Cauchy noise removal using group-based low-rank prior

Meng Ding [a], Ting-Zhu Huang [a,*], Tian-Hui Ma [b], Xi-Le Zhao [a,*], Jing-Hua Yang [c]

[a] *School of Mathematical Sciences/Research Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, PR China*
[b] *School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, PR China*
[c] *Department of Arts and Sciences, Chengdu College of University of Electronic Science And Technology of China, Chengdu, Sichuan 611731, PR China*

A B S T R A C T

Although the extensive research on Gaussian noise removal, few works consider the Cauchy noise removal problem. In this paper, we propose a novel group-based low-rank method for Cauchy noise removal. By exploiting the nonlocal self-similarity of natural images, we consider a group of similar patches as an approximate low-rank matrix, and formulate the denoising of each group as a low-rank matrix recovery problem. Meanwhile, we develop the alternating direction method of multipliers algorithm to solve the proposed nonconvex model with guaranteed convergence. Experiments illustrate that our method has superior performance over the state-of-the-art methods in terms of both visual and quantitative measures.

© 2019 Elsevier Inc. All rights reserved.

1. Introduction

As a fundamental problem in image processing, image denosing is an important preprocessing step of many image restoration tasks, such as super-resolution [45], inpainting [18], and deblurring [23]. Most existing image denoising techniques assume that images are corrupted by additive Gaussian noise, but there are many non-Gaussian noises in practical applications, such as impulsive noise [6] and multiplicative noise [14]. In particular, the Cauchy noise, as a kind of α-stable noise, widely arises in many real-world data, such as underwater acoustic signals [2], atmospheric radio signals [39], and radar clutter [24]. Recently, the problem of Cauchy noise removal has attracted much attention in image processing [35,38]. Mathematically, the observed image is modeled as

$$\mathbf{G}_{noi} = \mathbf{F}_{ori} + \mathbf{N}_{noi}, \tag{1}$$

where $\mathbf{G}_{noi} \in \mathbb{R}^{m_1 \times m_2}$ is the noisy image, $\mathbf{F}_{ori}$ is the original image, and $\mathbf{N}_{noi}$ is the noise following the Cauchy distribution whose probability density function is

$$p(x) = \frac{\gamma}{\pi (\gamma^2 + (x - \delta)^2)},$$

where δ is the localization parameter corresponding to the median of the distribution and $\gamma > 0$ is the scale parameter deciding the spread of the distribution around δ. Without loss of generality, in the following discussion, we consider $\delta = 0$.

Image denoising is a typical inverse problem; stable solution processes usually rely on the prior knowledge of the underlying image. One popular and efficient prior knowledge is piecewise smoothness, which is usually characterized by total variation (TV) [1,11,37,41,43,47,53], i.e.,

$$\|\mathbf{F}\|_{TV} = \|\nabla\mathbf{F}\|_2 = \sum_{1 \le p \le m_1, 1 \le q \le m_2} \sqrt{|(\nabla_x\mathbf{F})_{p,q}|^2 + |(\nabla_y\mathbf{F})_{p,q}|^2}$$

with

$$(\nabla_x\mathbf{F})_{p,q} = \begin{cases} \mathbf{F}_{p+1,q} - \mathbf{F}_{p,q}, & \text{if} \quad p < m_1, \\ \mathbf{F}_{1,q} - \mathbf{F}_{m_1,q}, & \text{if} \quad p = m_1, \end{cases}$$

and

$$(\nabla_y\mathbf{F})_{p,q} = \begin{cases} \mathbf{F}_{p,q+1} - \mathbf{F}_{p,q}, & \text{if} \quad q < m_2, \\ \mathbf{F}_{p,1} - \mathbf{F}_{p,m_2}, & \text{if} \quad q = m_2, \end{cases}$$

for $p = 1, 2, \ldots, m_1$ and $q = 1, 2, \ldots, m_2$, where $\mathbf{F}_{p,q}$ denotes (p, q)th pixel of the image $\mathbf{F}$. Sciacchitano et al. [38] proposed the following TV-based Cauchy noise removal model:

$$\underset{\mathbf{F}_{\text{ori}}}{\arg\min} \frac{\lambda}{2} \langle \log(\gamma^2 + (\mathbf{F}_{\text{ori}} - \mathbf{G}_{\text{noi}})^2), \mathbf{1} \rangle + \|\mathbf{F}_{\text{ori}}\|_{TV}, \tag{2}$$

where $\lambda > 0$ is the regularization parameter controlling the balance between the fidelity term and the regularization term. Since the fidelity term derived from Cauchy distribution is non-convex, the solution of (2) depends on the initial guess. To overcome this difficulty, the authors [38] added a quadratic penalty term

$$\underset{\mathbf{F}_{\text{ori}}}{\arg\min} \frac{\lambda}{2} \big(\langle \log(\gamma^2 + (\mathbf{F}_{\text{ori}} - \mathbf{G}_{\text{noi}})^2), \mathbf{1} \rangle + \mu \|\mathbf{F}_{\text{ori}} - \mathbf{F}_0\|_F^2 \big) + \|\mathbf{F}_{\text{ori}}\|_{TV}, \tag{3}$$

where $\mu > 0$ is a penalty parameter, $\mathbf{F}_0$ is the image obtained by applying a median filter on $\mathbf{G}_{\text{noi}}$. The authors proved that (3) is strictly convex, if $8\mu\gamma^2 \ge 1$. However, the median filter usually produces poor results, so pushing $\mathbf{F}_{\text{ori}}$ close to $\mathbf{F}_0$ may lower the denoising performance. To overcome this drawback, Mei et al. [35] removed the quadratic term $\|\mathbf{F}_{\text{ori}} - \mathbf{F}_0\|_F^2$ and applied the alternating direction method of multiplies (ADMM) to solve the nonconvex model (2) with guaranteed convergence. Experiments in [35] illustrated that the superior performance of their method over the convex method [38] in details preservation. The above TV-based methods can preserve sharp image edges, but also produces the undesirable staircase effects [3,7,19,52]. For alleviating staircase effects, Ding et al. [12] proposed a method based on TV with overlapping group sparsity for recovering blurred images corrupted by Cauchy noise. In [25], Laus et al. proposed a generalized myriad filter to estimate the parameters of the Cauchy distribution and designed a nonlocal myriad filter for removing Cauchy noise.

In recent years, low-rankness has been emerged as an important prior knowledge of natural images. And methods based on low-rank matrix recovery (LRMR) have been widely studied in image processing [8,29,30,36]. LRMR-based methods aim at recovering the underlying low-rank matrix by minimizing the matrix rank. Since the direct rank minimization is NP-hard, this optimization problem is generally relaxed by the nuclear norm minimization [5]

$$\hat{\mathbf{X}} = \underset{\mathbf{X}}{\arg\min} \frac{\lambda}{2} \|\mathbf{Y} - \mathbf{X}\|_F^2 + \|\mathbf{X}\|_*, \tag{4}$$

where $\|\mathbf{X}\|_* = \sum_r (\sigma_r(\mathbf{X}))$ is the nuclear norm of matrix $\mathbf{X}$, $\sigma_r(\mathbf{X})$ is the rth largest singular value. Eq. (4) has a closed-form solution by the singular value thresholding (SVT) [5] operator of the observation matrix

$$\hat{\mathbf{X}} = \mathbf{U} S_{\frac{1}{\lambda}}(\mathbf{\Sigma}) \mathbf{V}^T, \tag{5}$$

where $\mathbf{Y} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ is the singular value decomposition of Y and $S_{\frac{1}{\lambda}}(\mathbf{\Sigma}) = \text{diag}(\max(\sigma_r - \frac{1}{\lambda}, 0))$. Zhang et al. [48] proposed a new LRMR-based hyperspectral images (HSI) mixed noise removal method, which applied the "Go Decomposition" on the matrix ordered by a patch of HSI, exploring the low-rankness of a clean HSI patch.

Another significant image prior is nonlocal self-similarity (NSS) [9,10,22], which depicts the redundancy of similar image structures. Inspired by the success of nonlocal means denoising filter [4], a number of NSS-based regularization methods have been emerged [15,31,32,50]. Dabov et al. [9] proposed the block matching and 3-D filtering (BM3D) denoising method, which performed a 3-D transform-domain collaborative filtering on groups of similar patches, exploiting simultaneously the sparsity and NSS priors. Mairal et al. [34] combined the NSS prior and sparse coding by jointly decomposing groups of similar signals on subsets of the learned dictionary for image demosaicking and denoising tasks.

