Applied Mathematics and Computation 341 (2019) 128–147

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

# Total variation with overlapping group sparsity for deblurring images under Cauchy noise

Meng Ding[a], Ting-Zhu Huang[a,*], Si Wang[a], Jin-Jin Mei[b], Xi-Le Zhao[a]

[a] School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, PR China
[b] School of Mathematics and Statistics, Fuyang Normal University, Fuyang, Anhui 236037, PR China

ARTICLE INFO

*Keywords:*
Cauchy noise
Overlapping group sparsity
Total variation
Alternating direction method with multipliers
Majorization minimization

ABSTRACT

The methods based on the total variation are effective for image deblurring and denoising, which can preserve edges and details of images. However, these methods usually produce some staircase effects. In order to alleviate the staircase effects, we propose a new convex model based on the total variation with overlapping group sparsity for recovering blurred images corrupted by Cauchy noise. Moreover, we develop an algorithm under the framework of the alternating direction method with multipliers, and use the majorization minimization to solve subproblems of the proposed algorithm. Numerical results illustrate that the proposed method outperforms other methods both in visual effects and quantitative measures, such as the peak signal-to-noise ratio and the structural similarity index.



## 1. Introduction

Image restoration is a fundamental issue in the image processing, such as denoising [14,21,50], deblurring [24,43], super-resolution [17,55], and image inpainting [5,38,57]. Mathematically speaking, the degraded images are the convolution of the point spread function (PSF) with true images and plus some noise. In most practical applications, obtained images are corrupted by additive Gaussian noise. Actually, in many cases, the noise does not satisfy the Gaussian assumption, such as impulse noise [9,45], multiplicative noise [19,22,63], and Cauchy noise [31,36,48]. In this paper, we consider recovering blurred images corrupted by Cauchy noise, which usually arises in radar and sonar applications, biomedical images, atmospheric and synthetic aperture radar images [31,36,49]. We consider the degradation model as the following linear system:

$$g = Hf + n,$$

where $f \in \mathbb{R}^{n\times n}$ is the original gray-scale image with compacted Lipschitz boundary, $g$ is the observed image, $H$ is a convolution operator, and $n$ is Cauchy noise. The probability density function of Cauchy distribution with $x$ is expressed as

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

where $\delta$ is a localization parameter which corresponds to the median of the distribution, and $\gamma > 0$ is a scale parameter that decides the spread of the distribution around $\delta$ [46]. Without loss of generality, in the following discussion, we consider $\delta = 0$.

* Corresponding author.
*E-mail addresses:* dingmeng56@163.com (M. Ding), tingzhuhuang@126.com, tzhuang@uestc.edu.cn (T.-Z. Huang), uestcsiwang@163.com (S. Wang), meijinjin666@126.com (J.-J. Mei), xlzhao122003@163.com (X.-L. Zhao).

https://doi.org/10.1016/j.amc.2018.08.014
0096-3003/

Recently, several researchers have paid an attention to Cauchy noise. Chang et al. [13] studied a recursive restoration algorithm based on Markov random field model to reconstruct images under Cauchy noise. Wan et al. [56] studied a novel image segmentation method for the color image corrupted by Cauchy noise.

In the literature, approaches based on the total variation (TV) regularization for image restoration have attracted great attention [10,33,35,51,54,59]. For a gray image $f \in \mathbb{R}^{n\times n}$, the discrete gradient operator $\nabla f : \mathbb{R}^{n\times n} \to (\mathbb{R}^{n\times n}, \mathbb{R}^{n\times n})$ is defined by

$$(\nabla f)_{i,j} = ((\nabla_x f)_{i,j}, (\nabla_y f)_{i,j})$$

with

$$(\nabla_x f)_{i,j} = \begin{cases} f_{i+1,j} - f_{i,j}, & \text{if } \; i < n, \\ f_{1,j} - f_{n,j}, & \text{if } \; i = n \end{cases}$$

and

$$(\nabla_y f)_{i,j} = \begin{cases} f_{i,j+1} - f_{i,j}, & \text{if } \; j < n, \\ f_{i,1} - f_{i,n}, & \text{if } \; j = n \end{cases}$$

for $i, j = 1, 2, \ldots, n$, where $f_{i,j}$ is the $(i, j)$th pixel in the image and $\nabla_x$ and $\nabla_y$ are the horizontal and vertical gradient operators, respectively. The discrete TV of $f$ is

$$TV(f) = \|\nabla f\|_1 = \sum_{1\le i,j\le n} |(\nabla f)_{i,j}| = \sum_{1\le i,j\le n} \sqrt{|(\nabla_x f)_{i,j}|^2 + |(\nabla_y f)_{i,j}|^2}.$$

Sciacchitano et al. [52] proposed a TV-based variational model for deblurring degraded images with Cauchy noise. The discrete TV-based model was described as follows:

$$\arg\min_f \frac{\lambda}{2}\left\langle \log\left(\gamma^2 + (Hf-g)^2\right), \mathbf{1}\right\rangle + TV(f), \tag{1}$$

where $\lambda > 0$ is the regularization parameter, which controls the balance between the fidelity term and the regularization term, $\langle\cdot\rangle$ denotes the standard inner product, and $\mathbf{1} \in \mathbb{R}^{n\times n}$ is a matrix whose elements equal 1. If $H$ is the identity operator, the degraded $f$ is only corrupted by Cauchy noise. Since the fidelity term based on Cauchy distribution is non-convex, the solution of (1) depends on the initial guess. To overcome this shortcoming, the authors [52] added a quadratic penalty term, then they introduced the following convex variational model:

$$\arg\min_f \frac{\lambda}{2}\left(\left\langle \log\left(\gamma^2 + (Hf-g)^2\right), \mathbf{1}\right\rangle + \mu \parallel Hf - \hat{f} \parallel_F^2\right) + TV(f), \tag{2}$$

where $\hat{f}$ denotes the image obtained by applying the median filter and $\mu$ is a positive penalty parameter. From [52], we know that if $8\mu\gamma^2 \ge 1$, the objective function in (2) is strictly convex, and the primal-dual algorithm [8] was used to solve the convex model. Mei et al. [44] focused on the non-convex model (1). They developed the alternating direction method with multipliers (ADMM) to solve the non-convex variational model (1) with convergence guarantees, and achieved better performance than the method proposed in [52]. But the solution strongly depends on the initial guess.

Although the TV regularization can preserve fine features and sharp edges, it also produces staircase effects [2,7,12,18,64]. There are many methods which focus on overcoming the limitations of TV, such as replacing the original TV regularization by high-order TV prior [11,41,42,58], framelet prior [6,34], or image restoration via deep learning [62]. In [53], Selesnich and Chen considered the total variation with overlapping group sparsity (OGS-TV) for one-dimensional signal denoising. They applied the majorization minimization (MM) method to solve their model. Then, Liu et al. [39] extended the OGS-TV regularizer to two-dimensional cases for images deblurring under Gaussian noise. Compared with the TV regularization based model, the numerical experiments showed that their method can preserve the edges and overcome the staircase effects effectively.

The TV-based denoising method in [52] considered the sparsity of the gradient of the image. But from Fig. 1, the gradient tends to be group sparse. Therefore, we utilize the sparsity and group sparsity of the gradient to reduce staircase effects. Moreover, in Fig. 2, we present the recovered images by the TV-based method proposed in [52] and our method for removing Cauchy noise. From this figure, it can be seen that OGS-TV leads to a result with fewer staircase effects compared to TV-based method, such as the parts labeled by blue boxes. And the labeled red boxes in the bottom panel of Fig. 2 are shown in detail.

