Here $\sigma_i(\mathbf{X})$ is the i^{th} singular value of matrix $\mathbf{X}$.

The closed-form solution for the WNNP was given in [21, Theorem 1 and Corollary 1]. Without loss of generality, we assume that $d \geq m$, and the thin singular value decomposition of $\mathbf{Y}$ is $\mathbf{U\Sigma V}^{\mathrm{T}}$, where $\mathbf{\Sigma} = diag(\sigma_1, \sigma_2, \ldots, \sigma_m) \in \mathbb{R}^{m\times m}$ with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m \geq 0$. Then the global optimum of the WNNP problem can be expressed as $\mathbf{X}^* = \mathbf{UDV}^{\mathrm{T}}$, where $\mathbf{D} = diag(d_1, d_2, \ldots, d_m)$ and

$$\begin{aligned} &\{d_i\}_{i=1,\ldots,m} = \arg\min_{\{x_i\}_{i=1,\ldots,m}} \sum_{i=1}^{m}\{w_i x_i + \tfrac{1}{2}\|x_i-\sigma_i\|^2\}, \\ &s.t. \quad x_1 \geq x_2 \geq \ldots \geq x_m \geq 0. \end{aligned} \tag{9}$$

Furthermore, if the weights $\{w_i\}_{i=1}^m$ satisfy $0 \leq w_1 \leq w_2 \leq, \ldots, \leq w_m$, we obtain $d_i = \max(\sigma_i - w_i, 0)$.

The weights in the WNNP bring flexibility and a potential good solution, but they are difficult to choose. In addition, setting an appropriate weight is critical for this model. Paper [21, Remark 1] suggests to set weights as

$$w_i = \frac{C}{|d_i| + \epsilon}.$$

It is inspired from the reweighted ℓ_1 norm for a sparse vector x from [23], where the weights for the ℓ_1 norm at the k^{th} iteration are based on

$$w_i^k = \frac{C}{|x_i^k| + \varepsilon}. \tag{10}$$

Here C is a non-zero constant. In the matrix case, a reweighted nuclear norm method iteratively solves the WNNP problem with

$$w_i^k = \frac{C}{|d_i^k| + \varepsilon}.$$

Therefore, the iteration is

$$d_i^{k+1} = \max(\sigma_i - \frac{C}{|d_i^k| + \varepsilon}, 0). \tag{11}$$

We can find the fixed point for the iteration, which is

$$d_i = \begin{cases} 0, & c_2 < 0; \\ \frac{c_1+\sqrt{c_2}}{2}, & \text{otherwise}, \end{cases} \tag{12}$$

where $c_1 = \sigma_i - \varepsilon$, $c_2 = (\sigma_i + \varepsilon)^2 - 4C$.

Though the shrinkage in (12) performs very well in practice, to the best of our knowledge, the exact function whose proximal mapping coincides with (12) is not given in literature.

In the following, we show that this shrinkage operator corresponds to the proximal operator of a nonconvex function. Since the constant parameter ε is very small (In [21], the authors use the floating-point relative accuracy, $2.2204e-16$), we let $\varepsilon = 0$ for simplicity and rewrite (12) as

$$d_i = \begin{cases} 0, & \sigma_i^2 - 4C < 0; \\ \frac{\sigma_i+\sqrt{\sigma_i^2-4C}}{2}, & \text{otherwise}, \end{cases} \tag{13}$$

We show that actually $d_i = \mathbf{prox}_{\mathcal{P}}(\sigma_i) := \arg\min_x \mathcal{P}(x) + (x-\sigma_i)^2/2$, where $\mathcal{P}$ is the continuous function defined as

$$\mathcal{P}(x) = \begin{cases} \frac{3}{2}(Cx)^{\frac{2}{3}}, & 0 \leq x \leq \sqrt{C}; \\ \frac{3}{2}C + C\ln(\frac{x}{\sqrt{C}}), & x > \sqrt{C}. \end{cases} \tag{14}$$

Proposition 1. *The proximal operator of $\mathcal{P}$ for non-negative y has a closed-form expression, and*

$$\mathbf{prox}_{\mathcal{P}}(y) \in \begin{cases} \{0\}, & \text{if } 0 \leq y < 2\sqrt{C}; \\ \{0, \sqrt{C}\}, & \text{if } y = 2\sqrt{C}; \\ \{\frac{y+\sqrt{y^2-4C}}{2}\}, & \text{if } y > 2\sqrt{C}. \end{cases}$$

Proof. The proximal $\mathbf{prox}_{\mathcal{P}}(y) := \arg\min_x \mathcal{Q}(x, y)$, where

$$\mathcal{Q}(x,y) = \begin{cases} f(x), & 0 \leq x \leq \sqrt{C}; \\ g(x), & x > \sqrt{C}, \end{cases} \tag{15}$$

$$\begin{aligned} f(x) &:= \frac{3}{2}(Cx)^{\frac{2}{3}} + \frac{1}{2}(x-y)^2, \\ g(x) &:= \frac{3}{2}C + C\ln(\frac{x}{\sqrt{C}}) + \frac{1}{2}(x-y)^2, \end{aligned} \tag{16}$$

where $f(x)$ is smooth over $[0, \sqrt{C}]$ and

$$\begin{aligned} f'(x) &= C^{\frac{2}{3}}x^{-\frac{1}{3}} + x - y, \\ f''(x) &= -\frac{1}{3}C^{\frac{2}{3}}x^{-\frac{4}{3}} + 1, \\ f'''(x) &\geq 0. \end{aligned} \tag{17}$$

Let $f''(\tilde{x}) = 0$, and we obtain $\tilde{x} = 3^{-3/4}\sqrt{C} < \sqrt{C}$. We have $f'(x) \to +\infty$ when $x \to 0+$. Thus,

$$f(x) \text{ is} \begin{cases} \text{concave}, & x \in [0, \tilde{x}]; \\ \text{convex}, & x \in [\tilde{x}, \sqrt{C}], \end{cases} \tag{18}$$

In the following, we will consider x in the intervals $[0, \tilde{x}]$, $[\tilde{x}, \sqrt{C}]$, and $[\sqrt{C}, +\infty]$ separately, and it is obvious that, to find the minimum, the concave part ($x \in [0, \tilde{x}]$) in (18) is easy, we only need to compare $f(0)$ and $f(\tilde{x})$.

We also compute the derivative of g,

$$\begin{aligned} g'(x) &= \frac{C}{x} + x - y, \\ g''(x) &= -\frac{C}{x^2} + 1 > 0, \ x \in (\sqrt{C}, +\infty]. \end{aligned} \tag{19}$$

(a) When $0 \le y < 2\sqrt{C}$, we have $g'(x) > 0$ for $x > \sqrt{C}$, and $f(\sqrt{C}) = g(\sqrt{C})$. Therefore, the minimum can be reached in $[0, \sqrt{C}]$. Recall in (17), $f'''(x) \ge 0$, then $f'(x)$ is convex.