Recently, as a combination of low-rankness and NSS, the group-based low-rank (GBLR) prior has achieved great success in image restoration [17,20,27,28,33]. The basic idea is that a matrix consisting with vectorized similar image patches can be approximated by a low-rank matrix. Dong [13] showed an intrinsic relationship between simultaneous sparse coding and low-rank approximation, and presented a low-rank approach toward modeling NSS in images. Zhang et al. [49] proposed a
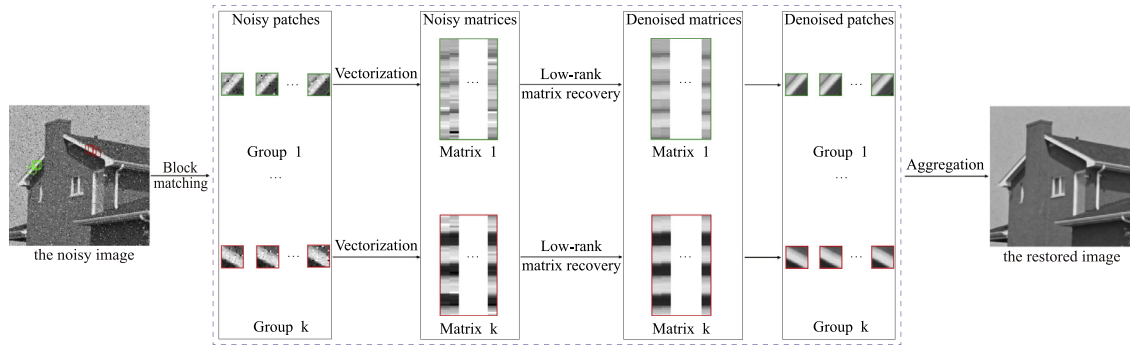
Fig. 1. Flowchart of the proposed denoising method.

group-based sparse representation method for image restoration, which exploitd the group composed of nonlocal patches with similar structures as the basic unit of sparse representation.

In this paper, we propose a novel GBLR method exploiting the NSS of images for Cauchy noise removal. We consider the group, i.e., a 2-D matrix consisting of similar image patches, as the basic unit of noise removal. The proposed method has three steps, grouping via stack similar images patches, denoising each group, aggregating the restored image via the denoised groups. The flowchart of the proposed method is illustrated in Fig. 1. Experiments show that our method achieves better denoising performance than compared methods.

In summary, our contributions are as follows:

• we propose a novel GBLR Cauchy noise removal method, where the denoising of each group is formulated as a LRMR problem;

• we develop an efficient algorithm under the framework of ADMM to solve the proposed model with guaranteed convergence.

The outline of this paper is as follows. Section 2 introduces the proposed GBLR method in details. Section 3 presents the denoising algorithm. Section 4 gives the experimental evaluation of the proposed method. Section 5 discusses the effects of parameters and the convergence behavior of the proposed algorithm. Finally, Section 6 summarizes the paper.

2. The proposed method

The proposed GBLR method involves three steps, which are presented in Algorithm 1. Below we give the details of each step.

Algorithm 1 GBLR method for Cauchy noise removal.

Step 1: Grouping: For each reference patch of the input image, stack the similar image patches

into a group.

Step 2: Denoising: For each group, obtain the denoised group by solving a GBLR model.

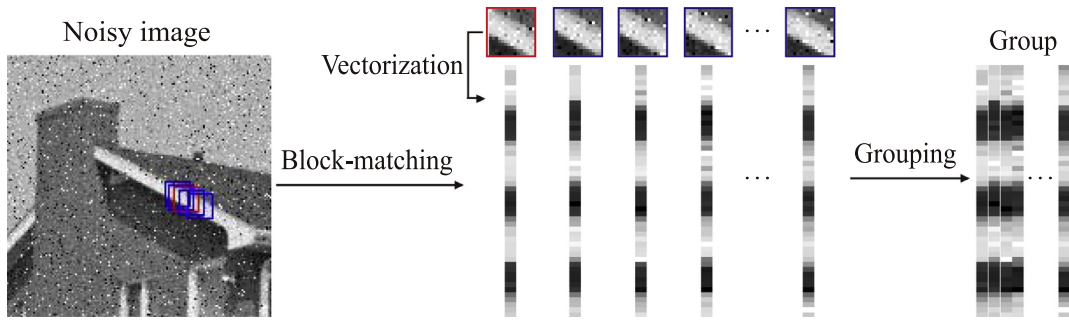**Step 3: Aggregation**: Compute the final denoised image by averaging all denoised groups.

In Step 1, we construct groups via block-matching. We use block-matching [9] to collect a group of image patches containing similar image structures. We employ the distance suggested in [25] to measure the similarity between two patches extracted from the noisy image. A smaller distance indicates higher similarity. Specifically, we initialize the locations of groups by performing block-matching on the noisy image, and update the grouping results using the current restored image to improve the accuracy of block-matching. This update occurs only once. Overlapped patches are allowed to avoid possible block effects. The grouping process is illustrated in Fig. 2.

A mathematical formulation of grouping is given as follows. For each reference patch $\mathbf{G}_{\text{ref}} \in \mathbb{R}^{s \times s}$ extracted from the noisy image $\mathbf{G}_{\text{noi}}$, we find the locations of its similar patches within a window of size $s_{\text{win}} \times s_{\text{win}}$. To measure the similarity between two patches $\mathbf{G}_1$ and $\mathbf{G}_2 \in \mathbb{R}^{s \times s}$ degraded by Cauchy noise, we use the following distance [25] derived from the noise distribution:

$$d(\mathbf{G}_1, \mathbf{G}_2) = \sum_{i=1}^{s} \sum_{j=1}^{s} \log\left(1 + \left(\frac{\mathbf{G}_{1(i,j)} - \mathbf{G}_{2(i,j)}}{2\gamma}\right)^2\right), \tag{6}$$

where $\mathbf{G}_{1(i,j)}$ and $\mathbf{G}_{2(i,j)}$ are the (i, j)th pixels in the patches $\mathbf{G}_1$ and $\mathbf{G}_2$, respectively. Assuming that t patches $\{\mathbf{G}_n\}_{n=1}^{t}$ similar to $\mathbf{G}_{\text{ref}}$ are found in $\mathbf{G}_{\text{noi}}$. Then by concatenating all columns of the patch $\mathbf{G}_n$ into a vector $\mathbf{g}_n \in \mathbb{R}^{s^2}$ and stacking all vectors

Fig. 2. Illustration of grouping via block-matching. The reference patch and matched patches are denoted by red and blur boxes, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in order, we found a noisy matrix $\mathbf{G} \in \mathbb{R}^{s^2 \times t}$ as follows:

$$\mathbf{G} = (\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_t). \tag{7}$$

In Step 2, we denoise each group by solving a GBLR model. We consider the degraded model in the unit of group

$$\mathbf{G} = \mathbf{F} + \mathbf{N}, \tag{8}$$

where $\mathbf{G}$ denotes a group of similar patches from the noisy image $\mathbf{G}_{\text{noi}}$, $\mathbf{F}$ denotes the corresponding clean group from the clean image $\mathbf{F}_{\text{ori}}$, and $\mathbf{N}$ denotes the noise embedded in $\mathbf{G}$.

Next, we provide a mathematical justification for the proposed model from the perspective of maximum a posteriori (MAP) estimator [35,38].

We assume that $\mathbf{N}$ follows the Cauchy distribution with $\delta = 0$ and consider $\mathbf{G}(x)$ and $\mathbf{F}(x)$ as random variables. In order to recover $\mathbf{F}$ from $\mathbf{G}$, we maximize the conditional probability $P(\mathbf{F}|\mathbf{G})$ using the MAP estimator. Based on Bayes' rule, we have

$$\begin{aligned} \arg\max_{\mathbf{F}} P(\mathbf{F}|\mathbf{G}) &\Longleftrightarrow \arg\max_{\mathbf{F}} \frac{P(\mathbf{G}|\mathbf{F})P(\mathbf{F})}{P(\mathbf{G})} \\ &\Longleftrightarrow \arg\min_{\mathbf{F}} -\log P(\mathbf{G}|\mathbf{F}) - \log P(\mathbf{F}) \\ &\Longleftrightarrow \arg\min_{\mathbf{F}} -\sum_x \log P(\mathbf{G}(x)|\mathbf{F}(x)) - \log P(\mathbf{F}), \end{aligned} \tag{9}$$

First, according to the assumptions, we have that

$$P(\mathbf{G}(x)|\mathbf{F}(x)) = \frac{\gamma}{\pi((\mathbf{G}(x) - \mathbf{F}(x))^2 + \gamma^2)}. \tag{10}$$

Second, to exploit the similarity between selected patches of the group, we impose a low-rank prior on $\mathbf{F}$, i.e., $P(\mathbf{F}) \sim \exp(\text{tr}(\sqrt{\mathbf{F}^T\mathbf{F}}))$, where tr denotes the trace and $\sqrt{\mathbf{F}^T\mathbf{F}}$ denotes a positive semidefinite matrix $\mathbf{E}$ such that $\mathbf{EE} = \mathbf{F}^T\mathbf{F}$. The low-rankness of $\mathbf{F}$ can also be seen from Fig. 3.