Based on these observations, we consider the total variation with overlapping group sparsity for image deblurring under Cauchy noise. By combining the fidelity term in (2) with the OGS-TV regularizer, we consider the following convex model for deblurring images under Cauchy noise:

$$\arg\min_f \frac{\lambda}{2}\left(\left\langle \log\left(\gamma^2 + (Hf-g)^2\right), \mathbf{1}\right\rangle + \mu \parallel Hf - \hat{f} \parallel_F^2\right) + \Phi(f), \tag{3}$$

where $\Phi(\cdot)$ denotes the OGS-TV function (see the definition in Section 2.1). Then, we develop an efficient algorithm to solve the model under the framework of ADMM. The MM is used for solving the subproblems of the proposed method.

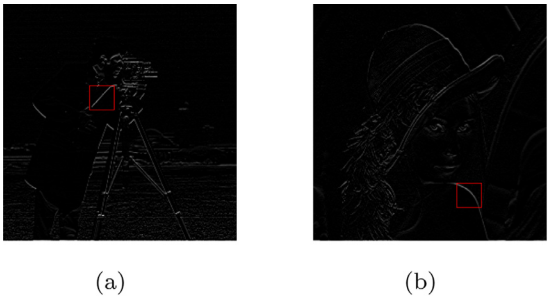



**Fig. 1.** Illustration of the overlapping group sparsity of the gradient of image. (a) the horizontal gradient image of Cameraman and (b) the horizontal gradient image of Lena. The large values of gradient are not isolated, but generally adjacent to other large values, which are shown in the red box. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

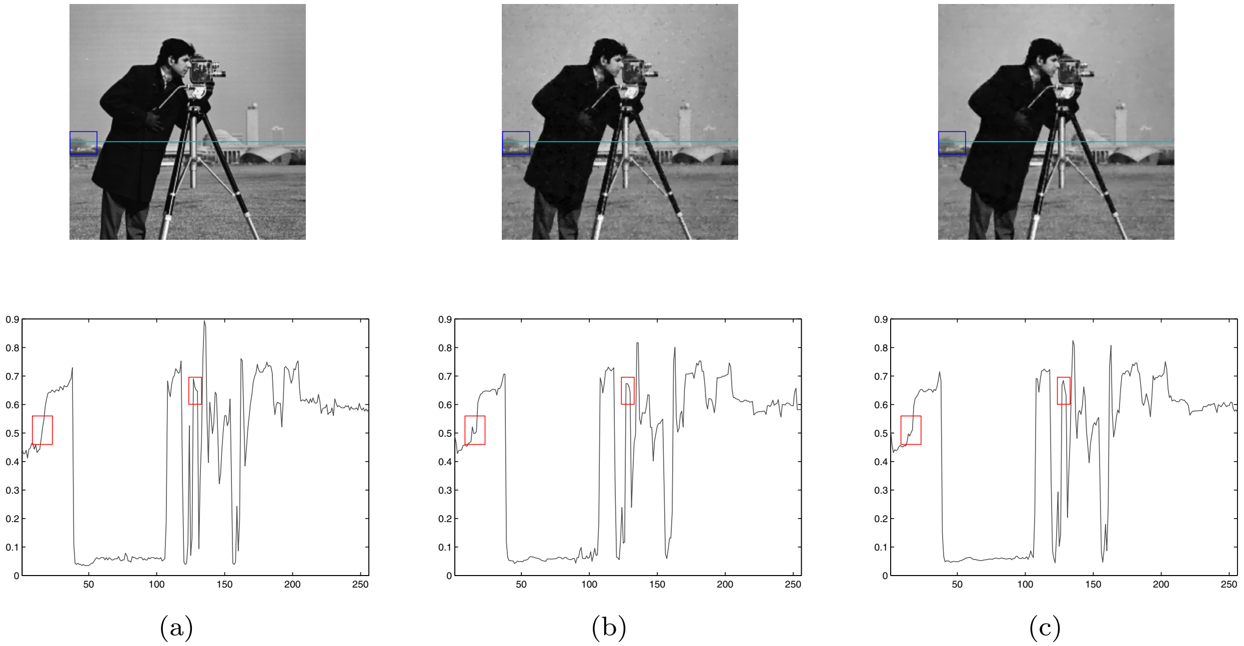



**Fig. 2.** Illustration of the ability of alleviating staircase effects by using the overlapping group sparsity of the gradient. First row: the denoising results on Cauchy noise. Second row: the intensities of a random row (the transversal line) of the image. (a) the original image, (b) the restored image by TV method, and (c) the restored image by our method.

According to our numerical results, we can observe that our new convex model using the OGS-TV regularizer maintains the edge-preserving property of the TV method and overcomes staircase effects.

The paper proceeds as follow. In Section 2, we briefly review the definition of the OGS-TV function, the ADMM algorithm, and MM method. In Section 3, we introduce the convex model for the deblurring image under Cauchy noise and apply ADMM to solve the model. In Section 4, numerical experiments demonstrate that the effectiveness of the proposed method. In Section 5, we analyze the experimental results and discuss the sensitivity of parameters. Finally, we summarize this paper in Section 6.

## 2. Preliminaries

In this section, we introduce the definition of OGS-TV, and review the ADMM and MM algorithms.

### *2.1. OGS-TV*

For the two-dimensional case, such as an image $v \in \mathbb{R}^{n\times n}$, we define a $K$-square-point group ($K$ determines the group size) by

$$\widetilde{v}_{i,j,K,K} = \begin{bmatrix} v_{i-a_1,j-a_1} & v_{i-a_1,j-a_1+1} & \cdots & v_{i-a_1,j+a_2} \\ v_{i-a_1+1,j-a_1} & v_{i-a_1+1,j-a_1+1} & \cdots & v_{i-a_1+1,j+a_2} \\ \vdots & \vdots & \ddots & \vdots \\ v_{i+a_2,j-a_1} & v_{i+a_2,j-a_1+1} & \cdots & v_{i+a_2,j+a_2} \end{bmatrix} \in \mathbb{R}^{K\times K}, \tag{4}$$

where $a_1 = [\frac{K-1}{2}]$, $a_2 = [\frac{K}{2}]$, and $[x]$ denotes the largest integer less than or equal to $x$. Let $v_{i,j,K,K}$ be a $K^2$-vector obtained by arranging all the elements of $\widetilde{v}_{i,j,K,K}$ in lexicographic order. Then the overlapping group sparsity functional of the two-dimensional array is defined by

$$\varphi(v) = \sum_{i=1}^{n}\sum_{j=1}^{n} \|v_{i,j,K,K}\|_2. \tag{5}$$

Consequently, we set the regularization term $\Phi(f)$ as

$$\Phi(f) = \varphi(\nabla_x f) + \varphi(\nabla_y f). \tag{6}$$

We call the regularizer $\Phi(f)$ as OGS-TV. Note that if $K = 1$, in the literature, $\Phi(f)$ is usually mentioned as the anisotropic TV function.

### 2.2. ADMM

ADMM is a simple but powerful algorithm that is well suited to solve the convex optimization problem [1,29]. The algorithm solves the following optimization problem with a linear constraint:

$$\begin{aligned} &\min \Psi_1(x_1) + \Psi_2(x_2), \\ &s.t.\ A_1x_1 + A_2x_2 = b, \\ &x_i \in \mathcal{X}_i, \quad i = 1, 2, \end{aligned} \tag{7}$$

where $\mathcal{X}_i \subseteq \mathbb{R}^{m_i}$ are the nonempty closed convex set, $\Psi_i : \mathcal{X}_i \to \mathbb{R}$ denote the closed convex function, $A_i \in \mathbb{R}^{l\times m_i}$ are the linear transform, and $b \in \mathbb{R}^l$ is a given vector.

By introducing a Lagrangian multiplier $p \in \mathbb{R}^l$ to the linear constraint, we get the following augmented Lagrangian function:

$$\mathcal{L}(x_1, x_2, p) = \Psi_1(x_1) + \Psi_2(x_2) + \langle p, b - A_1x_1 - A_2x_2\rangle + \frac{\beta}{2}\|b - A_1x_1 - A_2x_2\|_2^2, \tag{8}$$

where $\beta$ is a positive penalty parameter.