(11) If $f'(\tilde{x}) \ge 0$, $f(x)$ is non-decreasing and we have

$$\mathbf{prox}_{\mathcal{P}}(y) = 0.$$

(22) If $f'(\tilde{x}) < 0$, that is $C^{2/3}\tilde{x}^{-1/3} + \tilde{x} - y < 0$.
Recall $f'(x) \to +\infty$ when $x \to 0+$. $f'(\sqrt{C}) = 2\sqrt{C} - y > 0$, then there exists a unique $\hat{x} \in (\tilde{x}, \sqrt{C})$ such that $f'(\hat{x}) = 0$. Considering the shape of $f(x)$ described in (18), we know that the minimum can only be reached at $f(\hat{x})$ or $f(0)$. Note that

$$\begin{aligned} f(\hat{x}) - f(0) =& \frac{3}{2}(C\hat{x})^{2/3} + \frac{1}{2}(\hat{x} - y)^2 - \frac{1}{2}y^2 \\ =& \frac{3}{2}(C\hat{x})^{2/3} + \frac{1}{2}\hat{x}^2 - \hat{x}y = \hat{x}(\frac{3}{2}C^{2/3}\hat{x}^{-1/3} + \frac{1}{2}\hat{x} - y). \end{aligned}$$

Let $h(x) = \frac{3}{2}C^{2/3}x^{-1/3} + \frac{1}{2}x - y$. We have $h(\sqrt{C}) = 2\sqrt{C} - y > 0$. The derivative $h'(x) = -\frac{1}{2}C^{2/3}x^{-4/3} + \frac{1}{2}$ being negative over $(0, \sqrt{C})$ shows that $h(x) > 0$ for all $x \in (0, \sqrt{C})$. Therefore, we have $f(\hat{x}) > f(0)$. So we also have

$$\mathbf{prox}_{\mathcal{P}}(y) = 0.$$

(b) When $y = 2\sqrt{C}$, similar to (a), we have $g'(x) > 0$ for $x > \sqrt{C}$, and $f(\sqrt{C}) = g(\sqrt{C})$. Therefore, the minimum can be reached in $[0, \sqrt{C}]$. Furthermore,

$$f'(\tilde{x}) = C^{1/2}3^{1/4} + 3^{-3/4}\sqrt{C} - 2\sqrt{C} < 0,$$

and recall $f'(x) \to +\infty$ when $x \to 0+$, and $f'(x)$ is convex. So we only need to compare $f(0)$ and $f(\sqrt{C})$ because $f'(\sqrt{C}) = 0$. Note that

$$f(\sqrt{C}) - f(0) = \frac{3}{2}C + \frac{1}{2}(\sqrt{C} - 2\sqrt{C})^2 - \frac{1}{2}(2\sqrt{C})^2 = 0.$$

So we have

$$\mathbf{prox}_{\mathcal{P}}(y) \in \{0, \sqrt{C}\}.$$

(c) When $y > 2\sqrt{C}$, we have $f'(\sqrt{C}) = g'(\sqrt{C}) < 0$. The function f is decreasing on the convex part $[\tilde{x}, \sqrt{C}]$, and recall on the concave part $[0, \tilde{x}]$, the minimum can only be reached in $f(0)$ or $f(\tilde{x})$. Therefore, we only need to compare $f(0)$ and the minimal value of g at $\hat{x} := \frac{y + \sqrt{y^2 - 4C}}{2}$ ($g''(x) > 0$). In fact, we have

$$\begin{aligned} g(\sqrt{C}) =\ & f(\sqrt{C}) = 3C/2 + (\sqrt{C} - y)^2/2 = 2C - \sqrt{C}y \\ & + y^2/2 < y^2/2 = f(0). \end{aligned}$$

Therefore, we have

$$\mathbf{prox}_{\mathcal{P}}(y) = \frac{y + \sqrt{y^2 - 4C}}{2}.$$

□

3.2. Description of the patch group

Considering that the WNN is actually calculated on each group, we give a brief introduction about how to construct patch groups, and this process can be elaborated in Fig. 1.

Image $\mathbf{u} \in \mathbb{R}^{n_1 \times n_2}$ is divided into N overlapped patches $\bar{\mathbf{y}}_j$ of size $\sqrt{d} \times \sqrt{d}$, after the square patch is stretched into a column vector $\mathbf{y}_j \in \mathbb{R}^d$, among these N stretched patches $\{\mathbf{y}_j\}_{j=1,\dots,N}$, a subset $\{\mathbf{y}_j\}_{j=1,\dots,M}$, $M < N$ is fixed as key patches. For each $\mathbf{y}_j$, $j \in \{1, \dots, M\}$, its most similar m patches are found from all patches falling into a fixed-size search window, and together they form a matrix $\mathbf{Y}_j \in \mathbb{R}^{d \times m}$, the matrix is essentially a group of similar patches, and also be called a patch group, then the WNN is computed on each group, and promotes sparsity. Without loss of generality, we can write the procedure as $\mathbf{Lu} = \mathbf{Y} = \{\mathbf{Y}_j\}_{j=\{1,\dots,M\}}$, where the operator $\mathbf{L}$ extracts patches from image u, then regroups them, and $\mathbf{L} : \mathbb{R}^{n_1 \times n_2} \to \mathbb{R}^{d \times m \times M}$.

3.3. Group sparse representation

We describe the implementation of the group sparse representation in [21], since their main results will be used here.

The WNNP is adopted to compute the denoised patch group

$$\mathbf{X}_j^* = \mathbf{prox}_{\|\cdot\|_{\mathbf{w},*}}(\mathbf{Y}_j).$$

Considering the overlapping of pixels in patches, the authors then solve the following minimization problem for the denoised image $\mathbf{v}^*$

$$\mathbf{v}^* = \arg\min_{\mathbf{v}} \| \sum_j (\bar{\mathbf{X}}_j - \mathbf{X}_j^*) \|_2^2, \tag{20}$$

where $\{\bar{\mathbf{X}}_j\}_{j=\{1,\dots,M\}} = \mathbf{Lv}$.

3.4. Solving the WNN model

We apply the penalty method to solve the proposed WNN model and present some convergence results in the next section. We start by rewriting the WNN model (7) into its group based form

$$\begin{aligned} &\min_{\mathbf{u}} \|\mathbf{P}\|_{\mathbf{w},*} + \tfrac{\lambda}{2}\||\mathbf{z}| - \mathbf{b}\|_{\ell_{1-2},\tilde{\mathbf{\Omega}}} + \chi(\mathbf{v}), \\ &s.t. \quad \begin{cases} \mathbf{z} = \boldsymbol{FAu}, \\ \mathbf{P} = \mathbf{Lv}, \\ \mathbf{u} = \mathbf{v}, \end{cases} \end{aligned} \tag{21}$$

where $\mathbf{P} = \{\mathbf{P}_j\}_{j=\{1,\dots,M\}} \in \mathbb{R}^{d \times m \times M}$ are groups of patches, and $\|\mathbf{P}\|_{\mathbf{w},*} = \sum_j \sum_i w_i \sigma_i(\mathbf{P}_j)$, $\sigma_i(\mathbf{P}_j)$ denotes the i^{th} singular value of $\mathbf{P}_j$. $\mathbf{z} = (\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_2)$, and

$$\chi(\mathbf{v}) = \begin{cases} 0, & 0 \le \mathbf{v} \le 1; \\ \infty, & \text{otherwise.} \end{cases} \tag{22}$$

Based on variable splitting and the quadratic penalty method, we solve

$$\begin{aligned} \mathscr{P}_{WNN}(\mathbf{u}, \mathbf{P}, \mathbf{v}, \mathbf{z}) =\ & \|\mathbf{P}\|_{\mathbf{w},*} + \frac{\lambda}{2}\||\mathbf{z}| - \mathbf{b}\|_{\ell_{1-2},\tilde{\mathbf{\Omega}}} + \chi(\mathbf{v}) \\ & + \frac{\rho_1}{2}\|\mathbf{z} - \boldsymbol{FAu}\|_F^2 + \frac{\rho_3}{2}\|\mathbf{P} - \mathbf{Lv}\|_F^2 \\ & + \frac{\rho_2}{2}\|\mathbf{u} - \mathbf{v}\|_F^2, \end{aligned} \tag{23}$$

where ρ_1, ρ_2, and ρ_3 in the unconstrained objective are positive constants. The penalty method minimize $\mathscr{P}_{WNN}$ with respect to $\mathbf{u}, \mathbf{P}, \mathbf{v}$, and $\mathbf{z}$ alternatively while increasing the variables ρ_1, ρ_2, and ρ_3. The algorithm is summarized in Algorithm 1.

Algorithm 1 Penalty Method for solving the WNN model (7).

Initialize $\mathbf{u}^0 = \mathbf{v}^0$, $k = 0$, $r > 1$.
while stopping conditions are not satisfied **do**
  Update $\mathbf{z}_0^{k+1}$ via (33), similarly for $\mathbf{z}_1^{k+1}, \mathbf{z}_2^{k+1}$;
  Solve (28) via (12) to obtain $\mathbf{P}^{k+1}$;
  Update $\mathbf{v}^{k+1}$ via (32);
  Update $\mathbf{u}^{k+1}$ via (37);
  increase the variables ρ_1, ρ_2 and ρ_3: $\rho_i = r\rho_i$, $i = 1, 2, 3$;
  $k \leftarrow k + 1$.
end while
output the solution $\mathbf{u}^* = \mathbf{u}^{k+1}$.