Then, the proposed model for denoising each group is

$$\arg\min_{\mathbf{F}} \frac{\lambda}{2} \langle \log(\gamma^2 + (\mathbf{G} - \mathbf{F})^2), \mathbf{1} \rangle + \|\mathbf{F}\|_*, \tag{11}$$

where λ is the regularization parameter controlling the balance between the fidelity term and the regularization term, γ is the scale parameter, $\mathbf{1} \in \mathbb{R}^{s^2 \times t}$ is a matrix whose elements equal to 1, and $\langle \cdot, \cdot \rangle$ denotes the standard inner product.
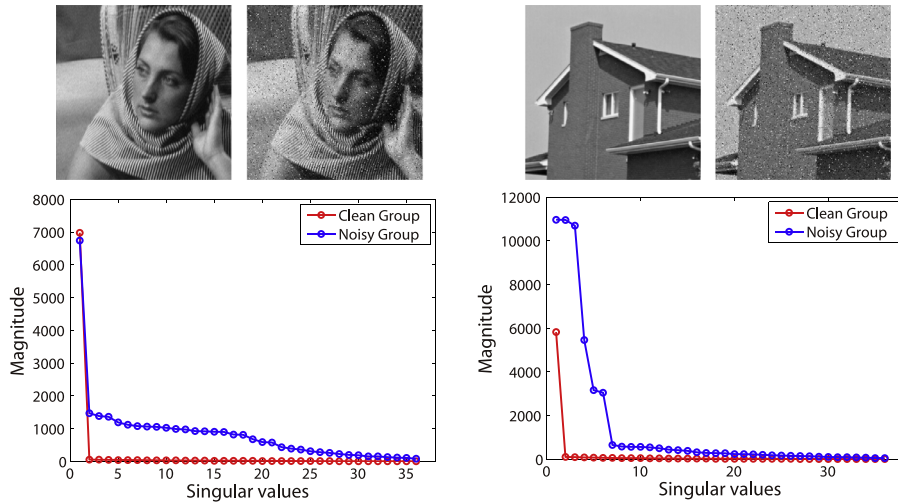
In Step 3, we aggregate the final denoised image. In the proposed method, since the patches are overlapped and one patch can appear in more than one group, each pixel may be covered by several denoised patches. The final value for each pixel is obtained by first returning denoised grouped matrices to their original positions and then taking the average of all the covered patches at this pixel.

3. Numerical algorithm

This section is divided into three parts. First, we review the framework of ADMM for nonconvex and nonsmooth problems. Second, we present the proposed algorithm in details. Third, we demonstrate the convergence of the proposed algorithm.

3.1. ADMM for nonconvex and nonsmooth problems

Before applying ADMM to solve the nonconvex model (11), we review its framework and convergence in nonconvex nonsmooth optimization [40].

Fig. 3. Group-based low-rank property of images Barbara (left) and Wall (right). Top row: original images and their degraded images. Bottom row: plots of singular values of the clean and noisy group extracted from the original and the degraded images. Each group is of size 36×60 containing 60 patches of size 6×6. One can see that the singular values of the clean group are much smaller and sparser than those of the noisy group.

Wang et al. [40] considers the following optimization problem with a linear constraint:

$$\min_{x,y} \quad \mathcal{F}(x) + \mathcal{H}(y)$$
$$s.t. \quad \mathbf{P}x + \mathbf{Q}y = 0, \tag{12}$$

where $\mathcal{F}(x)$ is a continuous, proper, possibly nonsmooth and nonconvex function, $\mathcal{H}(y)$ is a proper, differentiable and possibly nonconvex function, and $x \in \mathbb{R}^{n_1}$, $y \in \mathbb{R}^{n_2}$ are variables with the corresponding coefficients $\mathbf{P} \in \mathbb{R}^{l \times n_1}$, $\mathbf{Q} \in \mathbb{R}^{l \times n_2}$, respectively. By introducing a Lagrangian multiplier $z \in \mathbb{R}^l$ for the linear constraint, we have the following augmented Lagrangian function:

$$\mathcal{L}_\beta(x, y, z) = \mathcal{F}(x) + \mathcal{H}(y) + \langle z, \mathbf{P}x + \mathbf{Q}y \rangle + \frac{\beta}{2} \|\mathbf{P}x + \mathbf{Q}y\|_2^2,$$

where $\beta > 0$ is a penalty parameter.

ADMM [16,44] solves (12) in the following iterative way:

$$\begin{cases} x^{d+1} = \arg\min_x \mathcal{L}(x, y^d, z^d), \\ y^{d+1} = \arg\min_y \mathcal{L}(x^{d+1}, y, z^d), \\ z^{d+1} = z^d + \beta(\mathbf{P}x^{d+1} + \mathbf{Q}y^{d+1}). \end{cases} \tag{13}$$

The following lemma [40] establishes the convergence result of ADMM under the nonconvex and nonsmooth case.

Lemma 1. *Suppose that the following assumptions A1–A5 hold, for any sufficiently large β, then the sequence generated by (13) starting from any initialization has at least one limit point, and each limit point is a stationary point of the augmented Lagrangian function.*

A1 (**coercivity**) *Let $\mathcal{D} = \{(x, y) \in \mathbb{R}^{n_1 + n_2} : \mathbf{P}x + \mathbf{Q}y = 0\}$ be the nonempty feasible set and $\mathcal{F}(x) + \mathcal{H}(y)$ is coercive over $\mathcal{D}$;*

A2 (**feasibility**) *$Im(\mathbf{P}) \subseteq Im(\mathbf{Q})$, where $Im(\cdot)$ denotes the image of a matrix;*

A3 (**Lipschitz sub-minimization paths**)

(a) For any x, there exists a Lipschitz continuous map $H : Im(\mathbf{Q}) \to \mathbb{R}^l$ obeying $H(u) = \arg\min_y \{\mathcal{F}(x) + \mathcal{H}(y) : \mathbf{Q}y = u\}$,

(b) For any y, there exists a Lipschitz continuous map $H : Im(\mathbf{P}) \to \mathbb{R}^l$ obeying $F(u) = \arg\min_x \{\mathcal{F}(x) + \mathcal{H}(y) : \mathbf{P}x = u\}$;

A4 (**objective-$\mathcal{F}$regularity**) *$\mathcal{F}$ is lower semi-continuous or $\sup\{\|f\| : x \in S, f \in \partial\mathcal{F}(x)\}$ is bound for any bound set S;*

A5 (**objective-$\mathcal{H}$regularity**) *$\mathcal{H}$ is Lipschitz differentiable with the constraint $L_{\nabla\mathcal{H}} > 0$.*

3.2. Proposed algorithm

We adopt ADMM to minimize the proposed objective function. By introducing an auxiliary variable $\mathbf{A}$, we get the equivalent constrained problem

$$\arg\min_{\mathbf{F},\mathbf{A}} \quad \frac{\lambda}{2} \langle \log(\gamma^2 + (\mathbf{G} - \mathbf{A})^2), \mathbf{1} \rangle + \|\mathbf{F}\|_*$$
$$s.t. \quad \mathbf{F} = \mathbf{A}. \tag{14}$$

The augmented Lagrangian function of (14) is defined as

$$\mathcal{L}_\beta(\mathbf{F}, \mathbf{A}, \mathbf{B}) = \frac{\lambda}{2} \langle \log(\gamma^2 + (\mathbf{G} - \mathbf{A})^2), \mathbf{1} \rangle + \|\mathbf{F}\|_* + \langle \mathbf{F} - \mathbf{A}, \mathbf{B} \rangle + \frac{\beta}{2} \|\mathbf{F} - \mathbf{A}\|_F^2, \tag{15}$$

where $\mathbf{B}$ is the Lagrangian multiplier of the linear constraint and $\beta > 0$ is the penalty parameter. According to (13), the iterative scheme for solving (14) is as follows:

$$\begin{cases} \mathbf{F}^{d+1} = \arg\min_\mathbf{F} \mathcal{L}_\beta(\mathbf{F}, \mathbf{A}^d, \mathbf{B}^d), \\ \mathbf{A}^{d+1} = \arg\min_\mathbf{A} \mathcal{L}_\beta(\mathbf{F}^{d+1}, \mathbf{A}, \mathbf{B}^d), \\ \mathbf{B}^{d+1} = \mathbf{B}^d + \beta(\mathbf{F}^{d+1} - \mathbf{A}^{d+1}). \end{cases} \tag{16}$$

We give the details for solving the first two subproblems in (16).

