A mixed noise removal algorithm based on multi-fidelity modeling with nonsmooth and nonconvex regularization

Chun Li[1,2] · Yuepeng Li[1,3] · Zhicheng Zhao[1,2] · Longlong Yu[1,2] · Ze Luo[2]

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

In this article, we propose a mixed-noise removal model which incorporates with a non-smooth and nonconvex regularizer. To solve this model, a multistage convex relaxation method is used to deal with the optimization problem due to the nonconvex regularizer. Besides, we adopt the number of iteration steps as the termination condition of the proposed algorithm and select the optimal parameters for the model by a genetic algorithm. Several experiments on classic images with different level noises indicate that the robustness, running time, ISNR (Improvement in Signalto-Noise ratio) and PSNR (Peak Signal to Noise Ratio) of our model are better than those of other three models, and the proposed model can retain the local information of the image to obtain the optimal quantitative metrics and visual quality of the restored images.

Keywords Image restoration · Inverse problem · Alternating direction method of multipliers · Nonconvex optimization

1 Introduction

With the rapid development of computer technology, images are widely used in each field, especially in remote sensing and medical imaging. However, the images are usually polluted by all kinds of noises in the transmission. Due to the poor characteristics of noisy image, we cannot directly use noisy image in practical image applications (such as image classification, super-resolution, image enhancement and recognition). Therefore, image noise removal is an indispensable step in many practical applications, and an active topic in

✉ Chun Li
  lichun@cnic.cn

1   University of Chinese Academy of Sciences, Beijing, 100049 China

2   e-Science Technology and Application Laboratory, Computer Network Information Centre, Chinese Academy of Sciences, Beijing, 100190 China

3   Department of Big Data Technology and Application, Computer Network Information Centre, Chinese Academy of Sciences, Beijing, 100190 China

computer low-level vision. Formally speaking, image noise removal is an inverse problem which reconstructs the original image as much as possible from its polluted one, and the problem has been extensively studied in recent decades [8, 11, 18, 21–23, 25, 34, 39, 41–45, 47, 48, 68, 70]. Image denosing can be modeled as follows,

$$f = Au + n \tag{1}$$

Here, u denotes the origin image and Ω is a bounded open domain with a compact Lipschitz boundary. The observed image $f : \Omega \rightarrow R$ is partitioned into a true image and additive noise n. A can either be an identity operator or a blurring operator. In practical applications, two types of noises are often encountered during image transmission and image acquisition: white Gaussian noise (GN) and impulse noise (IN). Commonly, GN is brought up into images due to the thermal motion in camera, while IN is generally caused by the damaged pixels in camera transmission and sensors [6].

In recent decades, various techniques are proposed for image denoising, such as the methods in [63, 71, 74] for GN, the methods for IN removal [41, 44, 68, 70], and Poisson noise [28, 73]. Although those algorithms have powerful accomplishment for single noise reduction, they don't perform well in mixed noise denoising because different noises have different distributions.

In most previous works, Maximum a posteriori (MAP) estimation approach is often used for image denosing algorithms. We can formalize the problem as follows,

$$\hat{u} = \arg\max_{u} p(u|f) \tag{2}$$

Here, the observed image $f(x)$, and the original image $u(x)$ are two random variables for each $x \in \Omega$. Therefore, we can apply Bayes rule, and then obtain,

$$
\begin{aligned}
\hat{u} &= \arg\max_{u} p(u|f) \\
&= \arg\max_{u} \frac{p(f|u) p(u)}{p(f)} \\
&= \arg\min_{u} -\log(f|u) - \log(u) \\
&= \arg\min_{u} -\int_{\Omega} \log p(f(x)|u(x))dx - \log p(u)
\end{aligned}
\tag{3}
$$

Normally, the first term captures the noise-data distribution which is often called data fidelity term, and the second term denotes the prior information of origin image which is usually named regular term. In practical applications, the L_2-norm is used to fit the GN distribution, and L_1-norm data fidelity term is employed to simulate IN distribution, respectively [41, 44, 63, 68, 70, 71, 74]. Consequently, based on Bayes method, GN can be removed successfully by solving the following optimization problem,

$$\hat{u} = \arg\min_{u} \|f - u\|_2^2 + \lambda R(u) \tag{4}$$

here, $f \in R^N$, $u \in R^N$, λ, and $R(u)$ refer to noisy image, clean image, weight parameter, and the regularization term, respectively. Similarly, IN can be removed by solving the following minimization problem,

$$\hat{u} = \arg\min_{u} \left\{ \|f - Au\|_1 + \lambda R(u) \right\} \tag{5}$$

With the above derivations for the GN or IN image, now we can derive a model that links both types of noises. If we set up a multi-noise likelihood, according to Bayes' Theorem, and then we can obtain,

$$\hat{u} = \arg\min_{u} \left\{ \mu_1 \|f - u\|_2^2 + \mu_2 \|f - Au\|_1 + \lambda R(u) \right\} \tag{6}$$

here, μ_1 and μ_2 are also weight parameters.

Minimization problem (6) is a template of many variational optimization models. And in the past years, the variational optimization model for image processing is an active topic in computer vision all the time. Those models mainly consist of two parts: data-fidelity and regularization term. Regularization term can maintain the smoothness of the objective function and recover a clean image with sharp edges and fine image detail information. Model with regularization term is also called variational regularization [62], which have been applied to both statistical learning and variational workframe problems [4, 26, 30–32, 46], and if it is a convex regularization model, a unique solution of optimization problem exists. However, if it is a nonconvex regularization one, the solution of optimization problem is not necessarily unique. Recently, nonconvex regularization has become a popular method because it can restore high-quality images with well-kept detail information of local characteristics and edges. Therefore, Some sholars explored the theoretical solution and numerical solution in [53], [31, 32, 50, 55, 59], respectively. Since the nonconvex and nonsmooth regularization is a powerful method, there is still exploration space for improvement of performance in computer vison, especially for image noise reduction. Subsequently, in this paper, we introduce an efficient image mixed noised removal algorithm with multi-fidelity terms based on nonsmooth and nonconvex optimization.

Although a theoretical solution to the minimization problem with a nonconvex and nonsmooth regularization has not yet been proposed, many numerical algorithms, such as iteratively reweighted least squares algorithms [19, 31, 54, 66], half-quadratic algorithms [2, 14, 65], and graduated nonconvexity algorithms (GNC) [5, 55, 56], have been proposed. However, the GNC method, which accommodates the minimizer of nonsmooth and nonconvex potentials, requires considerable computational time. The iteratively reweighted (IR) method [15] was proposed to track the sparsity properties of the regularization comprehensive sensing problems, which has been introduced to process the image restoration problems [35, 57]. In consequence, in this paper, a nonsmooth and nonconvex regularization method is used to solve the image restoration problem. Meanwhile, due to the nonconvex regularization, we adopt a multistage nonconvex relaxation method [72] that gave a more better solution than the standard convex relaxation solution.

Apart from variational optimization models, the most often-used mixed noise removal approaches are based on the "detecting then filtering" strategy. Trilateral filter [29] combined a gradient bilateral filter and an intensity bilateral filter with a pyramidbased method to limit filter window. Nonlocal mean (NLM) [8, 9] filter averages all pixels in the image to remove image noise and the filter can also be set as a regular term in variational optimization models. Recently, with development of compressive sensing technology, sparse representation has become a powerful tool in various computer vision assignments. Some scientific researchers used the detection trick for mixed noise removal based on the sparse representation model. Tao et al. [64] have done a lot of precious works, and they used a powerful manifold ranking and optimization algorithm-based matrix factorization method for the saliency detection.

With the widespred appllication of deep learning, several discriminative learning methods have been developed to learning image prior models in the context of truncated inference

procedure. For example, Chen et al. [16, 17, 27] proposed a trainable nonlinear reaction diffusion model, which learnt a modified fields of experts [61] image prior by unfolding a fixed number of gradient descent inference steps. And He et al. [37] proposed a pulse-coupled neural network (PCNN)-based algorithm to underwater image enhancement scheme for robotic visual systems. The experimental results shown that the enhanced result improved the color and contrast of the source image and enhanced the details and edges of darker regions, which made a great contribution for deep-sea exploration tasks.

In this paper, we propose an efficient mixed GN and IN removal approach based on the L_1 and L_2 data-fidelity terms. Moreover, we also introduce a nonsmooth and nonconvex regularization term for solving the image restoration problem. Additionally, we use the multistage convex relaxation method in order to deal with the nonconvex regularization, which provides a good performance solution. Experimental results of simulated noisy images illustrate that our method outperforms current state-of-the-art mixed-noise removal methods.