According to the framework of the ADMM, a saddle point of (8) can be achieved by the alternative minimizing scheme. Then, we obtain the following ADMM algorithm:

**Algorithm 1** ADMM for the minimization problem (7).

1: Initialize $x_1^0$, $x_2^0$, $p^0$; set $\beta > 0$.
2: For $k = 1, 2, \ldots$, compute $x_1^{k+1}$, $x_2^{k+1}$, $p^{k+1}$ by

$$\begin{aligned} x_1^{k+1} &= \arg\min_{x_1\in\mathcal{X}_1} \Psi_1(x_1) + \frac{\beta}{2}\|b - A_1x_1 - A_2x_2^k + \frac{p^k}{\beta}\|_2^2 \\ x_2^{k+1} &= \arg\min_{x_2\in\mathcal{X}_2} \Psi_2(x_2) + \frac{\beta}{2}\|b - A_1x_1^{k+1} - A_2x_2 + \frac{p^k}{\beta}\|_2^2 \\ p^{k+1} &= p^k + \beta(b - A_1x_1^{k+1} - A_2x_2^{k+1}). \end{aligned}$$

3: Until satisfying a stopping criterion.

### 2.3. MM

Instead of minimizing a difficult minimization problem $O(v)$ directly, the MM solves a sequence of easier optimization problems $P(v, v^k)$ $(k = 0, 1, 2, \ldots)$. Generally, a MM iterative algorithm for minimizing $O(v)$ is written as

$$v^{k+1} = \arg\min_v P(v, v^k), \tag{9}$$

where $P(v, v')$ is a majorizer of $O(v)$, i.e., $P(v, v') \geq O(v)$ for all $v$, $v'$, and $P(v, v) = O(v)$. When $O(v)$ is convex, under some mild conditions, the sequence $\{v^k\}$ produced by (9) converges to the minimizer of $O(v)$ [25,47].

Next, we consider a minimization problem as

$$\min_{v} O(v), \tag{10}$$

where $O(v) = \frac{\beta}{2}\|v - v_0\|_2^2 + \varphi(v)$, $v \in \mathbb{R}^{n^2}$, $\beta$ is a positive parameter and the functional $\varphi(\cdot)$ is defined as (5). Obviously, the problem $O(v)$ in (10) is convex. In order to obtain an efficient solution of (10) by the MM, we first find a majorizer of $O(v)$. Here, we just need to find a majorizer of $\varphi(v)$. Note that

$$\frac{1}{2}\left(\frac{1}{\|u\|_2}\|v\|_2^2 + \|u\|_2\right) \geq \|v\|_2, \tag{11}$$

for all $v$ and $u \neq 0$, and with equality when $v = u$. Using (11) for each group, we get a majorizer of $\varphi(v)$,

$$Q(v, u) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left[\frac{1}{\|u_{i,j,K,K}\|_2}\|v_{i,j,K,K}\|_2^2 + \|u_{i,j,K,K}\|_2\right], \tag{12}$$

with $Q(v, u) \geq \varphi(v)$, $Q(u, u) = \varphi(u)$, provided that $\|u_{i,\, j,\, K,\, K}\|_2 \neq 0$ for all $i$, $j$. After a simple calculation, $Q(v, u)$ can be written as

$$Q(v, u) = \frac{1}{2}\|\Lambda(u)v\|_2^2 + C, \tag{13}$$

where $C$ is a constant that does not depend on $v$, and $\Lambda(u)$ is a diagonal matrix with each diagonal component

$$[\Lambda(u)]_{l,l} = \sqrt{\sum_{i=-a_1}^{a_2}\sum_{j=-a_1}^{a_2}\left[\sum_{k_1=-a_1}^{a_2}\sum_{k_2=-a_1}^{a_2}|u_{l-i+k_1,l-j+k_2}|^2\right]^{-\frac{1}{2}}} \tag{14}$$

with $l = 1,\ 2,\ \ldots,\ n^2$. The entries of $\Lambda$ can be easily computed by the convolution operation.

Then the majorizer of $O(v)$ can be expressed as

$$P(v, u) = \frac{\beta}{2}\|v - v_0\|_2^2 + Q(v, u) = \frac{\beta}{2}\|v - v_0\|_2^2 + \frac{1}{2}\|\Lambda(u)v\|_2^2 + C, \tag{15}$$

with $P(v, u) \geq O(v)$ for all $v$, $u$, and $P(u, u) = O(u)$. To minimize $O(v)$, the MM aims to iteratively solve

$$v^{k+1} = \arg\min_{v} \frac{\beta}{2}\|v - v_0\|_2^2 + \frac{1}{2}\|\Lambda(v^k)v\|_2^2, \quad k = 0, 1, 2, \ldots. \tag{16}$$

Then we get the solution of (16) easily

$$v^{k+1} = \left(I + \frac{1}{\beta}\Lambda^2(v^k)\right)^{-1} v_0, \quad k = 1, 2, \ldots, \tag{17}$$

where $I$ is an identity matrix with the same size of $\Lambda^2(v^k)$. To summarize, we obtain Algorithm 2 for solving the problem (10).

**Algorithm 2** The MM method for solving (10).

1: Initialize $v^0 = v_0$, $k = 0$, set $\beta$, $K$, maximum iteration $N$.

2: Iteration

$$[\Lambda(u)]_{l,l} = \sqrt{\sum_{i=-a_1}^{a_2}\sum_{j=-a_1}^{a_2}\left[\sum_{k_1=-a_1}^{a_2}\sum_{k_2=-a_1}^{a_2}|u_{l-i+k_1,l-j+k_2}|^2\right]^{-\frac{1}{2}}}$$

$$v^{k+1} = \left(I + \frac{1}{\beta}\Lambda^2(v^k)\right)^{-1} v_0$$

$$k = k + 1.$$

3: If satisfies iteration $N$, return $v^{k+1}$ and stop.

## 3. The proposed algorithm

Since the fidelity term contains $\log\left(\gamma^2+(Hf-g)^2\right)$, the model (1) is non-convex. To deal with the difficulty on non-convexity, inspired by Sciacchitano et al. [52], we add an extra quadratic term and propose the following convex variation model:

$$\arg\min_f \frac{\lambda}{2}\left(\left\langle \log\left(\gamma^2+(Hf-g)^2\right),\mathbf{1}\right\rangle+\mu\parallel Hf-\hat{f}\parallel_F^2\right)+\varphi(\nabla_x f)+\varphi(\nabla_y f). \tag{18}$$

The following theorem establishes the existence and uniqueness of the solution of the proposed model (18).

**Theorem 1.** *Assume that $Null(H)\cap Null(\nabla)=\{0\}$ with $Null(\cdot)$ denotes the null space, the model (18) has at least one solution. If $8\mu\gamma^2\geq 1$, there exists a unique solution.*

**Proof.** Denote by $E(f)$ the objective function of (18). It is clear that $E(f)$ is proper, continuous. According to the Weierstrass' theorem [3], it remains only to show the coercivity of $E(f)$, i.e., for every sequence $\{f^k\}$ such that $\|f^k\|_F\to\infty$, we have $\lim_{k\to\infty}E(f^k)=\infty$. We prove it by contradiction. Suppose that there exists a subsequence of $\{f^k\}$ (also denoted as $\{f^k\}$) that $\{E(f^k)\}$ is bounded, we have that $\{\|Hf^k\|_F\}$ and $\{\Phi(f^k)\}$ are bounded, then $\{\|\nabla f^k\|_1\}$ is bounded. According to the assumption $Null(H)\cap Null(\nabla)=\{0\}$, the sequence $\{f^k\}$ is a bounded sequence, which is a contradiction. So the model (18) has at least one minimizer. From the previous definition (6), we can get that OGS-TV is convex. Similar to the proof in [52], when $8\mu\gamma^2\geq 1$, the objective function (18) is strictly convex and the solution is unique. □

As the proposed model is a convex optimization problem, there are many efficient algorithms to solve the model (18), such as the Bregman method [27,61], the primal-dual algorithm [8,23], the proximal splitting method [15], and the alternating direction method with multipliers (ADMM) [4,26,32]. Here we just discuss ADMM.