Below we present the closed-form solutions for minimizing the primal subproblems, respectively:

$$\mathbf{z}^{k+1} = \arg\min_{\mathbf{z}} \mathscr{P}_{WNN}(\mathbf{u}^k, \mathbf{P}^k, \mathbf{v}^k, \mathbf{z}), \tag{24}$$

$$\mathbf{P}^{k+1} = \arg\min_{\mathbf{u}} \mathscr{P}_{WNN}(\mathbf{u}^k, \mathbf{P}, \mathbf{v}^k, \mathbf{z}^{k+1}), \tag{25}$$

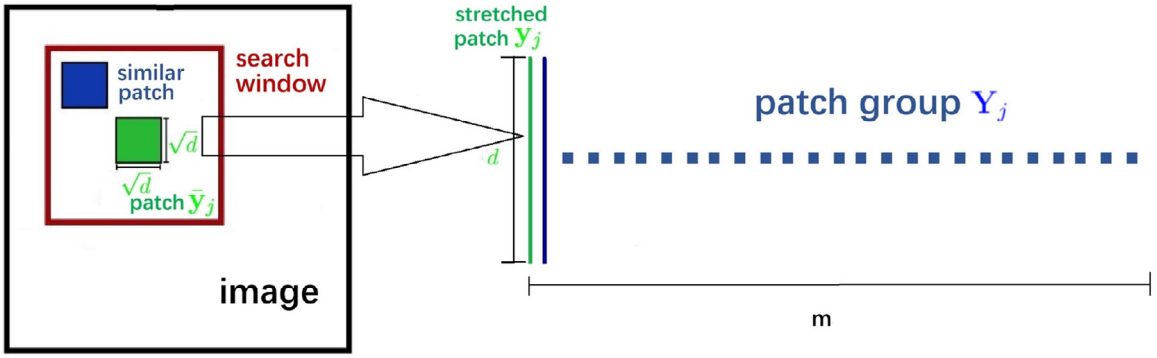

Fig. 1. Formulation of a patch group.

Fig. 2. 1st: 30% mask; the rest: clean image .

Table 1
SNR values of the recovered images under mixed impulse noise (both salt-and-pepper noise and random-valued impulse noise) by different methods.

SP +RI		dragonfly	swan	airplane	barbara	flowers	fruits	house	lena	peppers	stars	squares
	ER	18.85	13.80	17.48	14.69	15.48	18.34	19.16	15.64	14.86	25.32	26.56
0.2	PPA	20.94	17.43	18.82	15.75	17.61	19.72	20.32	15.89	15.92	26.94	28.36
+0.1	TV	21.73	18.20	19.45	16.16	18.14	20.15	20.75	16.32	16.54	27.52	29.13
	PEN	**23.89**	**20.87**	**20.91**	**17.71**	**19.77**	**21.95**	**24.17**	**18.11**	**18.55**	**28.86**	**30.37**
	ER	20.01	16.85	18.83	14.59	16.43	18.16	19.24	15.29	14.92	25.12	27.03
0	PPA	21.40	19.86	19.75	17.76	19.82	20.05	22.09	18.33	18.31	27.24	28.96
+0.3	TV	22.03	20.34	20.45	18.36	20.33	20.76	22.59	18.94	19.05	27.63	29.16
	PEN	**24.95**	**22.77**	**21.91**	**20.34**	**22.82**	**23.21**	**25.79**	**21.28**	**21.94**	**29.78**	**31.24**
	ER	23.38	20.19	20.54	18.69	20.35	22.45	23.24	20.77	19.80	26.21	28.15
0.1	PPA	24.71	22.94	23.15	20.41	23.26	23.31	25.52	21.27	22.29	28.12	29.56
+0.1	TV	25.14	23.52	23.70	20.90	23.66	23.82	26.22	21.93	22.72	28.72	30.12
	PEN	**27.29**	**24.93**	**24.64**	**22.42**	**25.74**	**25.14**	**28.32**	**23.81**	**25.28**	**30.21**	**31.64**
	ER	18.18	18.49	18.43	15.75	17.77	18.92	19.97	16.19	15.72	26.13	27.22
0.1	PPA	20.12	20.26	19.02	16.91	19.16	20.08	21.01	17.88	17.19	27.57	28.84
+0.2	TV	20.52	20.56	19.74	17.53	19.75	20.65	21.47	18.35	17.95	27.94	29.25
	PEN	**22.87**	**22.57**	**20.93**	**19.33**	**21.40**	**22.16**	**25.35**	**20.47**	**20.28**	**29.22**	**30.93**

$$\mathbf{v}^{k+1} = \arg\min_{\mathbf{v}} \mathscr{P}_{WNN}(\mathbf{u}^k, \mathbf{P}^{k+1}, \mathbf{v}, \mathbf{z}^{k+1}), \tag{26}$$

$$\mathbf{u}^{k+1} = \arg\min_{\mathbf{u}} \mathscr{P}_{WNN}(\mathbf{u}, \mathbf{P}^{k+1}, \mathbf{v}^{k+1}, \mathbf{z}^{k+1}). \tag{27}$$

There are closed-form solutions for all the four subproblem above. Recall $\mathbf{P} = \{\mathbf{P}_j\}_{j=\{1,\dots,M\}}$

$$\mathbf{P}_j^{k+1} = \arg\min_{\mathbf{P}_j} \|\mathbf{P}_j\|_{\mathbf{w},*} + \frac{\rho_3}{2}\|\mathbf{P}_j - (\mathbf{L}\mathbf{v}^k)_j\|_2^2. \tag{28}$$

Remark 1 in [21] gives the closed-form solution of the problem (28) with the weights chosen as (10).

$$\mathbf{v}^{k+1} = \arg\min_{\mathbf{v}} \rho_3\|\mathbf{L}\mathbf{v} - \mathbf{P}^{k+1}\|_2^2 + \rho_2\|\mathbf{v} - \mathbf{u}^{k+1}\|_2^2 + \chi(\mathbf{v}). \tag{29}$$

Considering the fact that a pixel in $\mathbf{v}$ can have K duplicate copies in $\mathbf{L}\mathbf{v}$, correspondingly, it related to K elements in the same position h in $\mathbf{P}^{k+1}$. In the following two formulas, we slightly abuse notations by omitting the index that indicates the position of a pixel, for example, v^{k+1} actually stands for a pixel,

$$\rho_3 \sum_{h=1}^{K} \left(v^{k+1} - P_h^{k+1}\right) + \rho_2\left(v^{k+1} - u^{k+1}\right) = 0, \tag{30}$$

and $0 \le v^{k+1} \le 1$, then we have

$$\hat{v}^{k+1} = \frac{1}{\rho_2 + \rho_3 K}\left(\rho_3 \sum_{h=1}^{K} P_h^{k+1} + \rho_2 u^{k+1}\right). \tag{31}$$

Followed by thresholding

$$v^{k+1} = \min\left\{1, \max\left\{0, \hat{v}^{k+1}\right\}\right\}. \tag{32}$$

Recall that $\bar{\mathbf{\Omega}} := (\mathbf{\Omega}_0, \mathbf{\Omega}_1, \mathbf{\Omega}_2)$, so the three components of $\mathbf{z}$ can be calculated separately with the same strategy [14], therefore, we only show the derivation of the closed-form solution for updating $\mathbf{z}_0$. The solution

$$\mathbf{z}_0^{k+1} = \arg\min_{\mathbf{z}_0} \lambda/\rho_1 \| |\mathbf{z}_0| - \mathbf{b}_0 \|_{\ell_{1-2},\mathbf{\Omega}_0} + \|\mathbf{z}_0 - \boldsymbol{F}\mathbf{u}\|_2^2,$$

where $\mathbf{g}_0 = \boldsymbol{F}\mathbf{u}$, obviously, $\mathbf{z}_0^{k+1}$ parallel to $\mathbf{g}_0$. (Otherwise, we can always project $\mathbf{z}_0^{k+1}$ onto $\mathbf{g}_0$, and get a better solution.) Thus let $\mathbf{x} = |\mathbf{z}_0| - \mathbf{b}_0$, then it is equivalent to solve the following minimization problem, which has a closed-form solution [34]