2 Preliminaries

2.1 Alternating direction method of multipliers

In this subsection, the theories underpinning the proposed image mixed noise removal algorithm are introduced. Firstly, the Alternating direction method of multipliers (ADMM) will be explained. Therefore, now we consider the following structured constrained convex optimization problem, called the Alternating Direction Method [7, 26, 36],

$$\min_{x \in X, y \in Y} \{\theta_1(x) + \theta_2(y) \,|\, Ax + By = b, x \in X, y \in Y\} \tag{7}$$

here, $\theta_1(x) : R^{n_1} \to R, \theta_2(x) : R^{n_2} \to R$ are convex functions (but not necessary smooth). $A \in R^{m \times n_1}, B \in R^{m \times n_2}$ and $b \in R^m$, $X \in R^{n_1}$, and $Y \in R^{n_2}$ are given closed convex sets. Let λ be the Lagrangian multiplier for the linear constraints $Ax + By = b$ in (7). Then, the augmented Lagrangian function is as follows,

$$L(x, y, \lambda) = \arg\min \left\{ \begin{array}{l} \theta_1(x) + \theta_2(y) - \left(\lambda^k\right)^T (Ax + By - b) \\ + \frac{\beta}{2} \|Ax + By - b\|_2^2 \end{array} \;\middle|\; \begin{array}{l} x \in X \\ y \in Y \end{array} \right\} \tag{8}$$

which is defined on $X \times Y \times R^m$. Thirdly, let (x^*, y^*, z^*) be a saddle point of the Lagrangian function. So, $(x^*, y^*, z^*) \in \Omega$, and it satisfies,

$$\left\{ \begin{array}{l} \theta_1(x) - \theta_1(x^*) + (x - x^*)^T \left(-A^T \lambda^*\right) \geq 0 \\ \theta_2(y) - \theta_2(y^*) + (y - y^*)^T \left(-B^T \lambda^*\right) \geq 0 \quad , \forall (x, y, \lambda) \in \Omega \\ (\lambda - \lambda^*)^T (Ax^* + By^* - b) \geq 0 \end{array} \right. \tag{9}$$

here, $\Omega = X \times Y \times R^m$. By defining

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, F(w) = \begin{pmatrix} -A^T \lambda \\ -B^T \lambda \\ Ax + By - b \end{pmatrix} \tag{10}$$

and $\theta(u) = \theta_1(x) + \theta_2(y)$. The first-order optimal condition (9) can be written in a compact form as,

$$w^* \in \Omega, \theta(u) - \theta(u^*) + (w - w^*)^T F(w^*) \geq 0, \forall w \in \Omega \tag{11}$$

Note that the mapping F is monotone. We use Ω^* to denote the solution set of the variational inequality (11). For convenience, we use the notations, namely,

$$v = \begin{pmatrix} y \\ x \end{pmatrix}, V^* = \left\{ (x^*, y^*) \,\middle|\, (x^*, y^*, \lambda^*) \in \Omega \right\} \tag{12}$$

For a given λ^k, $u^{k+1} = \left(x^{k+1}, y^{k+1} \right)$ is the solution of the following problem,

$$\left(x^{k+1}, y^{k+1} \right) = \arg\min \left\{ \begin{array}{l} \theta_1(x) + \theta_2(y) \\ -(\lambda^k)^T (Ax + By - b) \\ +\frac{\beta}{2} \|Ax + By - b\|_2^2 \end{array} \,\middle|\, \begin{array}{l} x \in X \\ y \in Y \end{array} \right\} \tag{13}$$

The multiplier update is defined by,

$$\lambda^{k+1} = \lambda^k - \beta \left(Ax^{k+1} + By^{k+1} - b \right) \tag{14}$$

For a given $\left(y^k, \lambda^k \right)$, x^{k+1} is the solution of the following problem,

$$x^{k+1} = \left\{ \begin{array}{l} \theta_1(x) - \left(\lambda^k \right)^T \left(Ax + By^k - b \right) \\ + \frac{\beta}{2} \left\| Ax + By^k - b \right\|_2^2 \end{array} \right\} \tag{15}$$

Using operator separation technique for λ^k, x^{k+1}, and y^{k+1} respectively, we have the solution of the following problem,

$$y^{k+1} = \left\{ \begin{array}{l} \theta_2(y) - \left(\lambda^k \right)^T \left(Ax^{k+1} + By - b \right) \\ + \frac{\beta}{2} \left\| Ax^{k+1} + By - b \right\|_2^2 \end{array} \right\} \tag{16}$$

and

$$\lambda^{k+1} = \lambda^k - \beta \left(Ax^{k+1} + By^{k+1} - b \right) \tag{17}$$

Note that (15) and (16) are equivalent to (18) and (19), respectively.

$$x^{k+1} = \left\{ \theta_1(x) + \frac{\beta}{2} \left\| \left(Ax + By^k - b \right) - \frac{1}{\beta} \lambda^k \right\|_2^2 \right\} \tag{18}$$

$$y^{k+1} = \left\{ \theta_2(y) + \frac{\beta}{2} \left\| \left(Ax^{k+1} + By - b \right) - \frac{1}{\beta} \lambda^k \right\|_2^2 \right\} \tag{19}$$

2.2 The multi-stage convex relaxation method

In this subsection, because the multistage convex relaxation method will be used, so we first introduce it as following. From the multi-stage convex relaxation method first proposed in [72], we have the following optimization problem,

$$\hat{w} = \underset{w \in R^d}{\operatorname{argmin}} \left\{ g(w) = g_0(w) + \sum_{i=1}^{I} g_i(w) \right\} \tag{20}$$

where $g_0(w)$ is convex in w and each $g_i(w)$ is nonconvex. The convex/concave duality that was first defined in [72] can solve this optimization problem efficiently.

Let $h_i(w) : R^d \to \Omega_i \subset R^{d_i}$ be a vector function with range Ω_i, and assume that there is a function $\bar{g}_i$, which is defined on Ω_i, such that,

$$g_i(w) = \bar{g}(h_i(w)) \tag{21}$$

Suppose that exists h_i such that the function $\bar{g}_i(z_i)$ is concave on $z_i \in \Omega_i$. Then, the function $g_i(w)$ can be rewritten as,

$$g_i(w) = \min_{v_i \in R^{d_i}} \left\{ v_i^T h_i(w) + g_i^*(v_i) \right\} \tag{22}$$

Using the concave duality [60], if $g_i^*(v_i)$ is the concave dual of $\bar{g}_i(z_i)$, we define,

$$g_i^*(v_i) = \min_{z_i \in \Omega_i} \left[-v_i^T z_i + \bar{g}_i(z_i) \right] \tag{23}$$

The minimum of (22) can be obtained by solving the following equation,

$$\hat{V}_i = \nabla_z \bar{g}_i(z) \,|z = h_i(w) \tag{24}$$

Secondly, we present a simple convex relaxation of (20),

$$\hat{w} = \arg\min_{w \in R^d} \left\{ g_0(w) + \sum_{i=1}^{I} g_i(w)^T v_i \right\} \tag{25}$$

Therefore, we can summarize the multistage convex relation algorithm as Algorithm 1.

Algorithm 1 The multistage convex relation algorithm.

---

initialization:
    (1) Set parameters $\hat{v} = 1$;
    (2) Repeat the following two steps until convergence;
for $i = 1, 2, 3, ..., n$ **do**
    (1) Solve problem $\hat{w} = \underset{w \in R^d}{\arg\min} \left\{ g_0(w) + \sum_{i=1}^{I} g_i(w)^T v_i \right\}$;
    (2) Compute the problem $\hat{v}_i = \nabla_z \bar{g}_i(z) \,|z = h_i(w)$;
end

---

3 The related work

Numerous previous methods have been developed for image denoising in the past decades, and most of these works can be classified as two types methods, namely, deep learning Neural Network based method and regularization optimization based method.

3.1 Deep neural networks for image denoising

Several deep neural networks have been attempted to handle image noise reduction and image enhancement. Scientific scholars used sparse representation (SR) to deal with various image computer vision and image processing assignments in the past few years. In [69], the authors used sparse denoising auto-encoders technique to handle GN removal and obtained satisfactory results compared to K-SVD [25]. The multilayer perceptron (MLP) was victoriously used for image noise reduction in [10]. Convolutional neural networks (CNNs) for image noise removal has been used by Kingma and Ba [40], which have achieved comparable results. In [16] a trainable nonlinear reaction diffusion (TNRD) model was proposed which can be expressed as a feed-forward deep network by unfolding a fixed number of gradient descent inference steps. In [27] a trainable nonlinear reaction diffusion (TNRD)

model for passion noise is proposed and achieves promising performance. In [37] pulse-coupled neural network (PCNN)-based image enhancement and color transfer algorithms are combined to enhance the underwater image, which make some contribution for deep-sea exploration businesses. Kai Zhang and Lei Zhang et. [75] used residual learning and batch normalization to speed up the training process as well as boost the denoising performance. To the best of our knowledge, it remains uninvestigated to develop CNN for general image denoising.

3.2 Regularization optimization for image noise reduction

Traditional regularization optimization is still an active method in image processing and computer vision tasks, and most of these works use different regularization terms about the prior information to constraint image u. If an observed image u is corrupted with GN only, the noise distribution can be fitted successfully by the L_2-norm of the data fidelity term, thus we can solve the following optimization problem to remove advantageously the GN,

$$\hat{u} = \arg\min_u \| f - u \|_2^2 + R(u) \tag{26}$$

here, $R(u)$ is a regular function. If we formulate the Gaussian model as a MAP estimation, we can obtain the following general form,

$$\hat{u} = \arg\min_u \| f - u \|_2^2 + p(u) \tag{27}$$

here, $p(u) = R(u)$. Usually, we assume that u follows a Gibbs prior [3],

$$p(u) = \exp\left(-\gamma \int_\Omega \phi(u(x)) \right) dx \tag{28}$$

here, ϕ is a nonnegative function. Let $\phi(u(x)) = \| \nabla u(x) \|_1$, $\nabla u(x) = \left(\nabla_x u, \nabla_y u \right)$ is the gradient of u, and $\| \nabla u(x) \|_1$ is the total variation (TV) of u in [62].

The TV-L1 model [13] is obtained as a variant of the ROF model by replacing the squared-norm in the data fidelity term by the robust-norm. Although this was only a slight change, the TV-L1 model has some potential advantages over the ROF model. For example, 1) the TV-L1 model contrasts in variant; and 2) the TV-L1 model is much more effective in removing strong impulse noise (e.g., salt- and pepper noise). Because the Mumford-Shah (MS) model [49] is non-convex, therefore, it is difficult to solve the functional minimizer. To explore the true implicit solution of the MS function, Cai et al. [12] proposed the following equivalence model,

$$\hat{u} = \arg\min_u \left\{ \frac{\gamma_1}{2} \| f - Au \|_2^2 + \frac{\alpha}{2} \| \nabla u(x) \|_2^2 + \| \nabla u(x) \|_1 \right\} \tag{29}$$

here, γ_1 and α are positive weight parameters.