Note that the pixel values of any true digital image can attain only a finite number of values. So it is natural to require all pixels of the restored image in a certain interval $[a, b]$, see [10] for more details. For the easy computation and the certified results in [10], we only consider image located on the range [0,1]. We define a projection operator $\mathcal{P}_\Gamma$ on the set $\Gamma=\{f\in\mathbb{R}^{n\times n}|0\leqslant f_{i,j}\leqslant 1\}$,

$$\mathcal{P}_\Gamma(f)_{i,j}=\begin{cases}0, & f_{i,j}<0,\\ f_{i,j}, & f_{i,j}\in[0,1],\\ 1, & f_{i,j}>1.\end{cases} \tag{19}$$

And for convenience, we define an indication function for a set

$$\daleth_\Xi=\begin{cases}0, & x\in\Xi,\\ +\infty, & x\notin\Xi.\end{cases} \tag{20}$$

where $\Xi=[a,b]$.

For the model (18), by introducing four auxiliary variables $z$, $v_x$, $v_y$, and $w$ together with the indication function (20), we get the equivalent constrained minimization problem

$$\begin{aligned}&\arg\min_f \frac{\lambda}{2}\left(\left\langle \log\left(\gamma^2+(z-g)^2\right),\mathbf{1}\right\rangle+\mu\parallel z-\hat{f}\parallel_F^2\right)+\varphi(v_x)+\varphi(v_y)+\daleth_\Xi(w)\\ &s.t.\quad z=Hf, v_x=\nabla_x f, v_y=\nabla_y f, w=f.\end{aligned} \tag{21}$$

Then, (21) can be further expressed in a compact form

$$\begin{aligned}&\min \Psi_1(x_1)+\Psi_2(x_2),\\ &s.t.\ A_1x_1+A_2x_2=b,\end{aligned} \tag{22}$$

where

$$x_1=f,\ x_2=\begin{pmatrix}z\\ v_x\\ v_y\\ w\end{pmatrix},\ A_1=\begin{pmatrix}H\\ \nabla_x\\ \nabla_y\\ I\end{pmatrix},\quad A_2=\begin{pmatrix}-I&&&\\ &-I&&\\ &&-I&\\ &&&-I\end{pmatrix},\quad b=\begin{pmatrix}0\\0\\0\\0\end{pmatrix},$$

$\Psi_1(x_1)=0$, and $\Psi_2(x_2)=\frac{\lambda}{2}(\langle\log(\gamma^2+(z-g)^2),\mathbf{1}\rangle+\mu\parallel z-\hat{f}\parallel_F^2)+\varphi(v_x)+\varphi(v_y)+\daleth_\Xi(w)$.

Let $u_1$, $u_2$, $u_3$, and $u_4$ be the Lagrangian multipliers, we then form the augmented Lagrangian function as follows:

$$\begin{aligned}L(v_x,v_y,w,z,f;u_1,u_2,u_3,u_4)=&\ \varphi(v_x)+\langle u_1,\nabla_x f-v_x\rangle+\frac{\beta_1}{2}\|\nabla_x f-v_x\|_F^2\\ &+\varphi(v_y)+\langle u_2,\nabla_y f-v_y\rangle+\frac{\beta_1}{2}\|\nabla_y f-v_y\|_F^2\end{aligned}$$

$$+\frac{\lambda}{2}\left(\left\langle \log\left(\gamma^2+(z-g)^2\right),\mathbf{1}\right\rangle+\mu\parallel z-\hat{f}\parallel_F^2\right)+\langle u_3, Hf-z\rangle$$
$$+\frac{\beta_2}{2}\|Hf-z\|_F^2+\langle u_4, f-w\rangle+\frac{\beta_3}{2}\|f-w\|_F^2+\intercal_\Xi(w), \tag{23}$$

where $\beta_1,\ \beta_2,\ \beta_3>0$ are penalty parameters.

Then the ADMM for solving (21) is to update variables alternatively by minimizing the augmented Lagrange function (23) over $z,\ v_x,\ v_y,\ w,$ and $f$. The optimization problem is well structured since all the variables are separated into two groups, ($f$) and ($z,\ v_x,\ v_y,\ w$). The variables ($z,\ v_x,\ v_y,\ w$) are independent so that they can be solved respectively. In particular, the iterative scheme for solving (21) can be expressed as follows:

1. $v_x,\ v_y$-subproblem
   From the above discussion, it is obvious to see that the $v_x$-subproblem and the $v_y$-subproblem are independent, we handle them respectively,

$$\begin{aligned} v_x^{k+1} &= \arg\min_{v_x}\varphi(v_x)+\langle u_1^k, \nabla_x f^k-v_x\rangle+\frac{\beta_1}{2}\|\nabla_x f^k-v_x\|_F^2 \\ &= \arg\min_{v_x}\varphi(v_x)+\frac{\beta_1}{2}\|v_x-\nabla_x f^k-\frac{u_1^k}{\beta_1}\|_F^2, \end{aligned} \tag{24}$$

$$\begin{aligned} v_y^{k+1} &= \arg\min_{v_y}\varphi(v_y)+\langle u_2^k, \nabla_y f^k-v_y\rangle+\frac{\beta_1}{2}\|\nabla_y f^k-v_y\|_F^2 \\ &= \arg\min_{v_y}\varphi(v_y)+\frac{\beta_1}{2}\|v_y-\nabla_y f^k-\frac{u_2^k}{\beta_1}\|_F^2. \end{aligned} \tag{25}$$

   Since the problems (24) and (25) match the framework of the problem (10), thus we can apply Algorithm 2.
2. $z$-subproblem

$$z^{k+1}=\arg\min_z\frac{\lambda}{2}\left(\left\langle \log\left(\gamma^2+(z-g)^2\right),\mathbf{1}\right\rangle+\mu\parallel z-\hat{f}\parallel_F^2\right)+\frac{\beta_2}{2}\|z-Hf^k-\frac{u_3^k}{\beta_2}\|_F^2. \tag{26}$$

   As there exists the second order derivative for the objective function in (26), we determine to solve the Eq.(26) by using the Newton method. The Newton method will converge, provided the initial point is close enough to the root of the derivative of the $z$-subproblem.
3. $w$-subproblem

$$\begin{aligned} w^{k+1} &= \arg\min_w\langle u_4^k, f^k-w\rangle+\frac{\beta_3}{2}\|f^k-w\|_F^2+\intercal_\Xi(w) \\ &= \arg\min_w\frac{\beta_3}{2}\|w-f^k-\frac{u_4^k}{\beta_3}\|_F^2+\intercal_\Xi(w). \end{aligned} \tag{27}$$

   By the sample projection, we get the minimizer explicitly by

$$w^{k+1}=\mathcal{P}_\Gamma\left[f^k+\frac{u_4^k}{\beta_3}\right]. \tag{28}$$

4. $f$-subproblem
   The $f$-subproblem corresponds to the following optimization problem:

$$\begin{aligned} f^{k+1} &= \arg\min_f\frac{\beta_1}{2}\|\nabla_x f-v_x^{k+1}+\frac{u_1^k}{\beta_1}\|_F^2+\frac{\beta_1}{2}\|\nabla_y f-v_y^{k+1}+\frac{u_2^k}{\beta_1}\|_F^2 \\ &\quad+\frac{\beta_2}{2}\|Hf-z^{k+1}+\frac{u_3^k}{\beta_2}\|_F^2+\frac{\beta_3}{2}\|f-w^{k+1}+\frac{u_4^k}{\beta_3}\|_F^2. \end{aligned} \tag{29}$$