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} \|\mathbf{x} - (\mathbf{g}_0 - \mathbf{b}_0)\|^2 + \tfrac{\lambda}{\rho_1}\|\mathbf{x}\|_{\ell_{1-2},\mathbf{\Omega}_0},$$

then

$$\mathbf{z}_0^{k+1} = (\mathbf{x}^{k+1} + \mathbf{b}_0)\mathbf{n}. \tag{33}$$

To make this paper self-contained, we list the derivatives

$$\partial_{\mathbf{u}}\left(\|\mathbf{z}_0 - \boldsymbol{F}\mathbf{u}\|^2/2\right) = \partial_{\mathbf{u}}\left(\|\boldsymbol{F}^*\mathbf{z}_0 - \mathbf{u}\|^2/2\right) = \mathbf{u} - \Re(\boldsymbol{F}^*\mathbf{z}_0), \tag{34}$$

$$\partial_{\mathbf{u}}\left(\|\mathbf{z}_1 - \boldsymbol{F}(\mathbf{u} + \mathbf{D}^{\mathbf{s}}\mathbf{u})\|^2/2\right) = (2 + 2\Re(\mathbf{D}^{\mathbf{s}}))\mathbf{u} - \Re(\boldsymbol{F}^*\mathbf{z}_1 + \mathbf{D}^{\mathbf{s}}\overline{\boldsymbol{F}^*\mathbf{z}_1}), \tag{35}$$

$$\partial_{\mathbf{u}}\left(\|\mathbf{z}_2 - \boldsymbol{F}(\mathbf{u} - i\mathbf{D}^{\mathbf{s}}\mathbf{u})\|^2/2\right) = (2 + 2\Im(\mathbf{D}^{\mathbf{s}}))\mathbf{u} - \Re(\boldsymbol{F}^*\mathbf{z}_2 - i\mathbf{D}^{\mathbf{s}}\overline{\boldsymbol{F}^*\mathbf{z}_2}), \tag{36}$$

where $\boldsymbol{F}^*$ is the inverse Fourier transform, $\bar{\mathbf{z}}$ is the complex conjugate of $\mathbf{z}$. Therefore, to solve (27), we only need to solve a linear equation with invertible and diagonal matrix $\mathbf{H}$,

$$u = \mathbf{H}^{-1}(\rho_1(\Re(\hat{\mathbf{z}}_0^k + \hat{\mathbf{z}}_1^k + \mathbf{D}^{\mathbf{s}}\bar{\hat{\mathbf{z}}}_1^k + \hat{\mathbf{z}}_2^k) + \Im(\mathbf{D}^{\mathbf{s}}\bar{\hat{\mathbf{z}}}_2^k)) + \rho_2\mathbf{v}^k), \tag{37}$$

where $\mathbf{H} = (\rho_1(5\boldsymbol{I} + 2\Re(\mathbf{D}^{\mathbf{s}}) + 2\Im(\mathbf{D}^{\mathbf{s}})) + \rho_2\boldsymbol{I})$, $\hat{\mathbf{z}}_i^k = \boldsymbol{F}^*\mathbf{z}_i^k$ for$0 \le i \le 2$.

In Algorithm 1, we set a constant $r > 1$, to increase the penalty coefficients ρ_1, ρ_2 and ρ_3.

3.5. Convergence results

We apply the convergence theorems listed in [36], and verify the objection function satisfies the required conditions, therefore the result of Algorithm 1 is a stationary point of $\mathscr{P}_{WNN}(\mathbf{u}, \mathbf{P}, \mathbf{v}, \mathbf{z})$.

In [36], the authors consider the following optimization problem,

$$\underset{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_N}{\text{minimize}}\ f(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_N), \tag{38}$$

where the nondifferentiable part of f is separable,

$$f(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_N) = f_0(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_N) + \sum_{k=1}^{N} f_k(\mathbf{x}_k), \tag{39}$$

the function f is proper, $f_0 : \mathbb{R}^{n_1+\cdots+n_N} \mapsto \mathbb{R} \cup \{\infty\}$ and $f_k : \mathbb{R}^{n_k} \mapsto \mathbb{R} \cup \{\infty\}$, $k = 1, \ldots, N$. Here, $N, n_1, \ldots, n_N$ are positive integers, and each $\mathbf{x}_k$, $k = 1, \ldots, N$, is a coordinate block of $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$.

Lemma 1. *Suppose that* $\operatorname{dom} f_0$ *is open and* f_0 *is Gâteaux-differentiable on* $\operatorname{dom} f_0$*, then* f *is regular at each* $\mathbf{z} \in \operatorname{dom} f$.

Theorem 1. *Assume that the level set* $\mathbf{X}^0 = \left\{\mathbf{x} : f(\mathbf{x}) \le f\left(\mathbf{x}^0\right)\right\}$ *is compact and that* f *is continuous on* $\mathbf{X}^0$*. Then, the sequence* $\left\{\mathbf{x}^r = \left(\mathbf{x}_1^r, \ldots, \mathbf{x}_N^r\right)\right\}_{r=0,1,\ldots}$ *generated by the penalty (Block Coordinate Descent) method using the essentially cyclic rule is defined and bounded. Moreover, the following statements hold: If* $f(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ *has at most one minimum in* $\mathbf{x}_k$ *for* $k = 2, \ldots, N-1$ *and if the cyclic rule is used, then every cluster point* $\mathbf{z}$ *of* $\{\mathbf{x}^r\}_{r \equiv (N-1) \bmod N}$ *is a coordinatewise minimum point of* f*. In addition, if* f *is regular at* $\mathbf{z}$*, then* $\mathbf{z}$ *is a stationary point of* f.

Recall (23), we let

$$\begin{aligned} f(\mathbf{z}, \mathbf{P}, \mathbf{v}, \mathbf{u}) &= \mathscr{P}_{WNN}(\mathbf{u}, \mathbf{P}, \mathbf{v}, \mathbf{z}) \\ &= f_0 + \|\mathbf{P}\|_{\mathbf{w},*} + \frac{\lambda}{2}\| |\mathbf{z}| - \mathbf{b} \|_{\ell_{1-2},\bar{\mathbf{\Omega}}} + \chi(\mathbf{v}), \end{aligned} \tag{40}$$

where $f_0(\mathbf{z}, \mathbf{P}, \mathbf{v}, \mathbf{u}) = \frac{\rho_1}{2}\|\mathbf{z} - \boldsymbol{F}\boldsymbol{A}\mathbf{u}\|_F^2 + \frac{\rho_3}{2}\|\mathbf{P} - \mathbf{L}\mathbf{v}\|_F^2 + \frac{\rho_2}{2}\|\mathbf{u} - \mathbf{v}\|_F^2$, then from Lemma 1, we know f is regular. Moreover, f is continuous on the compact level set $\mathbf{X}^0 = \left\{\mathbf{x} : f(\mathbf{x}) \le f\left(\mathbf{x}^0\right)\right\}$. Recall the two minimization subproblem (28) and (29), they all have at most one minimum, then use Theorem 1, we know the output $(\mathbf{z}^{k+1}, \mathbf{P}^{k+1}, \mathbf{v}^{k+1}, \mathbf{u}^{k+1})$ is a stationary point of $\mathscr{P}_{WNN}$.

4. Numerical experiments

Here we assume the three masks are randomly generated, and we further assume that they are identical, *i.e.*, $\mathbf{\Omega}_0 = \mathbf{\Omega}_1 = \mathbf{\Omega}_2$, then we randomly generate those masks in the same way as in [14], and here in $\mathbf{\Omega}_i$, only 30% of pixels are available, the rest of pixels are considered lost, as shown in the 1st row of the top line in Fig. 2.