Because the TV-L1 model is more effective than the MS model at removing strong impulse noise (e.g., salt-and pepper noise), the L1-data fidelity term $\| f - Au \|_1$ plays an important role in the processing. The L1 data fidelity term has been extensively studied in [1, 51, 52]. Some researchers proposed the following model which is based on the TV-L1 model to remove strong impulse noise.

$$\hat{u} = \arg\min_u \left\{ \frac{\gamma_1}{2} \| f - Au \|_1 + \frac{\alpha}{2} \| \nabla u(x) \|_2^2 + \| \nabla u(x) \|_1 \right\} \tag{30}$$

To remove both GN and IN [38], a combined L1-L2 data fidelity terms were suggested by Shi, who combined (28) and (29) to obtain [43].

$$\hat{u} = \arg\min_u \left\{ \begin{array}{c} \frac{\gamma_1}{2} \| f - Au \|_1 + \frac{\gamma_2}{2} \| f - Au \|_2^2 \\ + \frac{\alpha}{2} \| \nabla u\,(x) \|_2^2 + \| \nabla u\,(x) \|_1 \end{array} \right\} \tag{31}$$

here, $\alpha > 0$, $\gamma_1 \geq 0$, and $\gamma_2 \geq 0$ are parameters that balance the data fidelity terms and the regularization terms.

Currently, the patch-based nonlocal low rank regularization was shown to have better performance than the nuclear norm regularization in image reconstruction [22, 24, 34, 39, 41, 42]. Because the rank minimization problem is the key subproblem in patch-based nonlocal rank regularization methods, and it can be approximately solved by minimizing the following optimization problem [24].

$$E\,(u, \varepsilon) = \log\,(u + \varepsilon I) \tag{32}$$

here, $u \in R^{n \times n}$ is a symmetric positive semidefinite matrix, I is the identity matrix. In [58], Osher et al. proposed a novel low dimensional manifold model (LDMM), not only applied it to image noise reduction, but also achieved successful accomplishment. Tao et al. [64] used Manifold Ranking-Based Matrix Factorization (MRMF) for Saliency Detection. In the paper, they proved that the MRMF has good generalizability, and developed an efficient optimization algorithm based on the Nesterov method for Saliency Detection.

4 Proposed nonconvex image denoising model and algorithms

Based on the Shi-Model [43], we propose the following image denosing model,

$$\hat{u} = \arg\min_u \left\{ \begin{array}{c} \frac{\gamma_1}{2} \| f - Au \|_1 + \frac{\gamma_2}{2} \| f - Au \|_2^2 \\ + \frac{\alpha}{2} \| \nabla u\,(x) \|_2^2 + p\,(u) \end{array} \right\} \tag{33}$$

Using the Gibbs prior $p\,(u) = \exp\left(-\gamma \int_\Omega \phi\,(u\,(x))\right) dx$, Shi let $\phi\,(u\,(x)) = \| \nabla u\,(x) \|_1$, which resulted in [16]. To improve the image denoising effect, we define the Gibbs prior nonconvex regularizer as $\int \phi\,(|\nabla u|) dx$, here, ϕ is a nonconvex function, namely,

$$\left\{ \begin{array}{c} \phi_1\,(|t|) = \frac{|t|^p}{1 + \rho |t|^p} \\ \phi_2\,(|t|) = \frac{1}{\rho} \log\left(1 + \rho |t|^p\right) \end{array} \right. \tag{34}$$

for $p \geq 1$.

Enven though the existence and uniqueness of solution of nonconvex regularizing optimization is still an open problem, which can preserve the more detail information of image during image processing. Therefore, we propose the following model (based on the Shi-model),

$$\hat{u} = \arg\min_u \left\{ \begin{array}{c} \lambda \| f - Au \|_1 + \frac{\mu}{2} \| f - Au \|_2^2 \\ + \frac{\alpha}{2} \| \nabla u\,(x) \|_2^2 + \int_\Omega \phi\,(|\nabla u|)\, dx \end{array} \right\} \tag{35}$$

here, the Gibbs prior $p\,(u) = \exp\left(-\gamma \int_\Omega \phi\,(u\,(x))\right) dx = \int \phi\,(|\nabla u|) dx$, and ϕ is the nonconvex function defined in (34).

In the following subsections, we present the optimization algorithm to solve the proposed model (35), which is based on a variable splitting technique [67]. To deal with the nonconvex regularizer, a multistage convex relaxation method is used.

We exploit the algorithmic framework of the variable splitting technique [67] to solve the unconstrained problem was described by (35). To obtain variable splitting, we introduce the auxiliary variables h and d under the constraints $h = f - Au$, $\nabla u = d$ subject to

$$\begin{cases} \min_{u,h,d} \begin{Bmatrix} \int_\Omega |h|\,dx + \frac{\mu}{2} \int_\Omega |f - Au|^2 dx \\ + \frac{\alpha}{2} \|\nabla u\|_2^2 + \int_\Omega \phi\,(|d|)\,dx \end{Bmatrix} \\ s.t.\, h = f - Au, \nabla u = d \end{cases} \tag{36}$$

Next, we exploit the algorithmic framework of the Split Bregman Iteration Method (SBIM) [33] to solve (36). The additional details of the SBIM algorithm are as follows,

$$\left(u^{k+1}, h^{k+1}, d^{k+1}\right) = \arg\min_{u,d,h} \begin{Bmatrix} F\,(h, d) + \frac{\gamma_1}{2} \|h - (f - Au) - b_1^k\|_2^2 \\ + \frac{\gamma_2}{2} \|d - \nabla u - b_2^k\|_2^2 \end{Bmatrix} \tag{37}$$

$$b_1^{k+1} = b_1^k + \left(f - Au^{k+1}\right) - h^{k+1} \tag{38}$$

$$b_2^{k+1} = b_2^k + \nabla u^{k+1} - d^{k+1} \tag{39}$$

here,

$$F\,(h, d) = \begin{pmatrix} \lambda \int_\Omega |h|\,dx + \frac{\mu}{2} \int_\Omega |f - Au|^2 dx \\ + \frac{\alpha}{2} \|\nabla u\|_2^2 + \int_\Omega \phi\,(|d|)\,dx \end{pmatrix} \tag{40}$$

To solve the minimization problem (37), an alternative iteration algorithm is used. Specifically, the solution of minimization problem $\left(u^{k+1}, h^{k+1}, d^{k+1}\right)$ are defined as follows,

$$u^{k+1} \in \arg\min_u \begin{Bmatrix} \frac{\mu}{2} \|f - Au\|_2^2 + \frac{\alpha}{2} \|\nabla u\|_2^2 \\ + \frac{\gamma_1}{2} \|h - (f - Au) - b_1^k\|_2^2 \\ + \frac{\gamma_2}{2} \|d^k - \nabla u - b_2^k\|_2^2 \end{Bmatrix} \tag{41}$$

$$h^{k+1} \in \arg\min_h \left\{ \lambda \int_\Omega |h|\,dx + \frac{\gamma_1}{2} \|h - (f - Au) - b_1^k\|_2^2 \right\} \tag{42}$$

$$d^{k+1} \in \arg\min_d \left\{ \int_\Omega \phi\,(|d|)\,dx + \frac{\gamma_2}{2} \|d^k - \nabla u - b_2^k\|_2^2 \right\} \tag{43}$$

It is noteworthy that the alternative (SBIM) is equivalent to the alternative ADMM [7] when ϕ is convex.

4.1 *u*-subproblem

Due to (41), the optimal value of u^{k+1} must satisfy the following Euler-Lagrange equation,

$$\left[(\mu + \gamma_1)\, A^T Au - (\alpha + \gamma_2)\, \Delta \right] u = \begin{pmatrix} \mu A^T f + \gamma_1 A^T \left(f + b_1^k - h^k\right) \\ - \gamma_2 div\,\left(d^k - b_2^k\right) \end{pmatrix} \tag{44}$$

We can use the Gauss-Seidel method to solve this equation as in [3]. Equation (44) may also be solved efficiently in the discrete Fourier domain if we assume periodic boundary

conditions. If we define $\mathcal{F}(u)$ as the Fourier transform of u, the close-form solution of (44) can be written as follows,

$$u = \mathcal{F}^{-1}\left(\frac{\mu\mathcal{F}\left(A^T f\right) + \gamma_1\mathcal{F}\left[A^T\left(f + b_1^k - h^k\right)\right]}{(\mu + \gamma_1)A^T A - (\alpha + \gamma_2)\mathcal{F}(\Delta)} \right) \tag{45}$$

4.2 h-subproblem

Similarly, due to the minimization problem of (42), and we can get its approximate solution by the following shrinkage operator,

$$h^{k+1} = shrink\left(f + b_1^k - Au^{k+1}, \frac{\lambda}{\gamma_1}\right) \tag{46}$$

here, $shrink(s, t)$ is the shrinkage operator [67], and which was defined at each point $\alpha \in [0, 1]^2$ by

$$shrink(s_\alpha, t_\alpha) = \frac{s_\alpha}{|s_\alpha|}\max\left(|s_\alpha| - t_\alpha, 0\right) \tag{47}$$

4.3 d-subproblem

Moreover, resembling to the u-subproblem, we can obtain d^{k+1} by solving the following nonconvex minimization problem,

$$\min_d\left\{\int\phi\left(|d|\right)dx + \frac{\gamma_2}{2}\|d - s\|_2^2\right\} \tag{48}$$

here, ϕ is the nonconvex regularizer in (36), and $s = \nabla u^{k+1} + b_2^k$. To handle the nonconvex term, we apply the multistage convex relaxation technique, which was described above.Consequently, we directly use the result in [3]. Specifically, the solution of (48) can be easily obtained as follows,

$$d^{k+1} \in \arg\min_d\left\{\int_\Omega v^{k+1,l}\,|d|\,dx + \frac{\gamma_2}{2}\|d - s\|_2^2\right\} \tag{49}$$

$$d^{k+1,l+1} = shrink\left(s, \frac{v^{k+1,l}}{\gamma}\right) \tag{50}$$

$$v_1^{k+1,l+1} = \frac{1}{\left(1 + \rho\left|d^{k+1,l+1}\right|\right)^2}\ for\ \phi = \phi_1 \tag{51}$$

In this paper, we decompose the difficult optimization problem (37) into three subproblems (the u, h, and d-subproblems) based on the SBIM. In addition, all the subproblems have fast and accurate techniques for obtaining the solutions. For example, the u-subproblem can be solved efficiently by using the FFT technique or Gauss-Seidel iteration method. Furthermore, the h-subproblem and d-subproblem can be resolved successfully through applying the shrinkage operator. Eventually, the optimization procedure is summarized in Algorithm 2.