1. $\mathbf{F}$-subproblem

$$\mathbf{F}^{d+1} = \arg\min_\mathbf{F} \frac{\beta}{2} \left\| \mathbf{F} - \mathbf{A}^d + \frac{\mathbf{B}^d}{\beta} \right\|_F^2 + \|\mathbf{F}\|_*. \tag{17}$$

Using SVT, $\mathbf{F}^{d+1}$ has the following closed-form solution:

$$\mathbf{F}^{d+1} = \mathbf{U}^d \mathbf{S}_{\frac{1}{\beta}} (\mathbf{\Sigma}^d)(\mathbf{V}^d)^T, \tag{18}$$

where $\mathbf{A}^d - \frac{\mathbf{B}^d}{\beta} = \mathbf{U}^d \mathbf{\Sigma}^d (\mathbf{V}^d)^T$ and $\mathbf{S}_{\frac{1}{\beta}}(\mathbf{\Sigma}^d) = diag(\max(\sigma_r^d - \frac{1}{\beta}, 0))$.

2. $\mathbf{A}$-subproblem

$$\mathbf{A}^{d+1} = \arg\min_\mathbf{A} \frac{\lambda}{2} \langle \log(\gamma^2 + (\mathbf{G} - \mathbf{A})^2), \mathbf{1} \rangle + \frac{\beta}{2} \left\| \mathbf{F}^{d+1} - \mathbf{A} + \frac{\mathbf{B}^d}{\beta} \right\|_F^2. \tag{19}$$

Given fixed $\mathbf{G}$, $\mathbf{F}^{d+1}$, and $\mathbf{B}^d$, the optimal $\mathbf{A}$ consists of solving $s^2 t$ one-variable minimization problems

$$\arg\min_{\mathbf{A}_{m,n}} \frac{\lambda}{2} \log\left(\gamma^2 + (\mathbf{G}_{m,n} - \mathbf{A}_{m,n})^2\right) + \frac{\beta}{2} \left(\mathbf{F}_{m,n}^{d+1} - \mathbf{A}_{m,n} + \frac{\mathbf{B}_{m,n}^d}{\beta} \right)^2, \tag{20}$$

where $\mathbf{A}_{m,n}$, $\mathbf{G}_{m,n}$, $\mathbf{F}_{m,n}^{d+1}$, and $\mathbf{B}_{m,n}^d$ are the (m, n)th pixels in $\mathbf{A}, \mathbf{G}, \mathbf{F}^{d+1}$, and $\mathbf{B}^d$, respectively, and $m = 1, \ldots, s^2, n = 1, \ldots, t$.

Let

$$\mathbf{W}(\mathbf{A}_{m,n}) = \frac{\lambda}{2} \log\left(\gamma^2 + (\mathbf{G}_{m,n} - \mathbf{A}_{m,n})^2\right) + \frac{\beta}{2} \left(\mathbf{F}_{m,n}^{d+1} - \mathbf{A}_{m,n} + \frac{\mathbf{B}_{m,n}^d}{\beta} \right)^2, \tag{21}$$

then the optimality condition for $\mathbf{A}$-subproblem is given by

$$\mathbf{W}'(\mathbf{A}_{m,n}) = 0, \tag{22}$$

where $\mathbf{W}'$ is the first order derivative of $\mathbf{W}$. As there exists the second order derivative for (21), we get the solution of (22) by the Newton method

$$\mathbf{A}_{m,n}^{d+1,h+1} = \mathbf{A}_{m,n}^{d+1,h} - \frac{\mathbf{W}'(\mathbf{A}_{m,n}^{d+1,h})}{\mathbf{W}''(\mathbf{A}_{m,n}^{d+1,h})}, \tag{23}$$

where $\mathbf{W}''$ is the second order derivative of $\mathbf{W}$. $\mathbf{A}_{m,n}^{d+1,h+1}$ is the result of the $(h + 1)$th iteration in the $(d + 1)$th outer iteration. In the proposed method, we initialize the Newton algorithm using the current $\mathbf{G}_{m,n}$. However, we have not yet succeed in establishing a rigorous convergence of the Newton method, since it is difficult to check the technical conditions for the initialization point, which is changing during iteration. We empirically find that the Newton method rapidly approaches the minimizer to a high accuracy in few iterations.

The whole denoised algorithm is summarized in Algorithm 2.

3.3. Convergence

In the following, we study the convergence of the proposed algorithm. Inspired by [40], as a special case of (12), we have the following convergence result for Algorithm 2.

Theorem 1. *For sufficiently large β, the sequence* $(\mathbf{F}^d, \mathbf{A}^d, \mathbf{B}^d)$ *generated by Algorithm 2 has limit points and all of its limit points are stationary points of the augmented Lagrangian $\mathcal{L}_\beta$.*

Algorithm 2 Cauchy Noise Removal using Group-based Low-rankness Image Prior.

Input: noisy image $\mathbf{G}_{\text{noi}}$, parameters λ, β, s, t, and s_{win}, maximum iteration d_{max}, and relative error ε.

Initialization: Perform grouping on $\mathbf{G}_{\text{noi}}$ to get each group $\mathbf{G}$.

Out loop: For each group, do

    Initialization: set $\mathbf{A}^0 = \mathbf{G}$, $\mathbf{B}^0 = 0$.

    Inner loop: while $d \leq d_{\text{max}}$ **or** $\frac{\|\mathbf{F}^{d+1} - \mathbf{F}^d\|_F^2}{\|\mathbf{F}^d\|_F^2} \geq \varepsilon$, **do**

    1: solve $\mathbf{F}^{d+1}$ via (18).

    2: solve $\mathbf{A}^{d+1}$ via Newton method. The details are omitted.

    3: update $\mathbf{B}^{d+1}$ by (16).

    End Inner loop, and output $\mathbf{F}^{d+1}$.

End Out loop, and output denoised groups.

Output: aggregate the restored image $\tilde{\mathbf{F}}$ via denoised groups.

Proof. We rewrite the proposed model (11) as

$$\arg\min_{\mathbf{F},\mathbf{A}} \quad \mathcal{F}(\mathbf{F}) + \mathcal{H}(\mathbf{A})$$
$$s.t. \quad \mathbf{F} - \mathbf{A} = 0,$$

where $\mathcal{F}(\mathbf{F}) = \|\mathbf{F}\|_*$, $\mathcal{H}(\mathbf{A}) = \frac{\lambda}{2}\langle \log(\gamma^2 + (\mathbf{G} - \mathbf{A})^2), \mathbf{1}\rangle$. It is clear that our model fits the framework of (12).

Now we check A1–A5. A1 holds because of the coercivity of $\mathcal{F}(\mathbf{F}) + \mathcal{H}(\mathbf{A})$. A2 and A3 hold because both P and Q are identity matrices. A4 holds because nuclear norm is lower semi-continuous. Hence, it remains to verify A5 that $\mathcal{H}(\mathbf{A}) = \frac{\lambda}{2}\langle \log(\gamma^2 + (\mathbf{G} - \mathbf{A})^2), \mathbf{1}\rangle$ is Lipschitz differentiable.

Since $\mathcal{H}$ is smooth, we can calculate its second derivative

$$\partial_{\mathbf{A}}^2 \mathcal{H} = \lambda \frac{\gamma^2 - (\mathbf{G} - \mathbf{A})^2}{(\gamma^2 + (\mathbf{G} - \mathbf{A})^2)^2} \leq \frac{\lambda}{\gamma^2},$$

and $L_{\nabla \mathcal{H}} = \frac{\lambda}{\gamma^2} > 0$ is Lipschitz constant for $\nabla \mathcal{H}$. This verifies A5 and completes the proof. $\square$

4. Numerical experiments

In this section, several numerical experiments are presented to demonstrate the performance of the proposed method in Cauchy noise removal. All numerical experiments are performed on Windows 10 64-bit and MATLAB R2012a running on a desktop equipped with an Intel(R) Core(TM) i7-6700M CPU with 3.40 GHz and 8 GB of RAM.

Fig. 4 shows six 256-by-256 gray-scale testing images with an intensity range [0, 255]. We simulate a Cauchy noise by the ratio of two independent standard normal variables, and generate the noisy image $\mathbf{G}_{\text{noi}}$ by using the following degradation:

$$\mathbf{G}_{\text{noi}} = \mathbf{F}_{\text{ori}} + \gamma \frac{\eta_1}{\eta_2},$$

where $\gamma > 0$ is the noise level, η_1 and η_2 are two independent random variables following Gaussian distribution with zero mean and unit variance.

We compare the proposed method ("GBLR" for short) with four compared methods, including the convex variational method [38] ("CTV" for short), the non-convex variational method [35] ("NTV" for short), the nonlocal generalized myriad filter-based method [25] ("NGMF" for short), and the generalized multivariate myriad filter-based method [26] ("GMMF" for short).