The signal-noise-ratio (SNR) is used to measure the reconstruction quality

$$\mathrm{RMSE}(\mathbf{u}, \mathbf{u}_g) = \frac{\sum_{j\in\Omega} |\mathbf{u}(j) - \mathbf{u}_g(j)|^2}{\sum_{j\in\Omega} |\mathbf{u}_g(j)|^2},$$

$$\mathrm{SNR}(\mathbf{u}, \mathbf{u}_g) = -10\log_{10}\mathrm{RMSE}(\mathbf{u}, \mathbf{u}_g),$$

where $\mathbf{u}_g$ is the ground truth image of size $n_1 \times n_2$ and $\mathbf{u}$ is the reconstructed image.

The SNR of the noisy measurements is defined as

$$\mathrm{SNR}(\mathbf{N}, \mathbf{M}) = -10\log_{10}\Big(\sum_{j\in\mathbf{\Omega}_0, 0\le k\le 2} |\hat{\mathbf{b}}_k(j) - \mathbf{b}_k(j)|^2 / \sum_{j\in\mathbf{\Omega}_0, 0\le k\le 2} |\mathbf{b}_k(j)|^2\Big),$$

where $\mathbf{N} = \{\hat{\mathbf{b}}_0, \hat{\mathbf{b}}_1, \hat{\mathbf{b}}_2\}$ is the noisy measurement with respect to the mask, and $\mathbf{M} = \{\mathbf{b}_0, \mathbf{b}_1, \mathbf{b}_2\}$ is the ground truth measurement.

Total variation (TV) is widely used as a regularizer [15,37,38], here we replace the WNN regularizer in the proposed method with the TV, and it is easy to see that we only need to change (27) and (26) a little bit to solve the TV model, then we compare with the proposed method, along with the ER method [14] and the newly proposed variational model [15] we mentioned before, we use PPA to note this method.

4.1. Noisy incomplete measurements

We consider both incomplete and noisy measurements, which are plotted in the first column in Figs. 3, 4, 5 and 6. After the mask (see Fig. 2) is applied, the observed intensity values suffer from a major information lost, only 30% data left. Then, the data is also corrupted by salt-and-pepper noise and random-valued impulse noise described in (4). Ground-truth images are presented in Fig. 2, where two sparse type images are considered: 10% nonzero pixels (stars) and 2% nonzero pixels (squares). From Table 1, we can see the proposed WNN model can outperform the ER model, PPA and TV based models in terms of SNR. To further demonstrate the effectiveness of the proposed model, beside the noise described in (4), we also test the proposed method to recover from

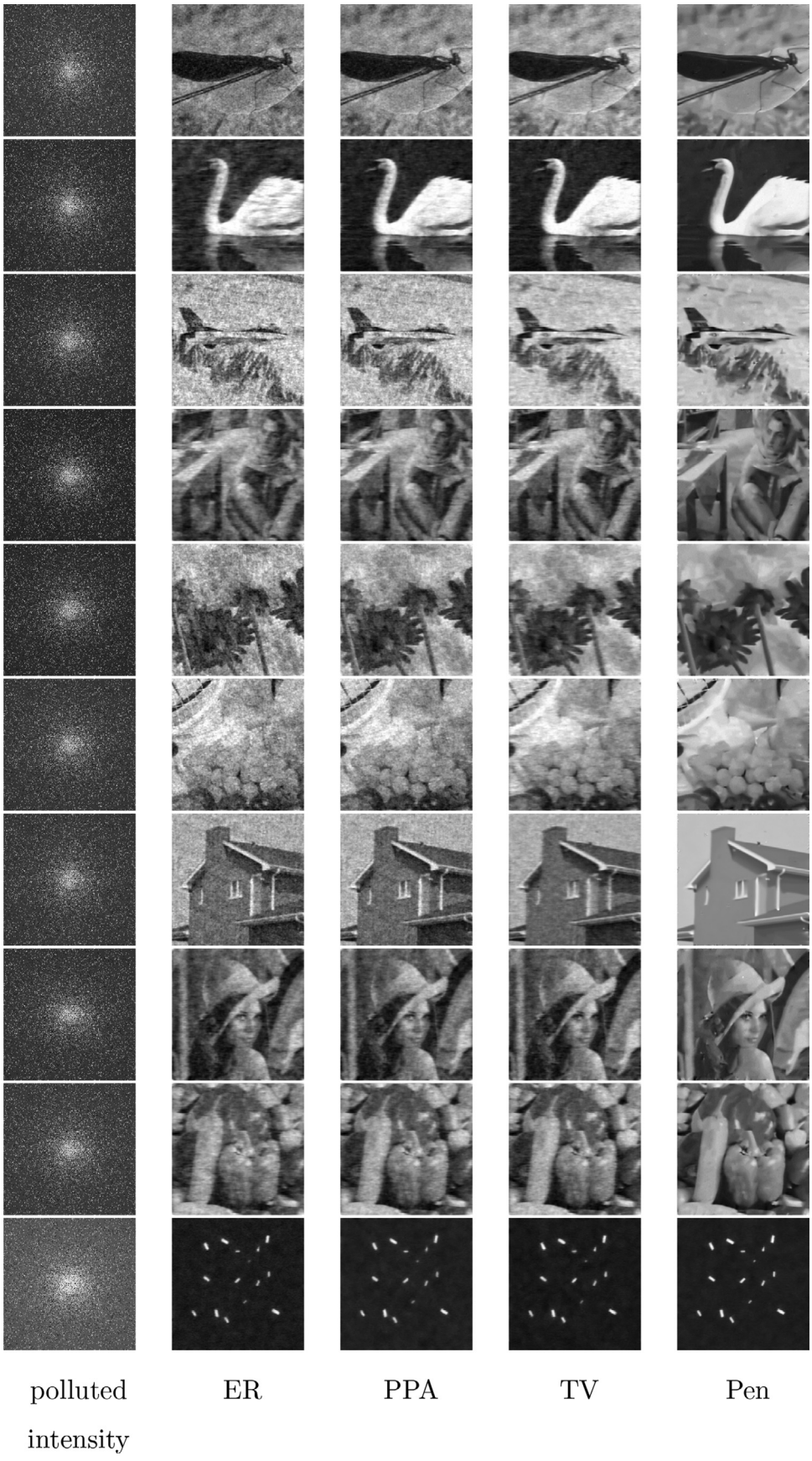

Fig. 3. Results of different methods when measurements are contaminated by mixed impulse noise. (Levels of salt-and-pepper noise and random-valued impulse noise are 0.2 and 0.1, respectively, refer to (4).).