Algorithm 2 Mixed noise removal algorithm with multi-fidelity terms based on nonsmooth and nonconvex regularization.

---

   initialization:
       (1) Set parameters $\lambda > 0$, $\mu > 0$, $\alpha > 0$, $\gamma_1 > 0$, $\gamma_2 > 0$;
       (2) Initialize u^0 via median filtering, and set $h^0 = b_1^0 = 0$, $d^0 = b_2^0 = 0$, $v = 1$;
   for $t = 1, 2, 3, ..., T$ **do**
      for $k = 1, 2, 3, ..., K$ **do**
          (1) Update h via (46);
          (2) Update d via (50);
          (3) Update b_1, b_2 via (38) and (39);
          (4) Update u via (45);
          (5) Output $u^{(t)} = u^{(k)}$, if $k = K$
      end
      (1) Group similar patches for each u_j using $u^{(t)}$;
      (2) Update λ, μ, α, γ_1, γ_2 with ρ ($0 < \rho < 1$);
      (3) Output the reconstruction image $\hat{u} = u^{(t)}$, if $t = T$;
   end

---

5 Experimental results

In this section, we present the results of our application of a nonconvex and nonsmooth mixed noise removal algorithm to the image restoration problem. Firstly, we employ a genetic algorithm aims to select the optimal parameters of the model. Secondly, we compare the noise removal effects of other models on multiple different levels noise images.

Because the proposed model with regularizer is nonconvex and nonsmooth, therefore, we could not give analysis convergence of the algorithm with an increasing number of iteration steps (Fig. 1b). Accordingly, we use the number of iterations of the algorithm as our termination condition. Additionally, we calculate the relationship between the number of algorithm iteration steps and the optimal ISNR on 100 images and chose to terminate with 16 iterations (Fig. 1a).

Moreover, we compare the models TVL1 [13], LRTM [20] and SHI [43] by applying them to several images, e.g., **Parrots, Monarch, Boats, Cameraman, House and Lena** (Fig. 2). Furthermore, we compare the models' ISNR and PSNR values after the image was denoised. The peak signal to noise ratio (PSNR) and improvement signal to noise ratio (ISNR) were chosen as the quantitative measures of image quality.

We utilize the improvement in the signal-to-noise ratio (ISNR) and the Peak Signal to Noise Ratio (PSNR) to evaluate the effectiveness of all methods for removing the noise, and the two metrics can be defined as follows,

$$ISNR\left(u_*, \hat{u}, \tilde{u}\right) = 10\log_{10}\left(\frac{\sum\left(u_* - \tilde{u}\right)^2}{\sum\left(u_* - \hat{u}\right)^2}\right)$$

$$PSNR\left(u_*, \hat{u}\right) = 10\log_{10}\left(\frac{255^2 mn}{\left\|u_* - \hat{u}\right\|_2^2}\right)$$

here, $u_* \in R^{m \times n}$ is the clean image, $\hat{u} \in R^{m \times n}$ is the restored image, and $\tilde{u} \in R^{m \times n}$ is the contaminated image. Note that a high value (ISNR and PSNR) indicates a better restored result.
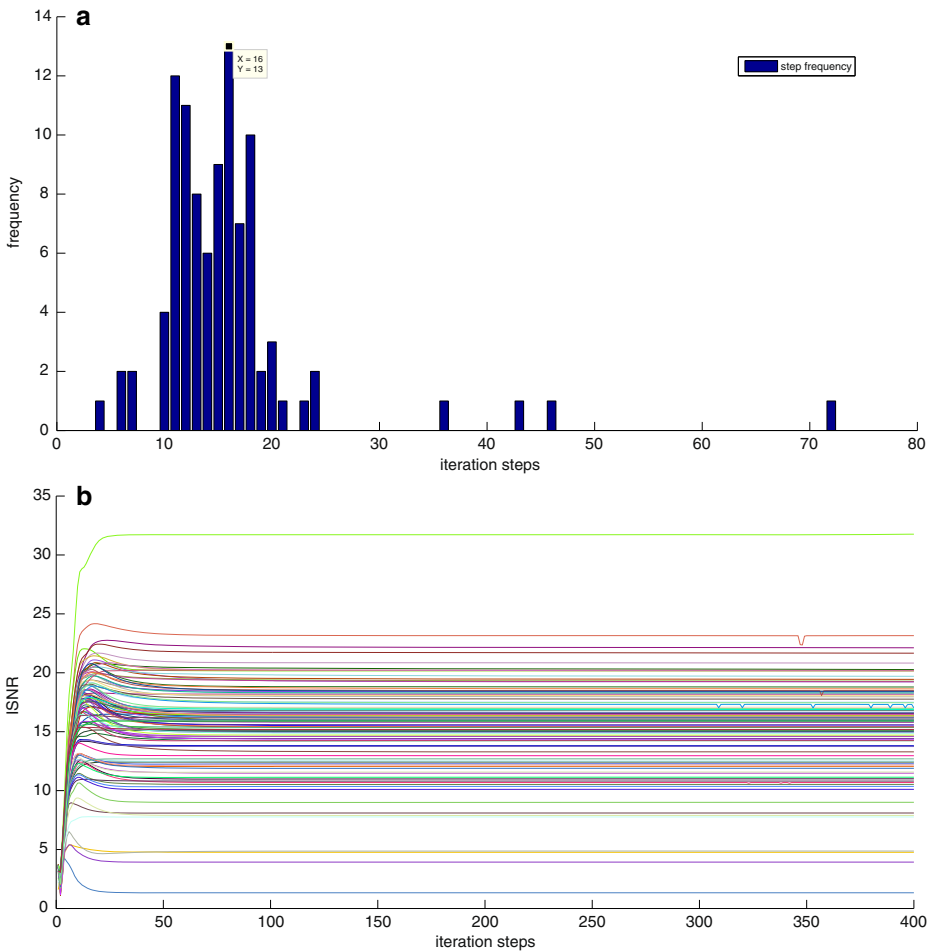
Fig. 1 a The relationship between the number of algorithm iteration steps and frequency of convergence; **b** The relationship between the number of algorithm iteration steps and the optimal ISNR on 100 images

5.1 Parameter selection

There are a total of 5 required parameters in our algorithm, and there are two methods with which to select them. In the first method, the parameters were chosen by their experience. In the second method, one fixes the values of one or more of the parameters and varies the values of other parameters. We used a genetic algorithm to find the best parameters. In the proposed model, the parameters (a, b, c, and d) should be satisfied the following conditions, namely, $a > 0$, $b < 0$, and $d > 0$. Our goal is to choose one of the best points in the parameter space. Therefore, we chose our parameters to maximize the average of the ISNR
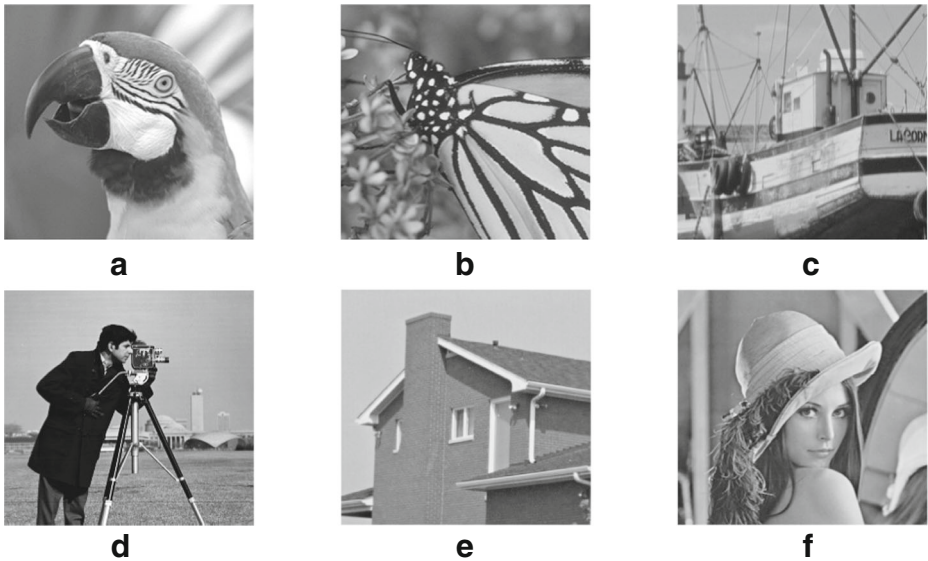
Fig. 2 The six classic Ground Truth (**a**–**f**): **a** is the Parrot, **b** is the Monarch, **c** is the Boat, **d** is the Cameraman, **e** is the House, **f** is the Lena

and PSNR values on the training set P, and such that the objective function of the genetic algorithm is defined by,

$$\begin{cases} O\left(a,b,c,d\right) = \sum_{p \in P} ISNR\left(M\left(p|a,b,c,d\right)\right) \\ O\left(a,b,c,d\right) = \sum_{p \in P} PSNR\left(M\left(p|a,b,c,d\right)\right) \\ \qquad P = \{p_1, p_2, p_3, p_4, ..., p_n\} \end{cases} \tag{52}$$

here, p_i is an image in training set, and a, b, c, and d are the model parameters that need to be selected. M represents the proposed model, and O is the objective function of genetic algorithm.