   The function is quadratic in $f$. Thus, the optimal solution of (29) satisfies

$$Af^{k+1}=B, \tag{30}$$

   where

$$\begin{aligned} A &= \beta_1(\nabla_x^T\nabla_x+\nabla_y^T\nabla_y)+\beta_2H^TH+\beta_3I, \\ B &= \nabla_x^T(\beta_1v_x^{k+1}-u_1^k)+\nabla_y^T(\beta_1v_y^{k+1}-u_2^k)+H^T(\beta_2z_y^{k+1}-u_3^k)+\beta_3\left(w^{k+1}-\frac{u_4^k}{\beta_3}\right). \end{aligned} \tag{31}$$

It should be pointed that $H$, $\nabla_x$, and $\nabla_y$ in image restoration are highly structured. In specific, their exact structures depend on the imposed boundary conditions. Here, we use the periodic boundary condition, so $H$, $\nabla_x$, and $\nabla_y$ are the block circulant with circulating block (BCCB) structure. After applying 2D fast Fourier transforms, the optimal $f$ is formed directly

$$f^{k+1} = \mathscr{F}^{-1}\left(\frac{\mathscr{F}(B)}{\mathscr{F}(A)}\right). \tag{32}$$

5. Updating multipliers via

$$\begin{cases} u_1^{k+1} = u_1^k + \beta_1(\nabla_x f^{k+1} - v_x^{k+1}), \\ u_2^{k+1} = u_2^k + \beta_1(\nabla_y f^{k+1} - v_y^{k+1}), \\ u_3^{k+1} = u_3^k + \beta_2(Hf^{k+1} - z^{k+1}), \\ u_4^{k+1} = u_4^k + \beta_3(f^{k+1} - w^{k+1}). \end{cases} \tag{33}$$

Finally, we summary the proposed ADMM in Algorithm 3.

**Algorithm 3** ADMM for solving (18).

1: Initialize $v_x^0$, $v_y^0$, $w^0$, $z^0$, $f^0$, $u_1^0$, $u_2^0$, $u_3^0$, $u_4^0$, $\lambda$, $\mu$, $\beta_1$, $\beta_2$, $\beta_3$, $K$, maximum inner iteration $N$; set $k = 0$.
2: Iteration
3: Compute $v_x^{k+1}$ according to (24) using Algorithm 2;
4: Compute $v_y^{k+1}$ according to (25) using Algorithm 2;
5: Compute $z^{k+1}$ using Newton method;
6: Compute $w^{k+1}$ according to (28);
7: Compute $f^{k+1}$ according to (32);
8: Update multipliers according to (33);
9: If $f^{k+1}$ satisfies the stopping criteria, return $f^{k+1}$ and stop.

The minimization problem (21) fits the framework (7) of ADMM. Owing to the convexity of the proposed model, the convergence of the proposed algorithm is theoretically guaranteed [20,30].

## 4. Numerical experiments

In this section, we present several numerical experiments to demonstrate the performance of the proposed method for restoring blurred images corrupted by Cauchy noise. Firstly, we focus on denoising cases, then we consider the deblurring cases under Cauchy noise. All test images are shown in Fig. 3, twelve 256-by-256 gray-scale images. All numerical experiments are performed on Windows 10 64-bit and Matlab R2012a running on a desktop equipped with an Intel(R) Core(TM) i7-8700K CPU with 3.7 GHz and 16 GB of RAM.

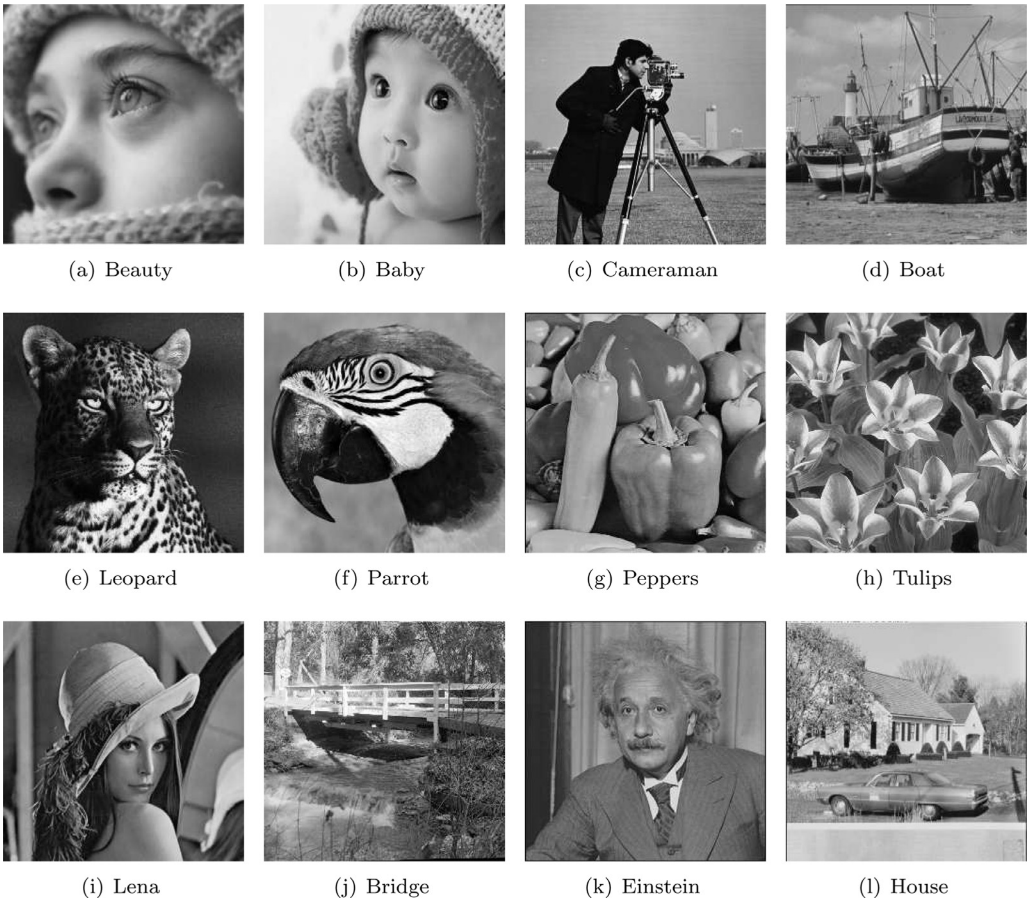
We compare our method OGS-TV with other two well-known methods: the median filter and the convex variational approach proposed in [52] ("TV" for short). Furthermore, we also compare our method with other methods, such as the myriad filter [28], the SURE-LET [40], the BM3D [16], and Mei's method [44].

The qualities of recovered images are measured by the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [60], which are defined as

$$\mathrm{PSNR} = 10\log_{10}\frac{Max_{f,\tilde{f}}^2}{\|f-\tilde{f}\|_F^2}, \quad \mathrm{SSIM} = \frac{2\mu_f\mu_{\tilde{f}}(2\sigma + c_2)}{(\mu_f^2+\mu_{\tilde{f}}^2+c_1)(\sigma_f^2+\sigma_{\tilde{f}}^2+c_2)},$$

where $f$ and $\tilde{f}$ are the original image and the restored image, respectively, $Max_{f,\tilde{f}}$ is the maximum possible pixel value of the image $f$ and $\tilde{f}$, $\mu_f$ and $\mu_{\tilde{f}}$ denote their means, $\sigma_f^2$ and $\sigma_{\tilde{f}}^2$ denote their variances, $\sigma$ is the covariance of $\tilde{f}$ and $f$, and $c_1$, $c_2 > 0$ are constants. The value of PSNR satisfies the human subjective sensation, and the higher PSNR value, the better quality of the restored image. The value of SSIM conforms with the quality perception of the human visual system. The characteristic of the restored image is more similar to the original image if the SSIM value is closer to 1.