(a) Barbara   (b) Cameraman   (c) Wall   (d) Parrot   (e) Bamboo   (f) Computer

Fig. 4. Original images.

Table 1
The PSNR and SSIM values obtained by different methods for testing images with $\gamma = 5$.

Image	PSNR						SSIM					
	Noisy	CTV	NTV	NGMF	GMMF	GBLR	Noisy	CTV	NTV	NGMF	GMMF	GBLR
Barbara	19.23	23.51	26.94	29.53	29.29	**31.20**	0.4797	0.7463	0.8404	0.8764	0.8751	**0.8964**
Cameraman	19.23	26.73	28.80	28.78	28.14	**29.77**	0.3604	0.8246	0.8431	0.8224	0.8136	**0.8647**
Wall	19.24	32.10	32.51	32.30	32.15	**33.48**	0.2954	0.8739	0.8710	0.8372	0.8361	**0.9001**
Parrot	19.20	30.36	30.93	31.96	31.43	**32.69**	0.3122	0.8986	0.8942	0.8565	0.8559	**0.9034**
Bamboo	19.17	23.88	25.78	26.90	27.12	**28.15**	0.7942	0.9148	0.9414	0.9479	0.9518	**0.9603**
Computer	19.22	34.03	33.75	33.84	33.90	**36.06**	0.2765	0.9291	0.9228	0.8679	0.8724	**0.9482**
Average	19.22	28.44	29.79	30.55	30.34	**31.89**	0.4197	0.8646	0.8855	0.8680	0.8675	**0.9122**

| (a) Noisy | (b) CTV | (c) NTV | (d) NGMF | (e) GMMF | (f) GBLR |

Fig. 5. Restored images by different methods with $\gamma = 5$. (a) noisy image, the results of (b) CTV, (c) NTV, (d) NGMF, (e) GMMF, and (f) GBLR.

Table 2
The PSNR and SSIM values obtained by different methods for testing images with $\gamma = 10$.

Image	PSNR						SSIM					
	Noisy	CTV	NTV	NGMF	GMMF	GBLR	Noisy	CTV	NTV	NGMF	GMMF	GBLR
Barbara	16.31	23.05	24.56	26.58	26.19	**28.29**	0.3422	0.6942	0.7525	0.7999	0.7901	**0.8414**
Cameraman	16.33	25.97	26.88	26.60	25.14	**27.49**	0.2470	0.7714	0.7860	0.6550	0.7368	**0.8034**
Wall	16.35	30.40	30.35	29.96	29.20	**31.16**	0.1889	0.8430	0.8173	0.7989	0.7770	**0.8518**
Parrot	16.31	28.87	28.99	28.88	28.00	**29.82**	0.2008	0.8510	0.8557	0.8111	0.7943	**0.8716**
Bamboo	16.28	22.61	23.20	24.86	23.04	**25.18**	0.6820	0.8843	0.8976	0.9163	0.8874	**0.9264**
Computer	16.31	31.82	31.56	31.02	30.45	**32.89**	0.1684	0.8950	0.8752	0.8278	0.8100	**0.9054**
Average	16.32	27.12	27.59	27.98	27.00	**29.14**	0.3049	0.8232	0.8307	0.8015	0.7993	**0.8667**

(a) Noisy   (b) CTV   (c) NTV   (d) NGMF   (e) GMMF   (f) GBLR

Fig. 6. Restored images by different methods with $\gamma = 10$. (a) noisy image, the results of (b) CTV, (c) NTV, (d) NGMF, (e) GMMF, and (f) GBLR.
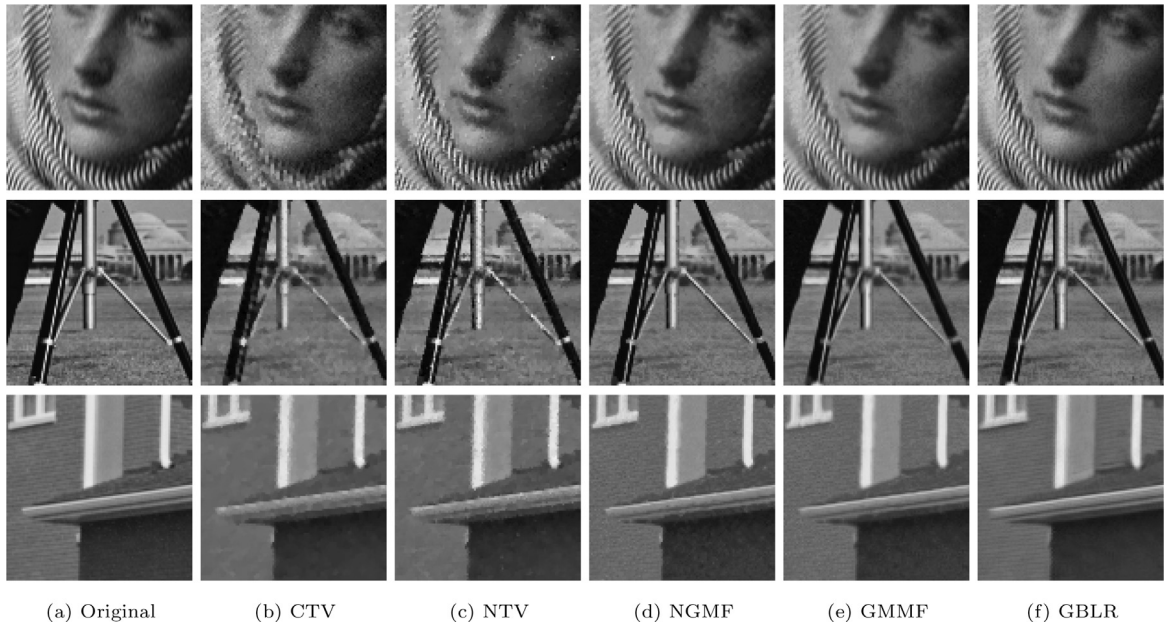
(a) Original   (b) CTV   (c) NTV   (d) NGMF   (e) GMMF   (f) GBLR

Fig. 7. Zoomed vision of restored images in Fig. 5. (a) original image, the results of (b) CTV, (c) NTV, (d) NGMF, (e) GMMF, and (f) GBLR.

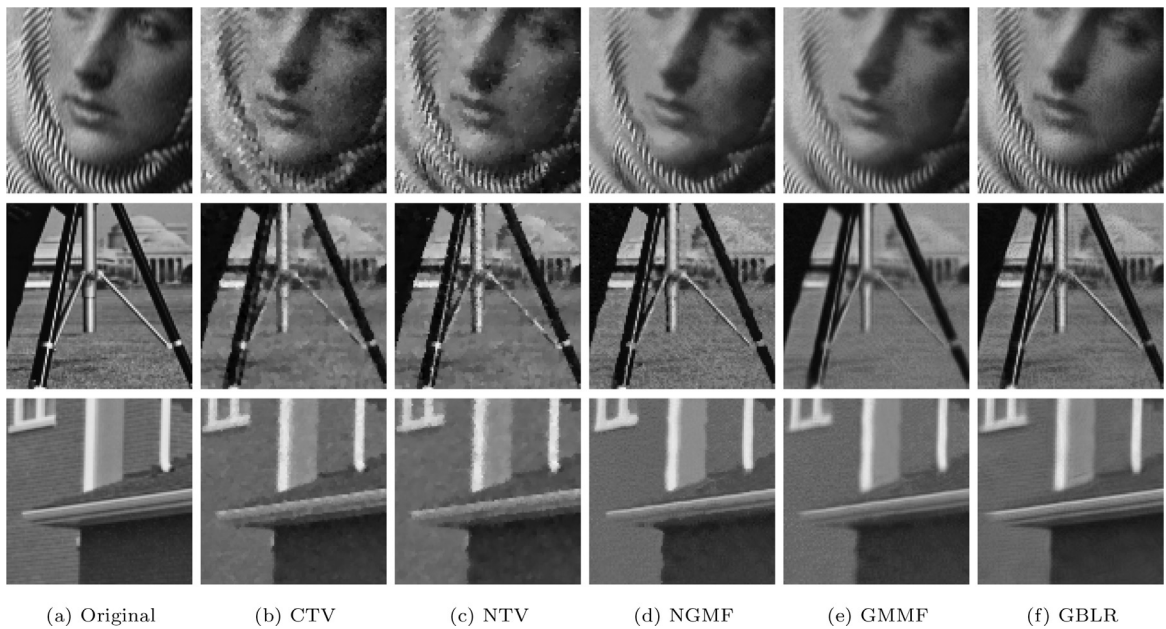(a) Original   (b) CTV   (c) NTV   (d) NGMF   (e) GMMF   (f) GBLR

Fig. 8. Zoomed vision of restored images in Fig. 6. (a) original image, the results of (b) CTV, (c) NTV, (d) NGMF, (e) GMMF, and (f) GBLR.