In the genetic algorithm, the number of chromosomes is 20, the heritability is 0.85, and the number of reproduction generation is 400. The following figure is the result of the parameter selection with 0.1 Gauss noise (Fig. 3).

According to the result of the operation, we choose the optimal parameters as follows, $a = 0.01$, $b = 0.20$,... It needs to be explained that in the following experiments, we used this same method to select the optimal parameters for different types of noise.

5.2 Comparison of pure noise experiments

In this subsection, we compare the denoising effects of the different models on pure noise images. Moreover, we compare Gauss noise with salt and pepper noise with the noise levels set at 0.01, 0.1, and 0.4, respectively. The denoising ISNR and PSNR results of the four models for different images are shown in Tables 1 and 2. The GN0.01 denoising visualization effect of **Parrots, Monarch, Boats, Cameraman, House and Lena** have been shown in Fig. 4. And The SP0.1 denoising visualization effect has been shown in Fig. 5.
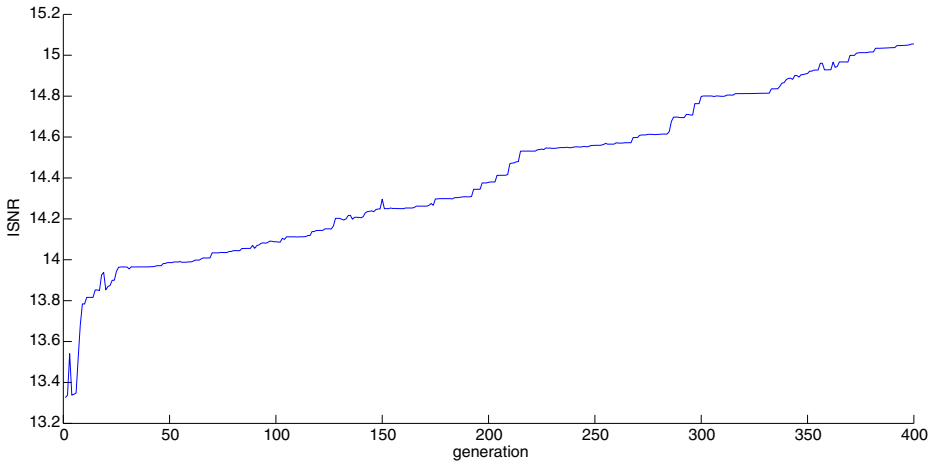
Fig. 3 The relationship between generation and ISNR parameter selection with 0.1 Gauss noise

Table 1 The denoising ISNR and PSNR results of the four models for different images with GN

| Images	Methods	GN0.01 ISNR|PSNR|TIME	GN0.1 ISNR|PSNR|TIME	GN0.4 ISNR|PSNR|TIME
parrots	ours	**8.31|76.61|1.77**	**2.4|67.96|1.97**	**0.19|58.04|2.02**
	SHI model [43]	8.09|76.38|10.14	2.27|67.82|8.89	0.06|56.95|7.67
	TVL1 [13]	5.02|73.32|8.67	1.45|67|8.91	−0.14|56.2|6.34
	LRTM [20]	7.08|75.37|9.14	2.34|67.94|8.94	0.15|57.22|9.19
monarch	ours	**7.49|75.68|1.8**	**2.31|67.59|1.81**	**0.18|56.93|1.77**
	SHI model [43]	7.17|75.36|9.91	2.2|67.48|9.13	0.04|56.72|7.88
	TVL1 [13]	4.69|72.87|8.05	1.68|66.96|6.88	−0.14|56.31|5.31
	LRTM [20]	6.32|74.47|8.98	2.21|67.47|8.88	0.12|56.76|8.83
boats	ours	**7.59|75.8|1.77**	2.31|67.5|1.72	**0.13|57.22|1.75**
	SHI model [43]	7.38|75.58|9.56	2.26|67.46|8.42	−0.09|55.93|8.17
	TVL1 [13]	5.24|73.44|6.97	1.78|66.97|7.22	−0.41|55.74|5.72
	LRTM [20]	6.49|74.68|9.23	2.36|67.54|8.92	0.11|56.98|8.8
cameraman	ours	**7.04|75.41|1.64**	**2.28|67.48|1.91**	**0.13|57.11|1.83**
	SHI model [43]	6.6|74.98|9.95	2.18|67.38|8.88	−0.05|56.72|7.59
	TVL1 [13]	3.9|72.28|9.8	1.61|66.82|8.63	−0.28|56.51|8.55
	LRTM [20]	6|74.39|9.13	2.27|67.46|8.63	0.12|56.92|8.56
house	ours	**9.71|77.9|1.81**	**2.56|67.83|1.7**	**0.17|57.31|1.91**
	SHI model [43]	9.66|77.84|9.19	2.51|67.79|9.08	0.05|57.15|6.61
	TVL1 [13]	8.28|76.47|9.86	2.3|67.57|8.8	−0.08|57.03|7.7
	LRTM [20]	8.11|76.27|9.34	2.53|67.84|8.89	0.17|57.25|8.58
Lena	ours	**7.82|75.99|1.84**	2.36|67.59|1.66	**0.15|56.84|1.75**
	SHI model [43]	7.63|75.8|9.83	2.31|67.54|8.81	0.02|56.8|7.89
	TVL1 [13]	5.83|74|6.7	1.98|67.21|8.13	-0.19|56.5|7.95
	LRTM [20]	6.74|74.91|9.25	2.37|67.61|8.78	0.15|56.82|8.84

Table 2 The denoising ISNR and PSNR results of the four models for different images with SP noise

Images	Methods	SP0.01 ISNR|PSNR|TIME	SP0.1 ISNR|PSNR|TIME	SP0.4 ISNR|PSNR|TIME
parrots	ours	**13.21|86.57|1.83**	**18.47|81.75|1.88**	**17.23|74.56|1.75**
	SHI model [43]	10.2|83.55|25.19	17.02|80.3|25.03	14.76|72.08|21.03
	TVL1 [13]	1.67|75.03|6.14	11.11|74.39|7.3	14.86|72.18|6.88
	LRTM [20]	8.52|81.76|9.3	12.8|76.21|9.48	15.49|72.81|9.23
monarch	ours	**14.93|87.99|1.7**	**17.33|80.75|1.88**	**15.85|73.21|1.86**
	SHI model [43]	11.44|84.5|24.67	15.69|79.11|25.23	13.12|70.48|21.84
	TVL1 [13]	3.23|76.29|6.09	11.15|74.57|6.89	13.25|70.61|6.45
	LRTM [20]	8.55|82.15|8.7	13.35|76.79|10.11	13.78|71.13|8.98
boats	ours	**20.91|94.69|1.7**	**20.07|83.61|1.77**	**18.35|75.89|1.64**
	SHI model [43]	12.79|86.57|24.48	17|80.54|22.28	13.98|71.52|21.31
	TVL1 [13]	2.94|76.72|7.27	12.33|75.86|7.64	15.48|73.02|7.66
	LRTM [20]	9.16|82.5|9.27	14.23|77.7|9.59	15.68|73.21|9.3
cameraman	ours	**8.06|81.23|1.61**	**14.81|78.01|1.73**	**15.01|72.24|1.81**
	SHI model [43]	5.91|79.07|28.5	13.51|76.71|27.13	13.41|70.65|23.5
	TVL1 [13]	0.13|73.3|6.83	9.58|72.78|7.3	13.61|70.84|7.44
	LRTM [20]	5.99|78.66|8.73	10.99|74.11|10.02	14|71.22|8.84
house	ours	**21.54|95.12|1.58**	**21.56|85.2|1.89**	**20.41|78|1.86**
	SHI model [43]	14.22|87.8|24.48	18.96|82.61|26.52	15.11|72.71|21.41
	TVL1 [13]	6.78| 80.36|6.06	15.84|79.49|7.53	18.64|76.23|7.38
	LRTM [20]	10.66|83.94|8.91	16.56|80.22|9.67	17.73|75.3|9.05
Lena	ours	**16.79|90.17|1.91**	**19.62|83.22|1.89**	**18.34|75.93|1.69**
	SHI model [43]	11.53|84.91|25.89	17.08|80.68|23.3	14.6|72.19|21.13
	TVL1 [13]	3.86|77.24|6.98	12.87|76.47|6.36	15.99|73.58|6.14
	LRTM [20]	8.23|81.82|8.98	13.55|77.05|9.55	15.84|73.37|8.92

The results in Tables 1 and 2 illustrate that our proposed model algorithm is superior to other models in terms of: running time, indices of Gaussian and salt-and-pepper noise images, and the local visualization quality after restoration.