*Parameters setting.* The parameter settings for our method are as follows. We remark that, in denoising case, $H$ is the identity operator, the degraded $f$ is only corrupted by Cauchy noise. So in (23), there are three parameters $\lambda$, $\beta_1$, and $\beta_2$ to tune by hand. We set the group size $K = 3$ in OGS-TV, the inner iteration $N = 10$ in the MM method, the iteration number as 5 in the Newton's method, $\lambda \in [4, 20]$, $\beta_1 = 600$, $\beta_2 = 100$, and $\beta_3 = 50$. Since $\gamma$ depends on the noise level, we use the same $\gamma$ for all experiments under the same noise level. Numerical experiments show that our method is robust with respect to $\mu$. Therefore, we choose $\mu$ such that the convexity condition is just satisfied, i.e., $8\mu\gamma^2 = 1$. In addition, since the regularization parameter $\lambda$ balances between the fidelity term and the regularization term, we manually tune it in order to obtain the highest PSNR value of the restored images.



(a) Beauty (b) Baby (c) Cameraman (d) Boat
(e) Leopard (f) Parrot (g) Peppers (h) Tulips
(i) Lena (j) Bridge (k) Einstein (l) House

**Fig. 3.** Original images.

We terminate our algorithm by the following stopping condition,

$$\frac{|E(f^{k+1}) - E(f^k)|}{|E(f^k)|} < 10^{-5},$$

where $E$ is the objective function of (18).

*4.1. Image denoising*

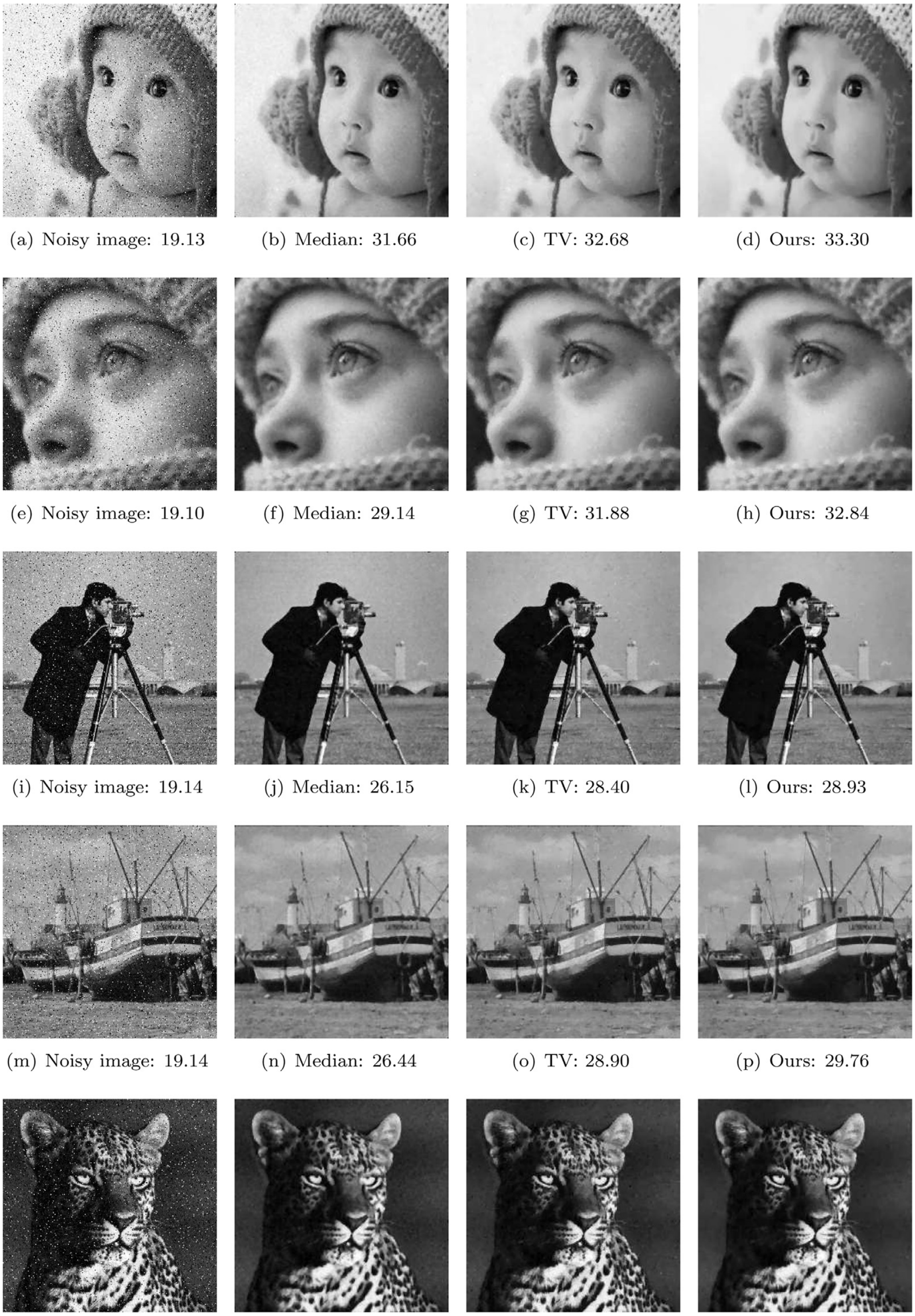
In this subsection, we focus on the denoising cases. Since a standard Cauchy random variable is obtained by the ratio of two independent standard normal variables, so we generate the noisy image $g$ by using the following degradation:

$$g = f + v = f + \xi \frac{\eta_1}{\eta_2},$$

where $v$ represents the Cauchy distribution, $\xi > 0$ represents the noise level, and $\eta_1$ and $\eta_2$ are independent random variables following Gaussian distribution with mean 0 and variance 1. Empirically, we set $\xi = \gamma^2$ for the good experimental performance.

In Figs. 4 and 6, we show the restored images by different methods for removing Cauchy noise with the noise level $\xi = 0.02$ and $\xi = 0.04$, respectively. Although the median filter reduces the noise effectively, it also oversmoothes the edges. From the restored images, it is obvious that the TV method produces staircase effects in the denoising results, while the proposed method overcomes this drawback. Especially, we can see that the background are rough in Fig. 4 (j), (k), (n), (o), (r), and (s). However, our method can keep the background smooth commendably.

To further illustrate clearly the superiority of the proposed OGS-TV method, we show the zoom-in regions of recovered images in Figs. 5 and 7. From these figures, we can see that the median filter recovers the background worst. The TV method generates staircase effects, such as the nose of Baby, the top of Cameraman, and the funnel of Boat. And our method can overcome staircase effects, while keeping the edges and details.



(a) Noisy image: 19.13 (b) Median: 31.66 (c) TV: 32.68 (d) Ours: 33.30

(e) Noisy image: 19.10 (f) Median: 29.14 (g) TV: 31.88 (h) Ours: 32.84

(i) Noisy image: 19.14 (j) Median: 26.15 (k) TV: 28.40 (l) Ours: 28.93

(m) Noisy image: 19.14 (n) Median: 26.44 (o) TV: 28.90 (p) Ours: 29.76

(q) Noisy image: 19.13 (r) Median: 25.47 (s) TV: 27.98 (t) Ours: 28.37

**Fig. 4.** Comparison of restored images from different methods for removing Cauchy noise. The number under images represents the PSNR (dB) of images. First column: noisy images ($\xi = 0.02$); second column: restored images by the median filter; third column: restored images by TV method ($\lambda = 0.8$ (Baby); 0.8 (Beauty); 0.9 (Cameraman); 0.9 (Boat); 0.9 (Leopard), $\mu = 6.25$, and $\gamma = \frac{\sqrt{2}}{10}$); fourth column: restored images by our method ($\lambda = 3$ (Baby); 4 (Beauty); 4 (Cameraman); 4 (Boat); 4 (Leopard), $\mu = 6.25$, and $\gamma = \frac{\sqrt{2}}{10}$).