Quantitative indices. For quantitative comparison, the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [42] are used to measure the quality of the restored image. They are defined as follows:

$$\mathrm{PSNR} = 10 \log_{10}\left(\frac{255^2}{\frac{1}{m_1 m_2}\|\mathbf{F}_{\mathrm{ori}} - \tilde{\mathbf{F}}\|_F^2} \right),$$

$$\mathrm{SSIM} = \frac{2\mu_{\mathbf{F}_{\mathrm{ori}}}\mu_{\tilde{\mathbf{F}}}(2\sigma + c_2)}{(\mu_{\mathbf{F}_{\mathrm{ori}}}^2 + \mu_{\tilde{\mathbf{F}}}^2 + c_1)(\sigma_{\mathbf{F}_{\mathrm{ori}}}^2 + \sigma_{\tilde{\mathbf{F}}}^2 + c_2)},$$
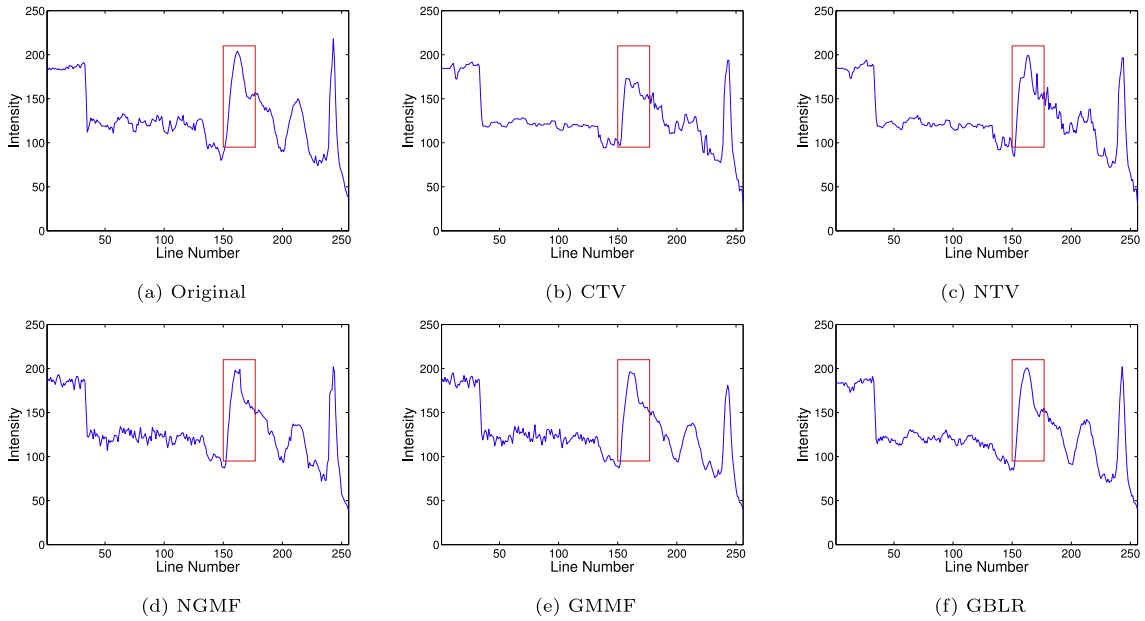
Fig. 9. Illustration of the ability of the proposed method to keep the smoothness. The intensities of a row of the recovered Wall images with $\gamma = 5$.

where $\mathbf{F}_{\text{ori}}$ and $\tilde{\mathbf{F}}$ are the original image and the restored image, respectively, $\mu_{\mathbf{F}_{\text{ori}}}$ and $\mu_{\tilde{\mathbf{F}}}$ are the mean values of $\mathbf{F}_{\text{ori}}$ and $\tilde{\mathbf{F}}$, $\sigma^2_{\mathbf{F}_{\text{ori}}}$ and $\sigma^2_{\tilde{\mathbf{F}}}$ denote their standard variances, σ is the covariance of $\mathbf{F}_{\text{ori}}$ and $\tilde{\mathbf{F}}$, and $c_1, c_2 > 0$ are constants. PSNR value and SSIM value measure the distance and the structural similarity between $\mathbf{F}_{\text{ori}}$ and $\tilde{\mathbf{F}}$, respectively. Higher values indicate better image quality.

Parameters setting. The parameter settings for our method are as follows. We set the iteration number of the Newton method to 10, the outer loop to $d_{\max} = 500$, the search window in block-matching to $s_{\text{win}} = 30$. For noise levels $\gamma = 5$ and $\gamma = 10$, we set the patch size to 6×6 and 8×8, the similar patch number to 60 and 80, respectively. We empirically set $\lambda \in [1.5, 3]$ with increment of 0.1 and $\beta \in [0.1, 1]$ with increment of 0.1 to get the highest PSNR value. For all denoised groups, we stop our algorithm according to the relative error between two successive iterations, i.e.,

$$\frac{\|\mathbf{F}^{d+1} - \mathbf{F}^d\|^2_F}{\|\mathbf{F}^d\|^2_F} < 10^{-5}.$$

As compared methods, there are some parameters which play an important role in restored results, such as the quadratic term parameter relied on convexity condition in CTV, the penalty parameter which controls the convergent speed in NTV, thus we tune these parameters according to the satisfied condition. For a fair comparison, we try our best to tune the parameters involved in compared methods according to the authors' suggestions to achieve the highest PSNR values.

Comparison results. Figs. 5 and 6 show the denoising results with noise levels $\gamma = 5$ and $\gamma = 10$. We can observe that the results recovered by both CTV and NTV have staircase effects in smooth regions and undesired pixels in the boundary. Especially, the results of NTV still contain some corrupted pixels when $\gamma = 5$, which can be seen in the face part of the Barbara image. Although NGMF and GMMF can effectively reduce noise, for a large noise level $\gamma = 10$, NGMF destroys the edges such as the eaves of the Wall image and GMMF oversmoothes the structures and loses many image details such as the scarf of the Barbara image and the face area of the Cameraman image. As a comparison, our method can effectively remove Cauchy noise and outperform both compared methods in terms of edge and detail preservation by well maintaining the most intact structures and boundary.

Figs. 7 and 8 show the zoom-in regions of restored images. From these restored images, we can see that the proposed method obtains smoother image with clearer texture and line pattern. Fig. 9 displays the intensities of a row of the recovered Wall images with $\gamma = 5$. It is clear that the intensity profile of the proposed method has almost the same curve as the original one and keeps the smoothness best.

Tables 1 and 2 present PSNR and SSIM values of restored images by all methods. The highest results for each quality index are labeled in bold. It is clear that for different noise levels, the proposed method achieves the highest PSNR and SSIM values.
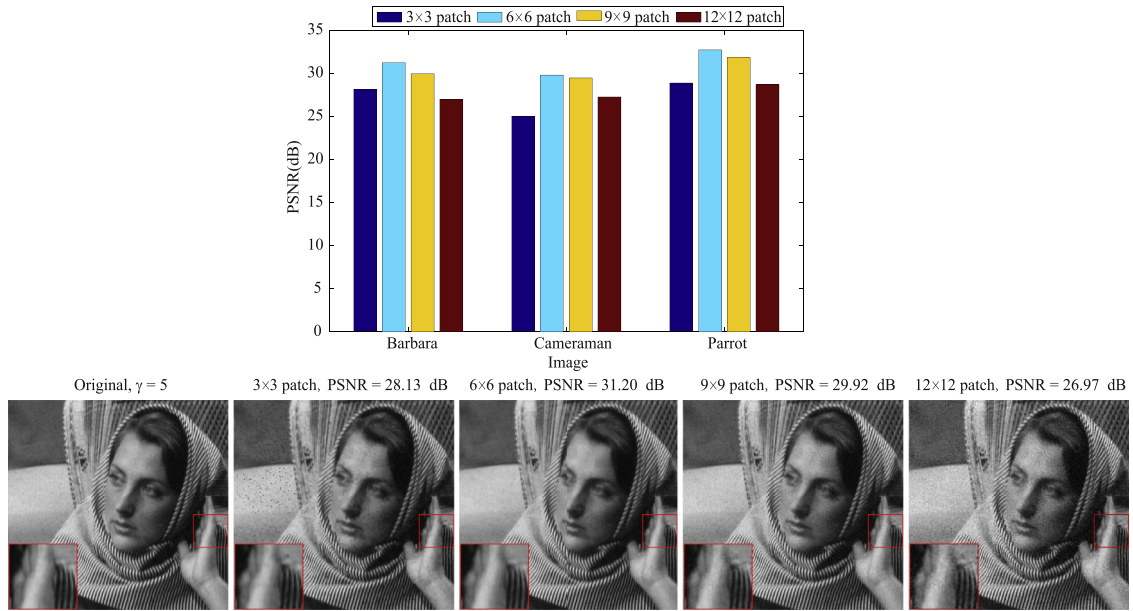
Original, γ = 5   3×3 patch, PSNR = 28.13 dB   6×6 patch, PSNR = 31.20 dB   9×9 patch, PSNR = 29.92 dB   12×12 patch, PSNR = 26.97 dB