5.3 Comparison of mixed noise experiments

In this subsection, in order to verify the robustness of the proposed model, we perform an experimental comparison on images with mixed noise. Where the mixed noise is a mixture of Gaussian noise and salt and pepper noise, the noise levels are GN0.01+SP0.01, GN0.01+SP0.1, GN0.1+SP0.01, and GN0.1+SP0.1, respectively. The experimental results are shown in Table 3, and the effective images before and after denoising of the different models are shown in Fig. 5.
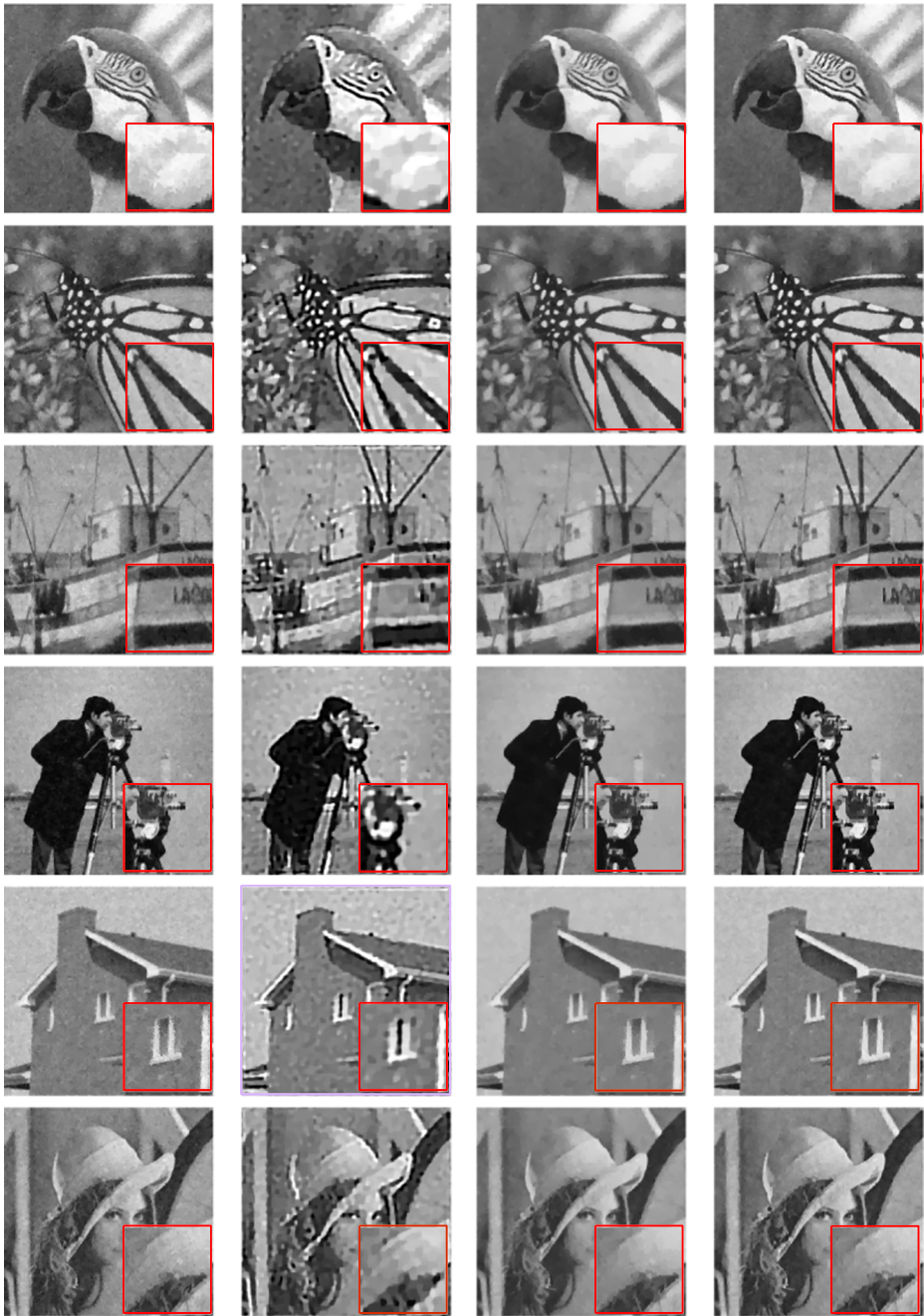
Fig. 4 Denoising effect of noisy images (GN0.01): The first column is the LRTM, the second column is the TVL1, the third column is the SHI model, and the fourth column is Ours
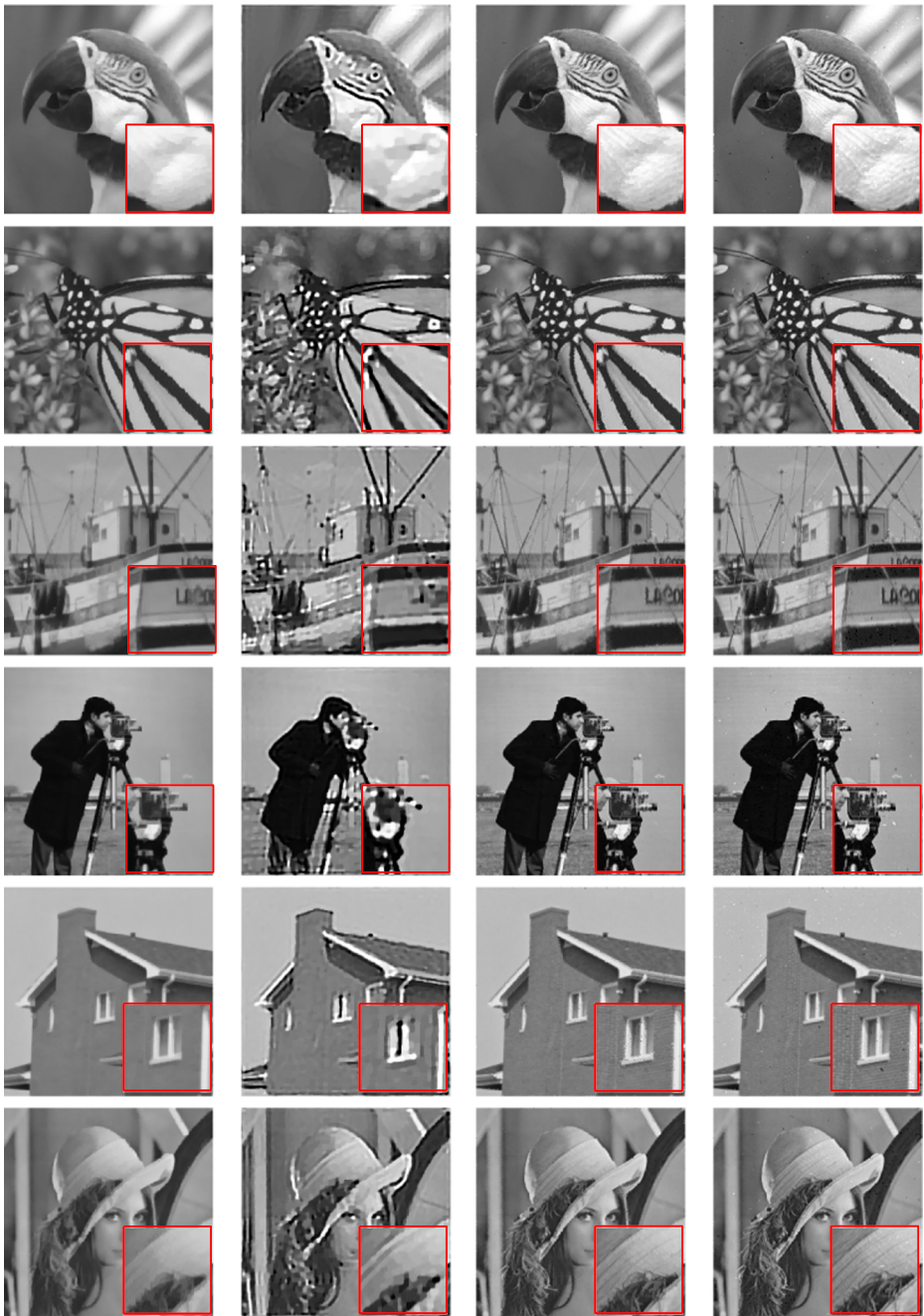
Fig. 5 Denoising effect of noisy images (SP0.1): The first column is the LRTM, the second column is the TVL1, the third column is the SHI model, and the fourth column is Ours

Table 3 The denoising ISNR and PSNR results of the four models for different images with MIXED noise