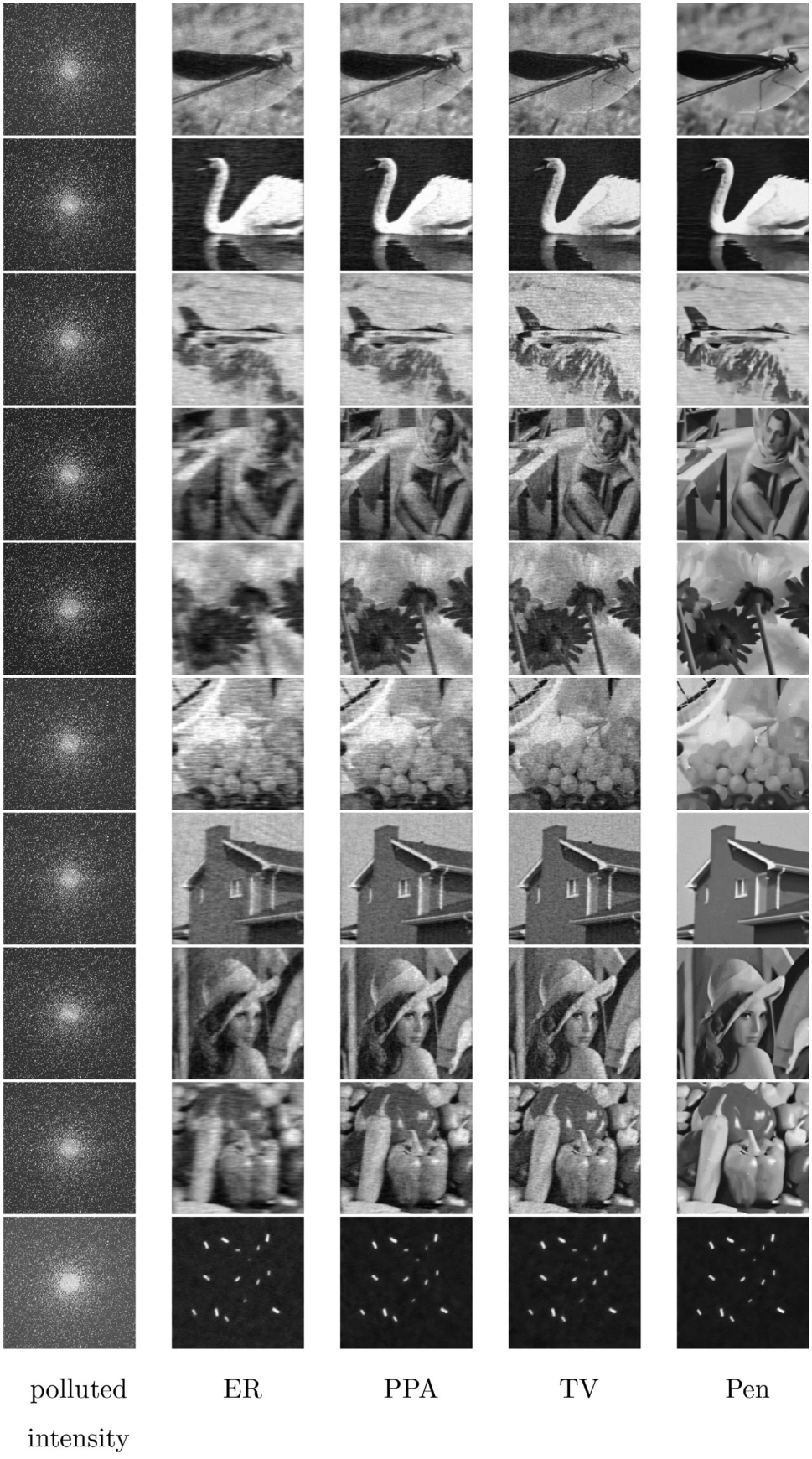

Fig. 4. Results of different methods when measurements are contaminated by random-valued impulse noise only with level 0.3, refer to (4).

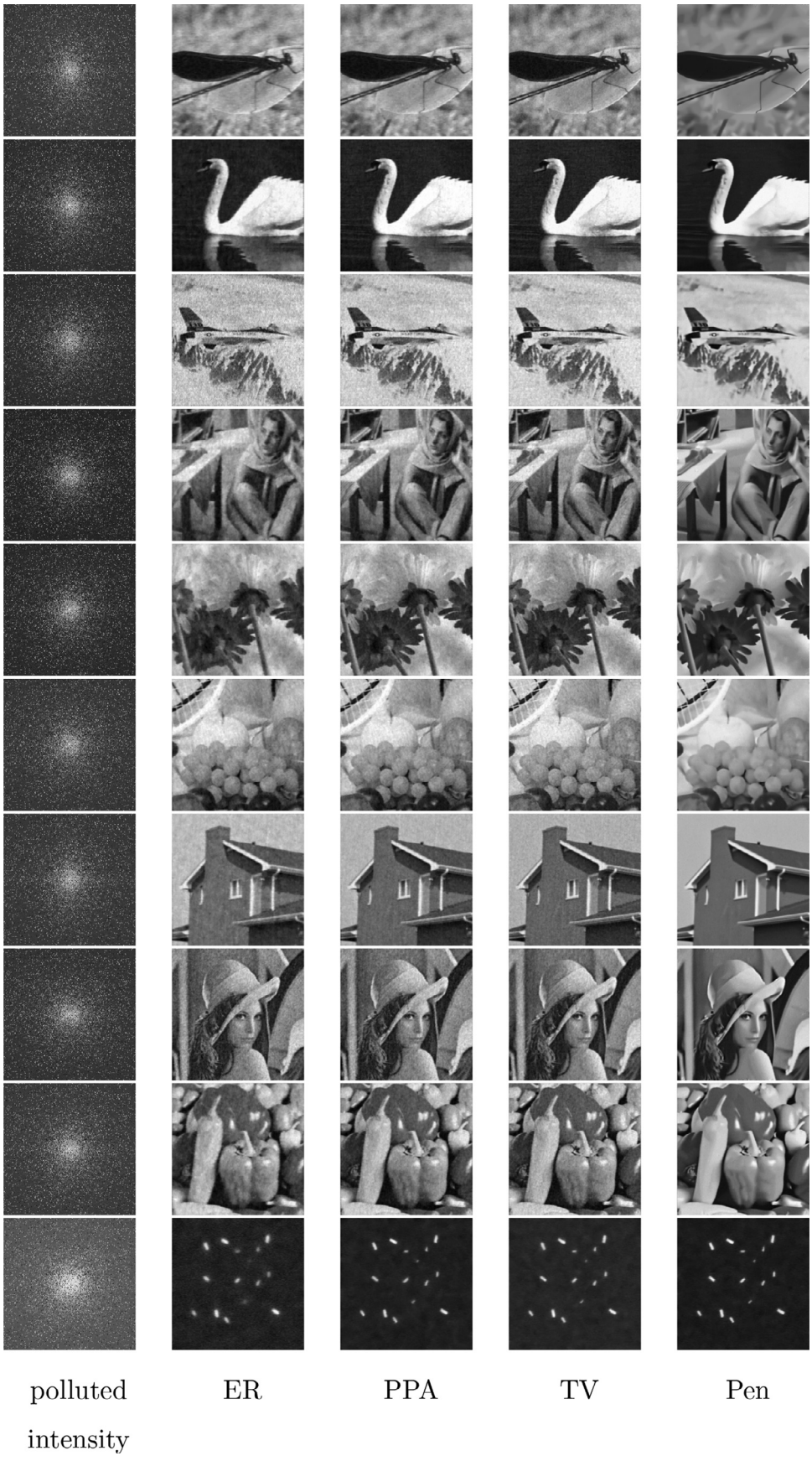

Fig. 5. Results of different methods when measurements are contaminated by mixed impulse noise (Levels of salt-and-pepper noise and random-valued impulse noise are 0.1 and 0.1, respectively, refer to (4).).