Fig. 10. Effects of the patch size. Top: PSNR values of the proposed method on three images with $\gamma = 5$ using different patch sizes. Bottom: Restored images of Barbara with different patch sizes.
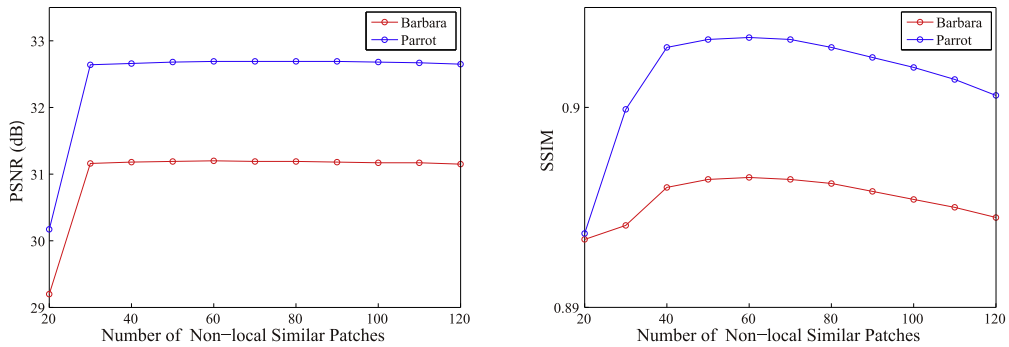

Fig. 11. Effects of the number of the non-local similar patches on denoising results.

5. Discussions

5.1. Effect of the patch size

We discuss the effect of the patch size s on the denoising performance. Fig. 10 plots the PSNR values of the proposed method on three images with $s = 3, 6, 9, 12$ and the noise level $\gamma = 5$. And we also provide a denoising example on Barbara. Regarding PSNR, we observe that the highest PSNR value is obtained when $s = 6$. Regarding visual quality, it is clear that the result of $s = 6$ exhibits the best image textures and details. Therefore, we empirically fix $s = 6$ on the noise level $\gamma = 5$. Similarly, we set the patch size $s = 8$ for $\gamma = 10$.

5.2. Number of non-local similar patches

We investigate the effects of the number of non-local similar patches t on the denoising performance. Fig. 11 shows the PSNR and SSIM values as a function of t when the noise level $\gamma = 5$. We observe that the PSNR value is stable with respect to t and the highest SSIM value is achieved around $t = 60$. Therefore, in the case of noise level $\gamma = 5$, we set $t = 60$ as the number of non-local similar patches. Similarly, we set $t = 80$ when $\gamma = 10$.

5.3. Numerical convergence

In this subsection, we empirically demonstrate the convergence of the proposed algorithm. In Section 3, Theorem 1 demonstrates that with a given initial value, Algorithm 2 converges to a unique stationary point. Fig. 12 plots of the relative
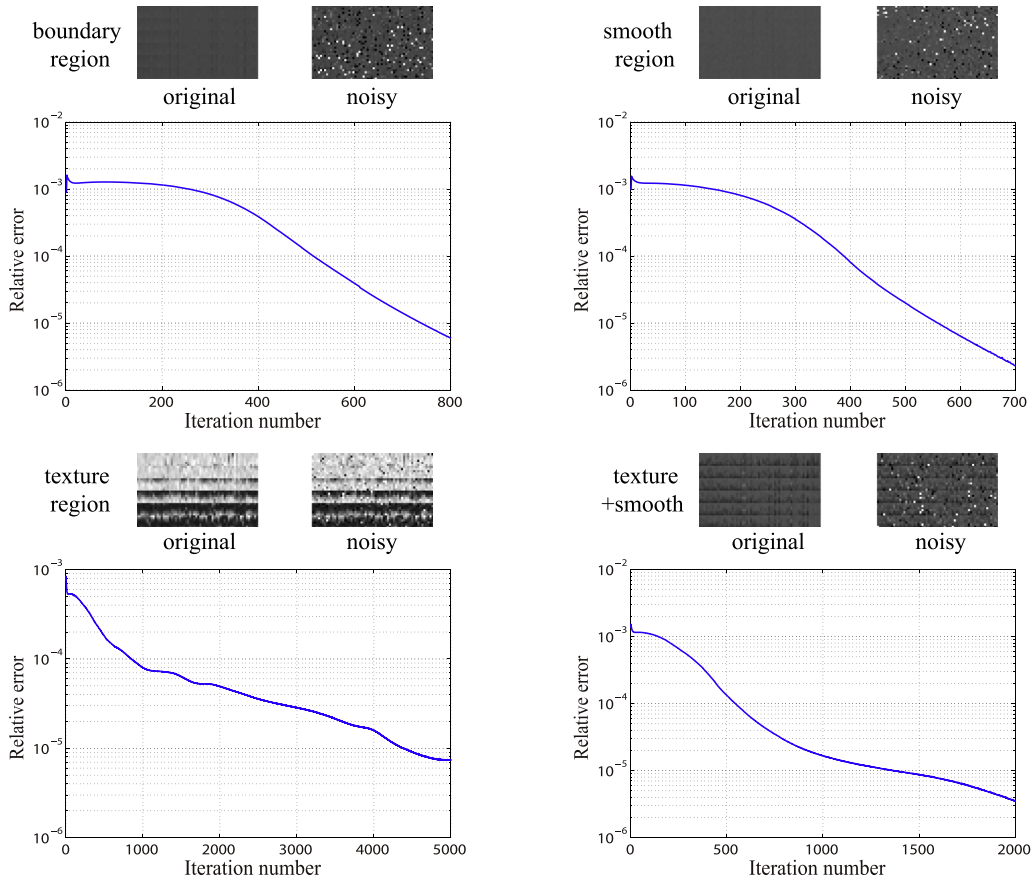
Fig. 12. Plots of the relative error values of restored groups versus iteration numbers, based on image Parrot with $\gamma = 5$ in four kinds of regions.

error values of restored groups versus iteration numbers, based on image Parrot with $\gamma = 5$ in four kinds of regions. We observe that as the iteration number increases, the relative errors between two successive restored groups approach to zero in all cases.

6. Conclusion

We propose a new method based on group-based low-rank image prior for Cauchy noise removal. We stack the similar patches from the noisy image as a matrix, and then formulate the Cauchy noise removal problem as the low-rank matrix recovery problem. We have developed an efficient numerical scheme based on ADMM for solving the proposed model with guaranteed convergence. Experiments demonstrate the effectiveness of the proposed method. In the future, we will extend the proposed method to other image processing applications [21,46,51].

Acknowledgments

References

[1] B. Amir, T. Marc, Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems, IEEE Trans. Image Process. 18 (11) (2009) 2419.

[2] S. Banerjee, M. Agrawal, Underwater acoustic communication in the presence of heavy-tailed impulsive noise with bi-parameter Cauchy-Gaussian mixture model, in: IEEE Conference on Ocean Electronics (SYMPOL), 2013, pp. 1–7.

[3] K. Bredies, K. Kunisch, T.T. Pock, Total generalized variation, SIAM J. Imaging Sci. 3 (3) (2013) 492–526.

[4] A. Buades, B. Coll, J.M. Morel, A non-local algorithm for image denoising, in: IEEE Conference on Computer Vision and Pattern Recognition, 2005, pp. 60–65.

[5] J.-F. Cai, E.J. Candès, Z. Shen, A singular value thresholding algorithm for matrix completion, SIAM J. Optim. 20 (4) (2008) 1956–1982.

[6] R.H. Chan, Y.-Q. Dong, M. Hintermuller, An efficient two-phase L1-TV method for restoring blurred images with impulse noise, IEEE Trans. Image Process. 19 (7) (2010) 1731–1739.

[7] T.F. Chan, S. Esedoglu, F.E. Park, Image decomposition combining staircase reduction and texture extraction, J. Vis. Commun. Image Represent. 18 (6) (2007) 464–486.

[8] Y.-Y. Chen, Y.-W. Guo, Y.-L. Wang, D. Wang, C. Peng, G.-P. He, Denoising of hyperspectral images using nonconvex low rank matrix approximation, IEEE Trans. Geosci. Remote Sens. 55 (9) (2017) 5366–5380.