Images	Methods	GN0.01+SP0.01 ISNR|PSNR|TIME	GN0.01+SP0.1 ISNR|PSNR|TIME	GN0.1+SP0.1 ISNR|PSNR|TIME
parrots	ours	**9.04|76.18|2.02**	**12.67|74.82|1.98**	**3.04|67.91|1.86**
	SHI model [43]	7.58|74.72|10.19	10.6|72.75|15.75	2.4|67.27|14.91
	TVL1 [13]	6.04|73.18|9.17	10.67|72.83|8.86	2.14|67.01|5.78
	LRTM [20]	8.01|75.36|9.3	11.37|73.98|9.17	2.84|67.9|8.64
monarch	ours	**8.06|75.21|1.88**	**11.5|73.78|1.81**	**2.83|67.56|1.73**
	SHI model [43]	6.98|74.13|9.97	9.78|72.06|15.34	2.25|66.98|14.58
	TVL1 [13]	5.65|72.8|6.58	9.62|71.9|7.72	2.2|66.93|8.38
	LRTM [20]	7.04|74.26|8.8	9.97|72.62|9.3	2.7|67.47|8.66
boats	ours	**8.18|75.37|1.83**	**11.99|74.37|1.83**	**2.79|67.45|1.78**
	SHI model [43]	7.26|74.45|9.52	9.97|72.35|16.44	2.37|67.03|14.58
	TVL1 [13]	6.21|73.39|8.05	10.5|72.89|8.02	2.32|66.98|6.59
	LRTM [20]	7.35|74.56|9.16	10.52|73.35|9.08	2.75|67.48|8.88
cameraman	ours	**7.84|75|1.75**	**11.26|73.46|1.81**	**2.76|67.39|1.77**
	SHI model [43]	6.41|73.57|9.83	9.71|71.91|16.22	2.32|66.96|14.8
	TVL1 [13]	4.95|72.11|9.3	9.55|71.75|8.92	2.12|66.76|8.05
	LRTM [20]	6.8|74.12|9.08	10.22|72.8|9.34	2.73|67.43|8.81
house	ours	**10.26|77.36|1.75**	**13.79|76.32|1.94**	**3.06|67.82|1.69**
	SHI model [43]	8.92|76.02|9.08	10.72|73.25|14.88	2.44|67.2|14.95
	TVL1 [13]	9.22|76.32|9.22	13.58|76.11|10.55	2.81|67.57|9.05
	LRTM [20]	8.76|75.97|8.95	11.85|74.73|9.28	2.97|67.83|8.69
Lena	ours	**8.56|75.58|1.63**	**12.12|74.51|1.75**	**2.88|67.55|1.78**
	SHI model [43]	7.52|74.54|9.84	10.07|72.46|17.06	2.4|67.07|15.05
	TVL1 [13]	6.9|73.92|8.28	11.01|73.4|6.61	2.52|67.19|6.98
	LRTM [20]	7.63|74.75|9.13	10.81|73.7|8.98	2.82|67.55|8.95

It can be seen from Table 3 that the proposed model algorithm is superior to the other three models in terms of running time and indexes on images with mixed noise, and the quality of local visualization after restoration is better than that of other models (Fig. 6).

6 Conclusion

In this paper, we proposed a nonconvex and nonsmooth regular image mixed noise removal algorithm based on the Shi model by solving the nonconvex minimization problem. we use the multistage nonconvex relaxation technique, which aims to deal with the nonconvex term. Furthermore, we employ a genetic algorithm to select the optimal parameters for the model and set the number of iteration steps as the termination condition for the algorithm. Several, experiments on different images with different noise levels illustrate that the model's robustness, running time, ISNR and PSNR are better than the other three models. Additionally, our algorithm can maintain the local information of the images with better visual quality.
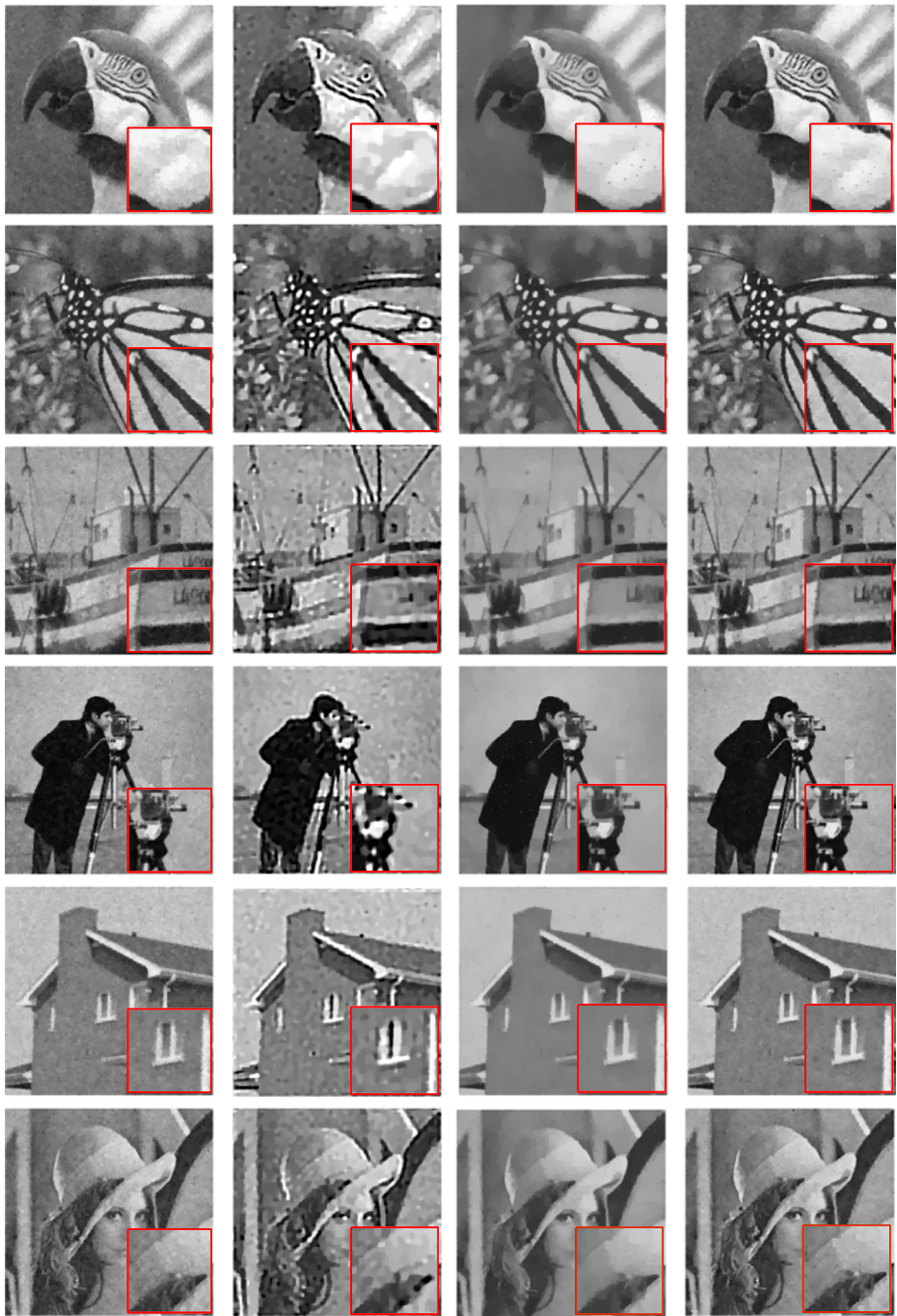
Fig. 6 Denoising effect of noisy images (GN0.01+SP0.01): The first column is the LRTM, the second column is the TVL1, the third column is the SHI model, and the fourth column is Ours

References

1. Alliney S (1997) A property of the minimum vectors of a regularizing functional defined by means of the absolute norm. IEEE Trans Signal Process 45(4):913–917
2. Aubert G, Kornprobst P (2006) Mathematical problems in image processing: partial differential equations and the calculus of variations, vol. 147. Springer Science & Business Media
3. Aubert G, Aujol J-F (2008) A variational approach to removing multiplicative noise. SIAM J Appl Math 68(4):925–946
4. Bar L, Chan TF, Chung G, Jung M, Kiryati N, Mohieddine R, Sochen N, Vese LA (2011) Mumford and shah model and its applications to image segmentation andimage restoration. In: Handbook of mathematical methods in imaging. Springer, pp 1095–1157
5. Blake A, Zisserman A (1987) Visual reconstruction. MIT Press, Cambridge
6. Bovik AC (2010) Handbook of image and video processing. Academic, New York
7. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J et al (2011) Distributed optimization and statistical learning via the alternating direction method of multipliers. Found Trends® Machine Learn 3(1):1–122
8. Buades A, Coll B, Morel J-M (2005) A non-local algorithm for image denoising. In: IEEE computer society conference on computer vision and pattern recognition, 2005. CVPR 2005, vol 2. IEEE, pp 60–65
9. Buades A, Coll B, Morel J-M (2005) A review of image denoising algorithms, with a new one. Multiscale Model Simul 4(2):490–530
10. Burger HC, Schuler CJ, Harmeling S (2012) Image denoising: can plain neural networks compete with bm3d? In: 2012 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 2392–2399
11. Cai J-F, Chan RH, Nikolova M (2008) Two-phase approach for deblurring images corrupted by impulse plus gaussian noise. Inverse Probl Imaging 2(2):187–204
12. Cai X, Chan R, Zeng T (2013) A two-stage image segmentation method using a convex variant of the mumford–shah model and thresholding. SIAM J Imag Sci 6(1):368–390
13. Chan TF, Esedoglu S (2005) Aspects of total variation regularized l 1 function approximation. J SIAM Appl Math 65(5):1817–1837
14. Charbonnier P, Blanc-Féraud L, Aubert G, Barlaud M (1997) Deterministic edge-preserving regularization in computed imaging. IEEE Trans Image Process 6(2):298–311
15. Chartrand R, Yin W (2008) Iterative reweighted algorithms for compressive sensing. Tech. Rep.
16. Chen Y, Pock T (2017) Trainable nonlinear reaction diffusion: a flexible framework for fast and effective image restoration. IEEE Trans Pattern Anal Mach Intell 39(6):1256–1272
17. Chen Y, Yu W, Pock T (2015) On learning optimized reaction diffusion processes for effective image restoration. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 5261–5269
18. Dabov K, Foi A, Katkovnik V, Egiazarian K (2007) Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans Image Process 16(8):2080–2095
19. Daubechies I, DeVore R, Fornasier M, Güntürk CS (2010) Iteratively reweighted least squares minimization for sparse recovery. Commun Pure Appl Math: A Journal Issued by the Courant Institute of Mathematical Sciences 63(1):1–38
20. Deng L, Zhu H, Li Y, Yang Z (2018) A low-rank tensor model for hyperspectral image sparse noise removal. IEEE Access 6:62120–62127