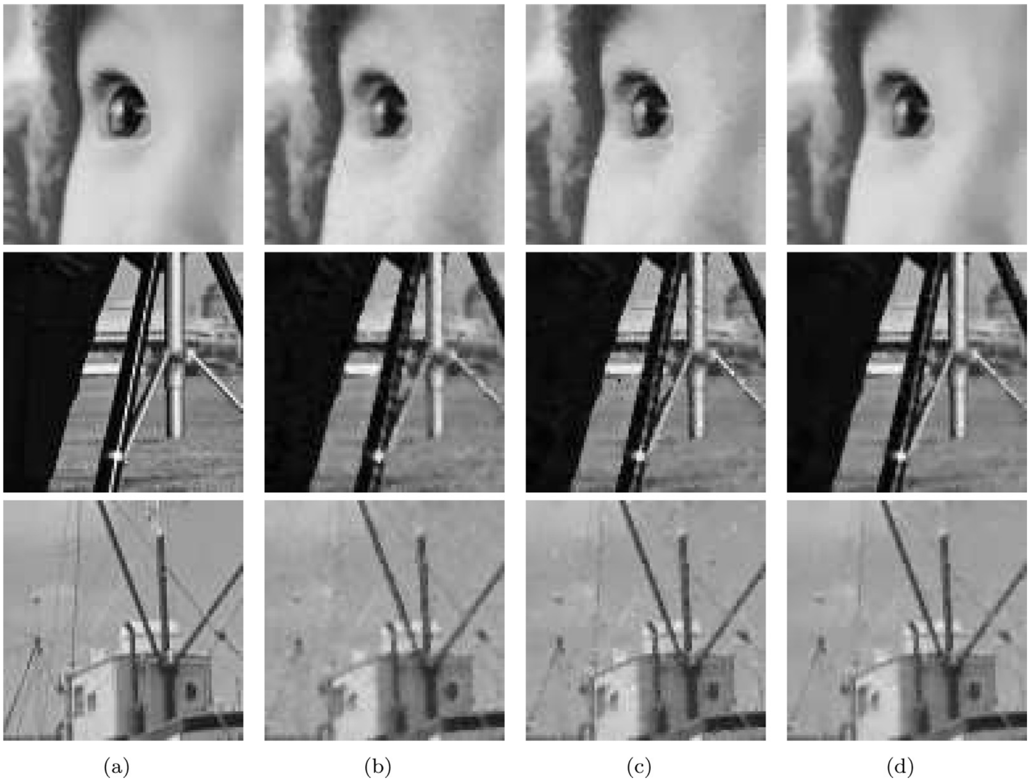



**Fig. 5.** Zoomed vision of restored images in Fig. 4. (a) original images, (b) restored images by the median filter, (c) restored images by TV method, and (d) restored images by our method.

**Table 1**
The PSNR (dB) and SSIM values for the noisy images and restored images by different methods ($\xi = 0.02$).

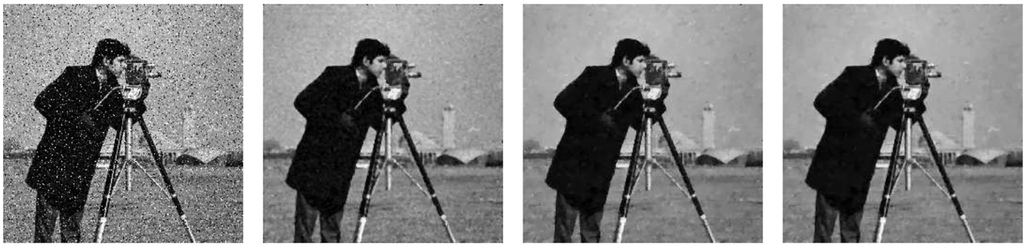
| Image | PSNR | | | | SSIM | | | |
|---|---|---|---|---|---|---|---|---|
| | Noisy | Median | TV | Ours | Noisy | Median | TV | Ours |
| Baby | 19.13 | 31.66 | 32.68 | **33.34** | 0.3060 | 0.8873 | 0.9075 | **0.9328** |
| Beauty | 19.10 | 29.14 | 31.88 | **32.84** | 0.3973 | 0.9134 | 0.9101 | **0.9275** |
| Parrot | 19.08 | 27.17 | 29.24 | **29.85** | 0.3973 | 0.8328 | 0.8738 | **0.8883** |
| Peppers | 19.17 | 29.41 | 30.79 | **31.27** | 0.3663 | 0.8531 | 0.8842 | **0.8900** |
| Cameraman | 19.14 | 26.15 | 28.40 | **28.93** | 0.3567 | 0.7941 | 0.8437 | **0.8781** |
| Tulips | 19.15 | 26.71 | 28.70 | **28.87** | 0.5083 | 0.8319 | 0.8654 | **0.8737** |
| Leopard | 19.13 | 25.47 | 27.98 | **28.37** | 0.4717 | 0.7757 | 0.8233 | **0.8435** |
| Lena | 19.14 | 28.31 | 29.68 | **30.67** | 0.3924 | 0.8443 | 0.8734 | **0.8874** |
| Bridge | 19.14 | 22.66 | 25.74 | **26.03** | 0.5938 | 0.6315 | 0.8045 | **0.8217** |
| Einstein | 19.12 | 29.76 | 30.88 | **31.48** | 0.3000 | 0.7856 | 0.8043 | **0.8312** |
| Boat | 19.14 | 26.44 | 28.90 | **29.76** | 0.4170 | 0.7797 | 0.8394 | **0.8624** |
| House | 19.06 | 24.67 | 27.85 | **28.22** | 0.4452 | 0.7485 | 0.8179 | **0.8466** |

For comparing the performance quantitatively, Tables 1 and 2 list the PSNR and SSIM values of the noisy images and the restored images by different methods. Obviously, comparing with the other two methods, we always obtain the highest values with respect to PSNR and SSIM. In Table 3, we show the comparison between the TV method and our method in terms of PSNR, SSIM, and Time (in seconds) values. These three images are corrupted by Cauchy noise with the noise level $\xi = 0.02$. We can see that although TV method is faster than ours, our method achieves higher PSNR and SSIM values.

For a fair comparison, we compare our method with Mei's method for solving model (1). In Fig. 8, we show the restored images by our method and Mei's method for removing Cauchy noise. We can see that Mei's method achieves better performance than our method. But when we set $\mu = 0$, i.e., our model is non-convex, our method outperforms Mei's method in keeping details, see Fig. 8 (e). To compare the sensitivity of our convex model and Mei's non-convex model to the initial guess, we test on three initial guesses $f_0$ including the random image, the filtered image (the recovered image by the median filter), and the observed image. In Table 4, we list the values of PSNR and SSIM by using different initial guesses for

(a) Noisy image: 16.23 (b) Median: 28.72 (c) TV: 30.41 (d) Ours: 30.99

(e) Noisy image: 16.21 (f) Median: 27.65 (g) TV: 30.32 (h) Ours: 31.19



(i) Noisy image: 16.25 (j) Median: 24.91 (k) TV: 26.76 (l) Ours: 27.31

(m) Noisy image: 16.26 (n) Median: 25.10 (o) TV: 26.63 (p) Ours: 27.16

(q) Noisy image: 16.27 (r) Median: 24.25 (s) TV: 25.87 (t) Ours: 26.28

**Fig. 6.** Comparison of restored images from different methods for removing Cauchy noise. The number under images represents the PSNR (dB) of images. First column: noisy images ($\xi = 0.04$); second column: restored images by the median filter; third column: restored images by TV method ($\lambda = 0.8$ (Baby); 0.9 (Beauty); 1 (Cameraman); 1.1 (Boat); 1 (Leopard), $\mu = 3.125$, and $\gamma = 0.2$); fourth column: restored images by our method ($\lambda = 4$ (Baby); 4 (Beauty); 5 (Cameraman); 5 (Boat); 5 (Leopard), $\mu = 3.125$, and $\gamma = 0.2$).