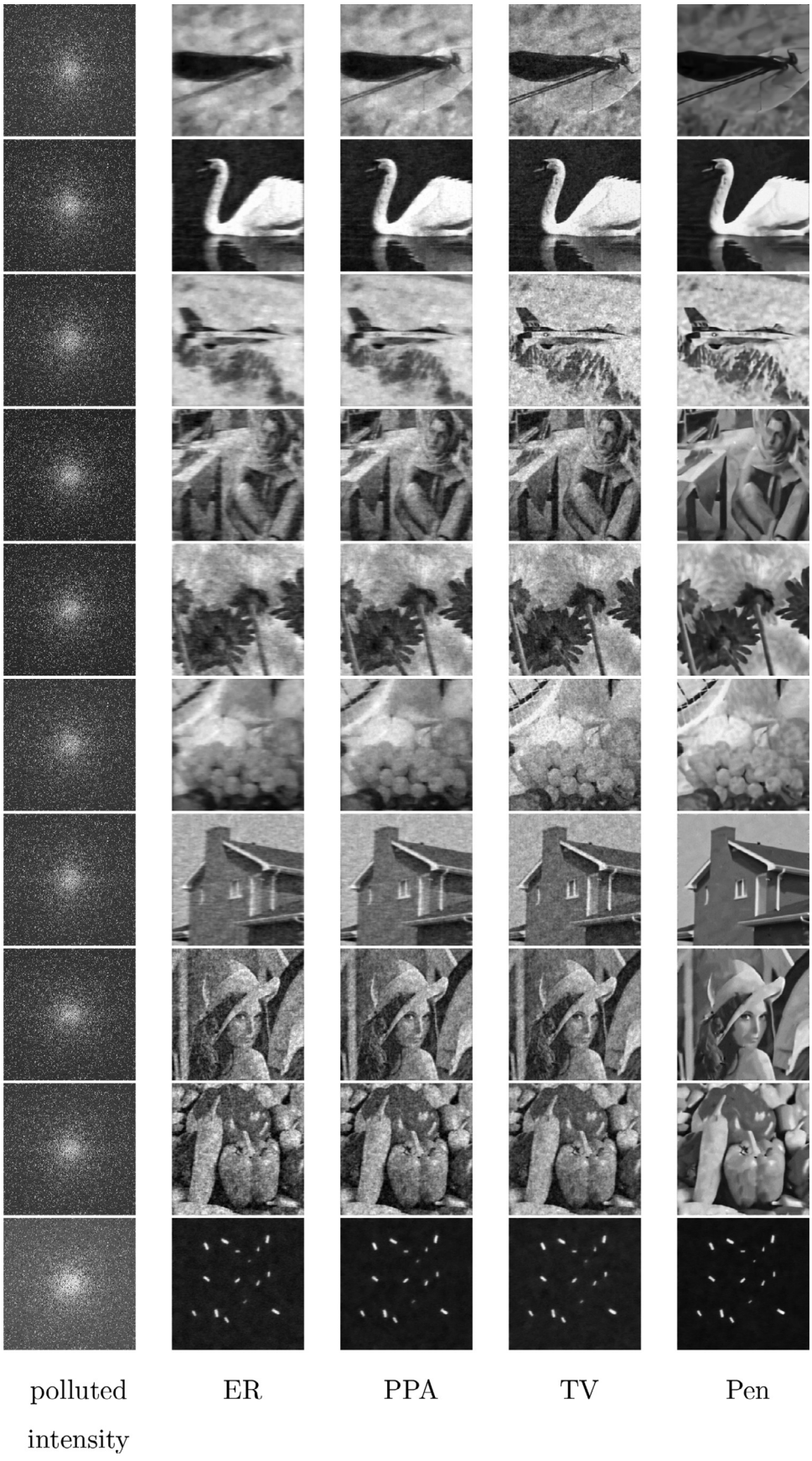

Fig. 6. Results of different methods when measurements are contaminated by mixed impulse noise (Levels of salt-and-pepper noise and random-valued impulse noise are 0.1 and 0.2, respectively, refer to (4).).

Table 2
SNR values of the recovered images under mixed impulse-Gaussian noise (salt-and-pepper noise, random-valued impulse noise and Gaussian noise) by different methods.

SP+RI+GS		swan	airplane	flowers	fruits	stars	squares
	ER	14.32	18.26	16.07	18.78	25.93	27.34
0.2	PPA	17.86	19.54	18.33	20.57	27.71	29.13
+0	TV	19.05	20.28	18.84	20.79	28.27	29.92
+0.1	PEN	**21.62**	**22.03**	**21.13**	**22.73**	**30.21**	**31.19**
	ER	19.53	20.11	19.84	22.02	25.86	27.58
0.1	PPA	22.23	22.65	22.63	22.87	27.84	29.16
+0.1	TV	22.87	23.21	23.15	23.31	28.19	29.82
+0.1	PEN	**24.32**	**24.18**	**25.15**	**24.82**	**29.91**	**31.28**
	ER	18.87	18.92	18.32	19.48	26.72	27.78
0	PPA	20.91	19.75	19.86	20.79	28.02	29.27
+0.2	TV	21.12	20.24	20.31	21.15	28.63	29.91
+0.1	PEN	**23.28**	**21.85**	**22.24**	**22.85**	**29.73**	**31.26**

mixed impulse-Gaussian noise, from Table 2, we can see that the proposed WNN model still can achieve better results.

As we can see, for example, from Fig. 4, we observe that ER can not restore clear results, for example, in the last line of Fig. 4, we can barely tell shapes of different peppers in the first column, while both WNN and TV models can recover images with clear contour of each pepper, however the TV restorations still have noticeable amount of noise left since collected intensity values are polluted and also suffer from a major data lost. We can see that the proposed model can recover the image with good visual quality, edge contours are clear, the homogeneous regions are reasonable smooth, see the pepper in the middle of the image located in the right corner of the last row of Fig. 4.

Also, as the noise gets heavier, the TV models can not yield clean restored images. As we can observe from Fig. 6, it is hard to tell shapes of peppers in the second and third columns of the last row. Also for the dragonfly and swan image, there are considerable amount of noise left. While from our result, in the first row of the last column, the noise are removed and the edge of the leaf where the dragonfly rests are shape, those thin legs of the dragonfly are also get restored.

5. Conclusion and future work

From compressive sensing perspective, we demonstrated that, if only a small portion of measurements are available and there are mixed impulse noise in the available data, we can still recover the image. The WNN regularization term is adopted to enforce the group sparsity patch-wisely, we provide a theoretical explanation for the approximate solution of the WNNM and also give the convergence proof, so that the widely used patch group algorithm family can benefit from our work, which sheds light on the way to explore the convergence behavior of these series of sparse coding approaches. Therefore, one can be more confident to use the closed-form solution of WNNM, since we provide analytical justification. Numerical experiments demonstrate the superior performance of the suggested model. However, the proposed approach has its limitations, on the condition that, for example, only 30% of data is available, then the impulse noise can not pollute more than one third of the remaining data, otherwise, the image can not be well restored by the proposed method.

Later on, we plan to extend the framework to deal with other types of pollution in the observed incomplete measurements, such as multiplicative noise [39]. Also, the analysis of the computational overhead of the penalty method is also an interesting topic to explore in the future.

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

This work was supported in part by the National Key R&D Program of China under Grant 2021YFE0203700, Grant NSFC/RGC N_CUHK 415/19, Grant ITF MHP/038/20, Grant RGC 14300219, 14302920, 14301121, CUHK Direct Grant for Research under Grant 4053405, 4053460, East China Normal University through startup funding grant, the Natural Science Foundation of China (Grant No. 61731009, 61961160734, 61971234, 11501301, 62001167), and the National Science Foundation (NSF) grants DMS- 1621798 and DMS-2012439.