[9] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Trans. Image Process. 16 (8) (2007) 2080–2095.

[10] A. Danielyan, V. Katkovnik, K. Egiazarian, BM3D frames and variational image deblurring, IEEE Trans. Image Process. 21 (4) (2012) 1715–1728.

[11] M. Ding, T.-Z. Huang, T.-Y. Ji, X.-L. Zhao, J.-H. Yang, Low-rank tensor completion using matrix factorization based on tensor train rank and total variation, J. Sci. Comput. 81 (2019) 941–964.

[12] M. Ding, T.-Z. Huang, S. Wang, J.-J. Mei, X.-L. Zhao, Total variation with overlapping group sparsity for deblurring images under Cauchy noise, Appl. Math. Comput. 341 (2019) 128–147.

[13] W.-S. Dong, G.-M. Shi, X. Li, Nonlocal image restoration with bilateral variance estimation: a low-rank approach, IEEE Trans. Image Process. 22 (2) (2013) 700–711.

[14] Y.-Q. Dong, T.-Y. Zeng, A convex variational model for restoring blurred images with multiplicative noise, SIAM J. Imaging Sci. 6 (3) (2013) 1598–1625.

[15] A. Elmoataz, O. Lezoray, S. Bougleux, Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing, IEEE Trans. Image Process. 17 (7) (2008) 1047–1060.

[16] T. Goldstein, B. O'Donoghue, S. Setzer, Fast alternating direction optimization methods, SIAM J. Imaging Sci. 7 (3) (2014) 225–231.

[17] S.-H. Gu, L. Zhang, W.-M. Zuo, X.-C. Feng, Weighted nuclear norm minimization with application to image denoising, in: IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 2862–2869.

[18] C. Guillemot, O.L. Meur, Image inpainting: overview and recent advances, IEEE Signal Process. Mag. 31 (1) (2013) 127–144.

[19] Y. Hu, M. Jacob, Higher degree total variation (HDTV) regularization for image recovery., IEEE Trans. Image Process. 21 (5) (2012) 2559–2571.

[20] H. Ji, S.-B. Huang, Z.-W. Shen, Y.-H. Xu, Robust video restoration by joint sparse and low rank matrix approximation, SIAM J. Imaging Sci. 4 (4) (2011) 1122–1142.

[21] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, L.-J. Deng, Y. Wang, FastDeRain: a novel video rain streak removal method using directional gradient priors, IEEE Trans. Image Process. 28 (4) (2019) 2089–2102.

[22] S. Kindermann, S. Osher, P.W. Jones, Deblurring and denoising of images by nonlocal functionals, SIAM J. Multiscale Model. Simul. 4 (4) (2005) 1091–1115.

[23] D. Krishnan, R. Fergus, Fast image deconvolution using hyper-Laplacian priors, in: Advances in Neural Information Processing Systems, 2009, pp. 1033–1041.

[24] E.E. Kuruoglu, W.J. Fitzgerald, P.J.W. Rayner, Near optimal detection of signals in impulsive noise modeled with a symmetric α-stable distribution, IEEE Commun. Lett. 2 (10) (1998) 282–284.

[25] F. Laus, F. Pierre, G. Steidl, Nonlocal myriad filters for Cauchy noise removal, J. Math. Imaging Vis. 60 (8) (2018) 1324–1354.

[26] F. Laus, G. Steidl, Multivariate myriad filters based on parameter estimation of student-t distributions, arXiv:1810.05594 (2019).

[27] F. Li, T.-Y. Zeng, A universal variational framework for sparsity-based image inpainting, IEEE Trans. Image Process. 23 (10) (2014) 4242–4254.

[28] X.-T. Li, X.-L. Zhao, T.-X. Jiang, Y.-B. Zheng, T.-Y. Ji, T.-Z. Huang, Low-rank tensor completion via combined non-local similarity and low-rank regularization, Neurocomputing 70 (2019) 677–695.

[29] G.-C. Liu, Z.-C. Lin, Y. Yu, Robust subspace segmentation by low-rank representation, in: International Conference on Machine Learning, 2010, pp. 663–670.

[30] R.-S. Liu, Z.-C. Lin, F. De la Torre, Z.-X. Su, Fixed-rank representation for unsupervised visual learning, in: IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 598–605.

[31] Y.-F. Lou, X.-J. Zhang, S. Osher, A. Bertozzi, Image recovery via nonlocal operators, J. Sci. Comput. 42 (2) (2010) 185–197.

[32] T.-H. Ma, T.-Z. Huang, X.-L. Zhao, Group-based image decomposition using 3-D cartoon and texture priors, Inf. Sci. 328 (2016) 510–527.

[33] T.-H. Ma, T.-Z. Huang, X.-L. Zhao, Y. Lou, Image deblurring with an inaccurate blur kernel using a group-based low-rank image prior, Inf. Sci. 408 (2017) 213–233.

[34] J. Mairal, F. Bach, J. Ponce, G. Sapiro, A. Zisserman, Non-local sparse models for image restoration, in: IEEE International Conference on Computer Vision, 2009, pp. 2272–2279.

[35] J.-J. Mei, Y.-Q. Dong, T.-Z. Huang, W.-T. Yin, Cauchy noise removal by nonconvex admm with convergence guarantees, J. Sci. Comput. 74 (2) (2018) 743–766.

[36] Y.-G. Peng, A. Ganesh, J. Wright, W.-L. Xu, Y. Ma, RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images, IEEE Trans. Pattern Anal. Mach.Intell. 34 (11) (2012) 2233–2246.

[37] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1–4) (1992) 259–268.

[38] F. Sciacchitano, Y.-Q. Dong, T.-Y. Zeng, Variational approach for restoring blurred images with Cauchy noise, SIAM J. Imaging Sci. 8 (3) (2015) 1894–1922.

[39] M. Shinde, S. Gupta, Signal detection in the presence of atmospheric noise in tropics, IEEE Trans. Commun. 22 (8) (1974) 1055–1063.

[40] Y. Wang, W.-T. Yin, J.-S. Zeng, Global convergence of ADMM in nonconvex nonsmooth optimization, J. Sci. Comput. 78 (1) (2019) 29–63.

[41] Y.-L. Wang, J.-F. Yang, W.-T. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, SIAM J. Imaging Sci. 1 (3) (2008) 248–272.

[42] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.

[43] P. Weiss, L. Blanc-Féraud, G. Aubert, Efficient schemes for total variation minimization under constraints in image processing, SIAM J. Sci. Comput. 31 (3) (2009) 2047–2080.

[44] C.-L. Wu, X.-C. Tai, Augmented lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models, SIAM J. Imaging Sci. 3 (3) (2010) 300–339.

[45] J. Yang, J. Wright, T.S. Huang, Y. Ma, Image super-resolution via sparse representation, IEEE Trans. Image Process. 19 (11) (2010) 2861–2873.

[46] J.-H. Yang, X.-L. Zhao, T.-Y. Ji, T.-H. Ma, T.-Z. Huang, Low-rank tensor train for tensor robust principal component analysis, Appl. Math. Comput. 367 (2020) 124783.

[47] J.-H. Yang, X.-L. Zhao, T.-H. Ma, Y. Chen, T.-Z. Huang, M. Ding, Remote sensing image destriping using unidirectional high-order total variation and nonconvex low-rank regularization, J. Comput. Appl. Math. 363 (2020) 124–144.

[48] H.-Y. Zhang, W. He, L.-P. Zhang, H.-F. Shen, Q.-Q. Yuan, Hyperspectral image restoration using low-rank matrix recovery, IEEE Trans. Geosci. Remote Sens. 52 (8) (2014) 4729–4743.

[49] J. Zhang, D.-B. Zhao, W. Gao, Group-based sparse representation for image restoration, IEEE Trans. Image Process. 23 (8) (2014) 3336–3351.

[50] X. Zhang, M. Burger, X. Bresson, S. Osher, Bregmanized nonlocal regularization for deconvolution and sparse reconstruction, SIAM J. Imaging Sci. 3 (3) (2010) 253–276.
[51] X.-L. Zhao, F. Wang, M.K. Ng, A new convex optimization model for multiplicative noise and blur removal, SIAM J. Imaging Sci. 7 (1) (2014) 456–475.
[52] X.-L. Zhao, W. Wang, T.-Y. Zeng, T.-Z. Huang, M.K. Ng, Total variation structured total least squares method for image restoration, SIAM J. Sci. Comput. 35 (6) (2013) 1304–1320.
[53] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, T.-X. Jiang, T.-H. Ma, T.-Y. Ji, Mixed noise removal in hyperspectral image via low-fibered-rank regularization, 2019, pp. 1–16, doi:10.1109/TGRS.2019.2940534.