21. Dong W, Li X, Zhang L, Shi G (2011) Sparsity-based image denoising via dictionary learning and structural clustering. In: 2011 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 457–464
22. 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
23. Dong W, Zhang L, Shi G, Li X (2013) Nonlocally centralized sparse representation for image restoration. IEEE Trans Image Process 22(4):1620–1630
24. Dong W, Shi G, Li X, Ma Y, Huang F (2014) Compressive sensing via nonlocal low-rank regularization. IEEE Trans Image Process 23(8):3618–3632
25. Elad M, Aharon M (2006) Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans Image Process 15(12):3736–3745
26. Esser E (2009) Applications of lagrangian-based alternating direction methods and connections to split bregman. CAM Report 9:31
27. Feng W, Qiao P, Chen Y (2018) Fast and accurate poisson denoising with trainable nonlinear diffusion. IEEE Trans Cybern 48(6):1708–1719
28. Figueiredo MA, Bioucas-Dias JM (2010) Restoration of poissonian images using alternating direction optimization. IEEE Trans Image Process 19(12):3133–3145
29. Garnett R, Huegerich T, Chui C, He W (2005) A universal noise removal algorithm with an impulse detector. IEEE Trans Image Process 14(11):1747–1754
30. Geman S, Geman D (1984) Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. IEEE Trans Pattern Anal Mach Intell 6:721–741
31. Geman D, Reynolds G (1992) Constrained restoration and the recovery of discontinuities. IEEE Trans Pattern Anal Mach Intell 3:367–383
32. Geman D, Yang C (1995) Nonlinear image recovery with half-quadratic regularization. IEEE Trans Image Process 4(7):932–946
33. Goldstein T, Osher S (2009) The split bregman method for l1-regularized problems. SIAM J Imag Sci 2(2):323–343
34. Gu S, Zhang L, Zuo W, Feng X (2014) Weighted nuclear norm minimization with application to image denoising. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2862–2869
35. Han Y, Feng X-C, Baciu G, Wang W-W (2013) Nonconvex sparse regularizer based speckle noise removal. Pattern Recognit 46(3):989–1001
36. He B, Liao L-Z, Han D, Yang H (2002) A new inexact alternating directions method for monotone variational inequalities. Math Program 92(1):103–118
37. He K, Wang R, Tao D, Cheng J, Liu W (2018) Color transfer pulse-coupled neural networks for underwater robotic visual systems. IEEE Access 6:32850–32860
38. Hintermüller M, Langer A (2012) Subspace correction methods for a class of non-smooth and non-additive convex variational problems in image processing
39. Huang T, Dong W, Xie X, Shi G, Bai X (2017) Mixed noise removal via laplacian scale mixture modeling and nonlocal low-rank approximation. IEEE Trans Image Process 26(7):3171–3186
40. Jain V, Seung S (2009) Natural image denoising with convolutional networks. In: Advances in neural information processing systems, pp 769–776
41. Ji H, Huang S, Shen Z, Xu Y (2011) Robust video restoration by joint sparse and low rank matrix approximation. SIAM J Imag Sci 4(4):1122–1142
42. Ji H, Liu C, Shen Z, Xu Y (2010) Robust video denoising using low rank matrix completion
43. Jia T, Shi Y, Zhu Y, Wang L (2016) An image restoration model combining mixed L1/L2 fidelity terms. J Vis Commun Image Represent 38:461–473
44. Jiang J, Zhang L, Yang J (2014) Mixed noise removal by weighted encoding with sparse nonlocal regularization. IEEE Trans Image Process 23(6):2651–2662
45. Jung M, Kang M (2015) Efficient nonsmooth nonconvex optimization for image restoration and segmentation. J Sci Comput 62(2):336–370
46. Li SZ (1994) Markov random field models in computer vision. In: European conference on computer vision. Springer, pp 361–370
47. Liu L, Chen L, Chen CP, Tang YY et al (2017) Weighted joint sparse representation for removing mixed noise in image. IEEE Trans Cybern 47(3):600–611
48. Mairal J, Bach F, Ponce J, Sapiro G, Zisserman A (2009) Non-local sparse models for image restoration. In: 2009 IEEE 12th international conference on computer vision. IEEE, pp 2272–2279
49. Mumford D, Shah J (1989) Optimal approximations by piecewise smooth functions and associated variational problems. Commun Pure Appl Math 42(5):577–685

50. Nikolova M (1999) Markovian reconstruction using a gnc approach. IEEE Trans Image Process 8(9):1204–1220
51. Nikolova M (2002) Minimizers of cost-functions involving nonsmooth data-fidelity terms. application to the processing of outliers. SIAM J Numer Anal 40(3):965–994
52. Nikolova M (2004) A variational approach to remove outliers and impulse noise. J Math Imaging Vision 20(1–2):99–120
53. Nikolova M (2005) Analysis of the recovery of edges in images and signals by minimizing nonconvex regularized least-squares. Multiscale Model Simul 4(3):960–991
54. Nikolova M, Chan RH (2007) The equivalence of half-quadratic minimization and the gradient linearization iteration. IEEE Trans Image Process 16(6):1623–1627
55. Nikolova M, Ng MK, Zhang S, Ching W-K (2008) Efficient reconstruction of piecewise constant images using nonsmooth nonconvex minimization. SIAM J Imag Sci 1(1):2–25
56. Nikolova M, Ng MK, Tam C-P (2010) Fast nonconvex nonsmooth minimization methods for image restoration and reconstruction. IEEE Trans Image Process 19(12):3073–3088
57. Ochs P, Dosovitskiy A, Brox T, Pock T (2013) An iterated l1 algorithm for non-smooth non-convex optimization in computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1759–1766
58. Osher S, Shi Z, Zhu W (2017) Low dimensional manifold model for image processing. SIAM J Imag Sci 10(4):1669–1690
59. Robini MC, Lachal A, Magnin IE (2007) A stochastic continuation approach to piecewise constant reconstruction. IEEE Trans Image Process 16(10):2576–2589
60. Rockafellar R (1997) Convex analysis, Princeton University Press, Princeton, 1970. MATH Google Scholar
61. Roth S, Black MJ (2009) Fields of experts. Int J Comput Vis 82(2):205–229
62. Rudin LI, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Physica D: Nonlinear Phenomena 60(1–4):259–268
63. Shao L, Yan R, Li X, Liu Y (2014) From heuristic optimization to dictionary learning: a review and comprehensive comparison of image denoising algorithms. IEEE Trans Cybern 44(7):1001–1013
64. Tao D, Cheng J, Song M, Lin X (2016) Manifold ranking-based matrix factorization for saliency detection. IEEE Trans Neural Netw Learn Syst 27(6):1122–1134
65. Teboul S, Blanc-Feraud L, Aubert G, Barlaud M (1998) Variational approach for edge-preserving regularization using coupled pdes. IEEE Trans Image Process 7(3):387–397
66. Vogel CR, Oman ME (1998) Fast, robust total variation-based reconstruction of noisy, blurred images. IEEE Trans Image Process 7(6):813–824
67. Wang Y, Yang J, Yin W, Zhang Y (2008) A new alternating minimization algorithm for total variation image reconstruction. SIAM J Imag Sci 1(3):248–272
68. Xiao Y, Zeng T, Yu J, Ng MK (2011) Restoration of images corrupted by mixed gaussian-impulse noise via l1–l0 minimization. Pattern Recognit 44(8):1708–1720
69. Xie J, Xu L, Chen E (2012) Image denoising and inpainting with deep neural networks. In: Advances in neural information processing systems, pp 341–349
70. Yan M (2013) Restoration of images corrupted by impulse noise and mixed gaussian impulse noise using blind inpainting. SIAM J Imag Sci 6(3):1227–1245
71. Yan R, Shao L, Liu Y (2013) Nonlocal hierarchical dictionary learning using wavelets for image denoising. IEEE Trans Image Process 22(12):4689–4698
72. Zhang T (2010) Analysis of multi-stage convex relaxation for sparse regularization. J Mach Learn Res 11:1081–1107
73. Zhang X, Lu Y, Chan T (2012) A novel sparsity reconstruction method from poisson data for 3d bioluminescence tomography. J Sci Comput 50(3):519–535
74. Zhang H, Yang J, Zhang Y, Huang TS (2013) Image and video restorations via nonlocal kernel regression. IEEE Trans Cybern 43(3):1035–1046
75. Zhang K, Zuo W, Chen Y, Meng D, Zhang L (2017) Beyond a gaussian denoiser: residual learning of deep cnn for image denoising. IEEE Trans Image Process 26(7):3142–3155

Chun Li studies optimization and machine learning. His work is focused on Semantic segmentation about medical images and computer vision.


Yuepeng Li studies machine learning. His work is focused on automated and distributing optimization routines for model fitting and data science.

Zhicheng Zhao received the B.E. degree from National Pilot School of software, Yunnan University, Yunnan China in 2016. He is currently pursuing the PHD degree at Computer Network Information Center, Chinese Academy of Science. His main research interests include deep learning, fine-grained image recognition.

Longlong Yu studies machine learning. His work is focused on the application of machine learning in ecological agriculture.
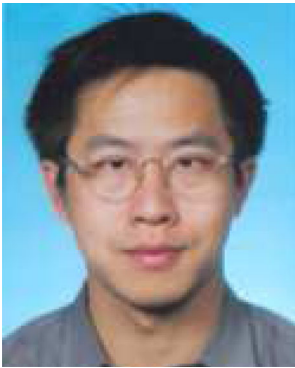
Ze Luo main research areas include data grids, data mining for spatio-temporal patterns, and eScience applications. In 2012, supported by the Chinese Academy of Sciences, as a visiting scholar, he went to the USGS Western Ecology Research Center and the Patuxent Wildlife Research Center of the US Geological Survey for a one-year visit. Based on massive data analysis, the impact of global climate change on the spatial distribution and migratory routes of migratory birds was studied.