References

[1] J. Dainty, J. Fienup, Phase retrieval and image reconstruction for astronomy, Image Recovery: Theory and Application, Academic Press (1987) 231–275.
[2] P. Skubák, Substructure determination using phase-retrieval techniques, Acta Crystallographica Section D: Structural Biology 74 (2) (2018) 117–124.
[3] Y. Shechtman, Y.C. Eldar, O. Cohen, H.N. Chapman, J. Miao, M. Segev, Phase retrieval with application to optical imaging: a contemporary overview, Signal Processing Magazine, IEEE 32 (3) (2015) 87–109.
[4] R.W. Gerchberg, W.O. Saxton, A practical algorithm for the determination of the phase from image and diffraction plane pictures, Optik (Stuttg) 35 (2) (1972) 237–246.
[5] J.R. Fienup, Phase retrieval algorithms: a comparison, Appl Opt 21 (15) (1982) 2758–2769.
[6] H.H. Bauschke, P.L. Combettes, D.R. Luke, Phase retrieval, error reduction algorithm, and Fienup variants: a view from convex optimization, Journal of the Optical Society of America A 19 (7) (2002) 1334–1345.
[7] H.H. Bauschke, P.L. Combettes, D.R. Luke, Hybrid projection-reflection method for phase retrieval, Journal of the Optical Society of America A 20 (6) (2003) 1025–1034.
[8] D.R. Luke, Relaxed averaged alternating reflections for diffraction imaging, Inverse Problem 21 (1) (2005) 37–50.
[9] Z. Wen, C. Yang, X. Liu, S. Marchesini, Alternating direction methods for classical and ptychographic phase retrieval, Inverse Probl 28 (11) (2012) 115010.
[10] T. Goldstein, C. Studer, Convex phase retrieval without lifting via phasemax, in: International Conference on Machine Learning, 2017, pp. 1273–1281.
[11] E.J. Candes, T. Strohmer, V. Voroninski, Phaselift: exact and stable signal recovery from magnitude measurements via convex programming, Commun Pure Appl Math 66 (8) (2013) 1241–1274.
[12] I. Waldspurger, A. Aspremont, S. Mallat, Phase recovery, maxcut and complex semidefinite programming, Mathematical Programming, Series A (2012) 1–35.
[13] R. Chandra, Z. Zhong, J. Hontz, V. McCulloch, C. Studer, T. Goldstein, Phasepack: A phase retrieval library, in: 2017 51st Asilomar Conference on Signals, Systems, and Computers, IEEE, 2017, pp. 1617–1621.
[14] H. Chang, Y. Lou, M.K. Ng, T. Zeng, Phase retrieval from incomplete magnitude information via total variation regularization, SIAM Journal on Scientific Computing 38 (6) (2016) A3672–A3695.
[15] H. Chang, S. Marchesini, Y. Lou, T. Zeng, Variational phase retrieval with globally convergent preconditioned proximal algorithm, SIAM J Imaging Sci 11 (1) (2018) 56–93.
[16] E.J. Candes, Y.C. Eldar, T. Strohmer, V. Voroninski, Phase retrieval via matrix completion, SIAM Rev. 57 (2) (2015) 225–251.
[17] J. Lu, H. Wang, J. Zhou, Y. Chen, Z. Lai, Q. Hu, Low-rank adaptive graph embedding for unsupervised feature extraction, Pattern Recognit 113 (2021) 107758.
[18] Z. Fu, Y. Zhao, D. Chang, Y. Wang, A hierarchical weighted low-rank representation for image clustering and classification, Pattern Recognit 112 (2021) 107736.
[19] B. Fan, X. Li, Y. Cong, Y. Tang, Structured and weighted multi-task low rank tracker, Pattern Recognit 81 (2018) 528–544.
[20] Z. Zha, X. Yuan, B. Wen, J. Zhang, J. Zhou, C. Zhu, Image restoration using joint patch-group-based sparse representation, IEEE Trans. Image Process. 29 (2020) 7735–7750.
[21] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, L. Zhang, Weighted nuclear norm minimization and its applications to low level vision, Int J Comput Vis 121 (2) (2017) 183–208.
[22] L. Ma, L. Xu, T. Zeng, Low rank prior and total variation regularization for image deblurring, J Sci Comput 70 (3) (2017) 1336–1357.
[23] E.J. Candes, M.B. Wakin, S.P. Boyd, Enhancing sparsity by reweighted ℓ_1 minimization, Journal of Fourier analysis and applications 14 (5–6) (2008) 877–905.
[24] X. Huang, B. Du, D. Tao, L. Zhang, Spatial-spectral weighted nuclear norm minimization for hyperspectral image denoising, Neurocomputing 399 (2020) 271–284.

[25] H. Deng, J. Tao, X. Song, C. Zhang, Estimation of the parameters of a weighted nuclear norm model and its application in image denoising, Inf Sci (Ny) 528 (2020) 246–264.
[26] G. Kim, J. Cho, M. Kang, Cauchy noise removal by weighted nuclear norm minimization, J Sci Comput 83 (1) (2020) 15.
[27] Y. Yu, Y. Zhang, S. Yuan, Quaternion-based weighted nuclear norm minimization for color image denoising, Neurocomputing 332 (2019) 283–297.
[28] N. Yair, T. Michaeli, Multi-scale weighted nuclear norm image restoration, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 3165–3174.
[29] S. Gu, L. Zhang, W. Zuo, X. Feng, Weighted nuclear norm minimization with application to image denoising, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2862–2869.
[30] M.L. Moravec, J.K. Romberg, R.G. Baraniuk, Compressive phase retrieval, in: Wavelets XII, volume 6701, International Society for Optics and Photonics, 2007, p. 670120.
[31] B. Aubin, B. Loureiro, A. Baker, F. Krzakala, L. Zdeborová, Exact asymptotics for phase retrieval and compressed sensing with random generative priors, in: Mathematical and Scientific Machine Learning, PMLR, 2020, pp. 55–73.
[32] C.A. Metzler, A. Maleki, R.G. Baraniuk, BM3D-PRGAMP: Compressive phase retrieval based on BM3D denoising, in: 2016 IEEE International Conference on Image Processing (ICIP), IEEE, 2016, pp. 2504–2508.
[33] Y. Dong, R.H. Chan, S. Xu, A detection statistic for random-valued impulse noise, IEEE Trans. Image Process. 16 (4) (2007) 1112–1120.
[34] Y. Lou, M. Yan, Fast ℓ_1–ℓ_2 minimization via a proximal operator, J Sci Comput 74 (2) (2018) 767–785.
[35] Y. Xiao, T. Zeng, J. Yu, M.K. Ng, Restoration of images corrupted by mixed gaussian-impulse noise via ℓ_1–ℓ_0 minimization, Pattern Recognit 44 (8) (2011) 1708–1720.
[36] P. Tseng, Convergence of a block coordinate descent method for nondifferentiable minimization, J Optim Theory Appl 109 (3) (2001) 475–494.
[37] C. Liu, M.K.-P. Ng, T. Zeng, Weighted variational model for selective image segmentation with application to medical images, Pattern Recognit 76 (2018) 367–379.
[38] Y. Yang, S. Han, T. Wang, W. Tao, X.-C. Tai, Multilayer graph cuts based unsupervised color–texture image segmentation using multivariate mixed student's t-distribution and regional credibility merging, Pattern Recognit 46 (4) (2013) 1101–1124.
[39] C. Shen, X. Bao, J. Tan, S. Liu, Z. Liu, Two noise-robust axial scanning multi-image phase retrieval algorithms based on pauta criterion and smoothness constraint, Opt Express 25 (14) (2017) 16235–16249.

Zhi Li received the B.S. degree and the M.S. degree from China University of Petroleum, Shandong, China, in 2007 and 2010, respectively. He also received the M.S. degree in applied science from Saint Maryś University, Halifax, NS, Canada, in 2012. After being awarded the Hong Kong Ph.D. Fellowship, he went to Hong Kong Baptist University, Hong Kong, where he received the Ph.D. degree in 2016. Then he worked as a Postdoctoral Researcher at Michigan State University, East Lansing, MI, USA, from 2016 to 2019. He is currently an associate researcher with the Department of Computer Science and Technology, East China Normal University, Shanghai, China.

Ming Yan received the B.S. and M.S. degrees from Uni- versity of Science and Technology of China, Hefei, China, and the Ph.D. degree fromUniversity of California, LosAnge- les,CA, USA, in 2012. He is currently an Assistant Professor with the Department of Computational Mathematics, Sci- ence and Engineering and the Department of Mathematics, Michigan State University. His research interests include op- timization methods and their applications in sparse recovery and regularized inverse problems, variationalmethods for im- age processing, parallel and distributed algorithms for solving big data problems.

Tieyong Zeng received the B.S. degree from Peking University, Beijing, China, in 2000, the M.S. degree from École Polytechnique, Palaiseau, France, in 2004, and the Ph.D. de- gree from the Université of Paris XIII, Paris, France, in 2007. He worked as a Postdoctoral Researcher with ENS de Cachan from 2007 to 2008 and an Assistant/Associate Professor with Hong Kong Baptist University from 2008 to 2018. He is currently a Professor with the Department of Mathematics, The Chinese University of Hong Kong.His re- search interests are image processing, machine learning, and scientific computing.

Guixu Zhang received the Ph.D. degree from the In- stitute of Modern Physics, Chinese Academy of Sciences, Lanzhou, China, in 1998. He is currently a Professor with the Department of Computer Science and Technology, East China Normal University, Shanghai, China. His research interests include hyperspectral remote sensing, image pro- cessing, artificial intelligence, machine learning, and scientific computing.