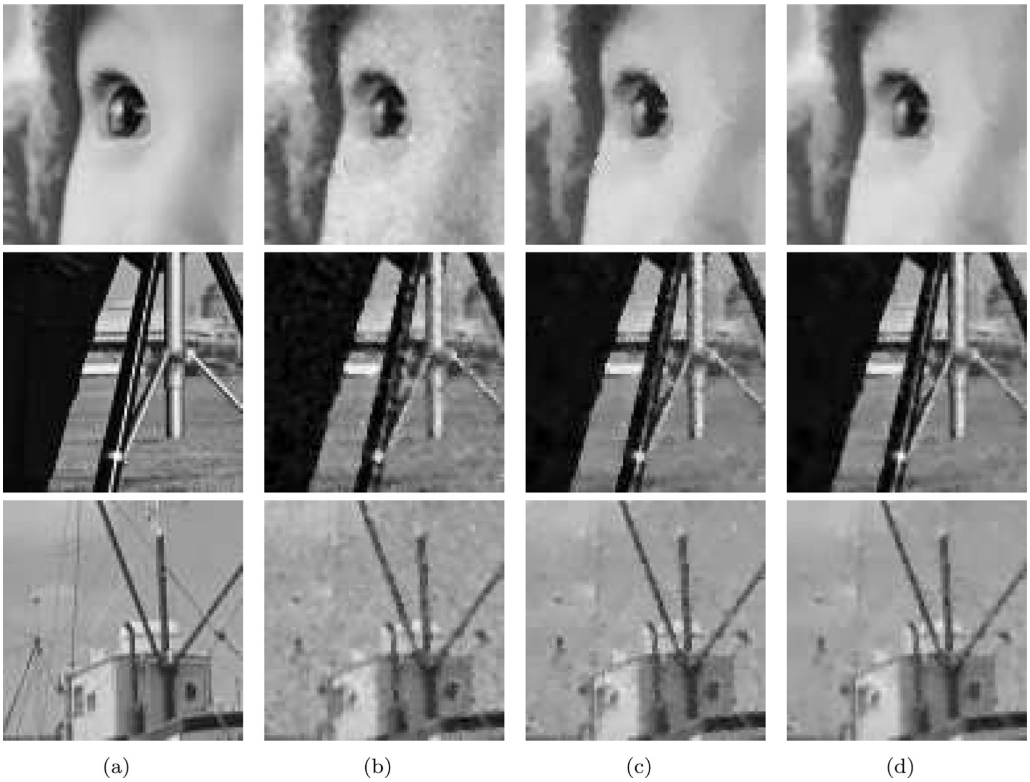


**Fig. 7.** Zoomed vision of restored images in Fig. 6. (a) original images, (b) restored images by the median filter, (c) restored images by TV method, and (d) restored images by our method.

**Table 2**
The PSNR (dB) and SSIM values for the noisy images and restored images by different methods ($\xi = 0.04$).

| Image | PSNR | | | | SSIM | | | |
|---|---|---|---|---|---|---|---|---|
| | Noisy | Median | TV | Ours | Noisy | Median | TV | Ours |
| Baby | 16.23 | 28.72 | 30.41 | **30.99** | 0.1897 | 0.7562 | 0.8779 | **0.8950** |
| Beauty | 16.21 | 27.65 | 30.32 | **31.18** | 0.1591 | 0.7914 | 0.8758 | **0.9126** |
| Parrot | 16.21 | 25.57 | 27.28 | **27.80** | 0.2805 | 0.7229 | 0.8169 | **0.8334** |
| Peppers | 16.28 | 27.18 | 28.60 | **28.91** | 0.2430 | 0.7474 | 0.8351 | **0.8545** |
| Cameraman | 16.25 | 24.91 | 26.76 | **27.31** | 0.2443 | 0.6715 | 0.8020 | **0.8214** |
| Tulips | 16.26 | 25.40 | 26.57 | **26.72** | 0.3608 | 0.7612 | 0.8039 | **0.8129** |
| Leopard | 16.27 | 24.25 | 25.87 | **26.28** | 0.3490 | 0.6912 | 0.7791 | **0.7941** |
| Lena | 16.26 | 26.91 | 27.53 | **28.26** | 0.2658 | 0.7433 | 0.8219 | **0.8308** |
| Bridge | 16.24 | 21.95 | 23.84 | **24.08** | 0.4337 | 0.5867 | 0.7053 | **0.7273** |
| Einstein | 16.25 | 27.47 | 28.51 | **28.99** | 0.1783 | 0.6760 | 0.7537 | **0.7648** |
| Boat | 16.26 | 25.10 | 26.63 | **27.16** | 0.2849 | 0.6858 | 0.7638 | **0.7929** |
| House | 16.18 | 23.92 | 25.84 | **26.11** | 0.3139 | 0.6604 | 0.7577 | **0.7900** |

**Table 3**
Comparisons of the performance of TV and our methods on PSNR (dB), SSIM, and Time (in seconds).

| Image | TV | | | Ours | | |
|---|---|---|---|---|---|---|
| | PSNR | SSIM | Time | PSNR | SSIM | Time |
| Boat | 28.90 | 0.8394 | **2.40** | **29.76** | **0.8624** | 5.50 |
| Lena | 29.68 | 0.8734 | **2.33** | **30.67** | **0.8874** | 5.54 |
| House | 27.85 | 0.8179 | **2.35** | **28.22** | **0.8466** | 4.72 |

(a) Original image (b) Noisy image: 19.17 (c) Ours: 27.55 (d) Mei's: 28.48 (e) Ours($\mu = 0$): 29.50

**Fig. 8.** Comparison between our method and Mei's method. The number under images denotes the PSNR (dB) value of images. (a) the original image, (b) the noisy image, (c) the recovered image by our method, (d) the recovered image by Mei's method, and (e) the recovered image by our method with $\mu = 0$.

**Table 4**
The PSNR and SSIM values of the results by Mei's model and our model with different initial guesses.

| Models | Initial guesses | PSNR | SSIM |
|---|---|---|---|
| Our model | The random image | 27.54 | 0.7773 |
| | The filtered image | 27.55 | 0.7772 |
| | The observed image | 27.55 | 0.7773 |
| Mei's model | The random image | 24.30 | 0.7185 |
| | The filtered image | 27.59 | 0.7850 |
| | The observed image | 28.48 | 0.8104 |
| Our model ($\mu = 0$) | The observed image | 29.50 | 0.8363 |

(a) SURE-LET: 22.92 (b) myriad filter: 26.15 (c) BM3D: 27.25 (d) Ours: 28.93

(e) SURE-LET: 23.08 (f) myriad filter: 28.56 (g) BM3D: 27.88 (h) Ours: 30.67

**Fig. 9.** Recovered images of different methods for removing Cauchy noise with the noise level $\xi = 0.02$. The number under images represents the PSNR (dB) of images. First column: restored images by SURE-LET; second column: restored images by the myriad filter; third column: restored images by BM3D; fourth column: restored images by our method.

removing Cauchy noise. It indicates that our method is stable with different initial guesses, while Mei's method depends on the initial guess.

Finally, we compare our method with other methods in image denoising including the myriad filter [28], the SURE-LET [40], and the BM3D [16]. Here, we test two images Cameraman and Lena degraded by Cauchy noise with the noise level $\xi = 0.02$. From Fig. 9, it is clearly that our method outperforms all of compared methods. Visually, we can see that the background are rough in Fig. 8 (b) and (f), and there is still some noise left in Fig. 8 (a), (c), (e), and (g).

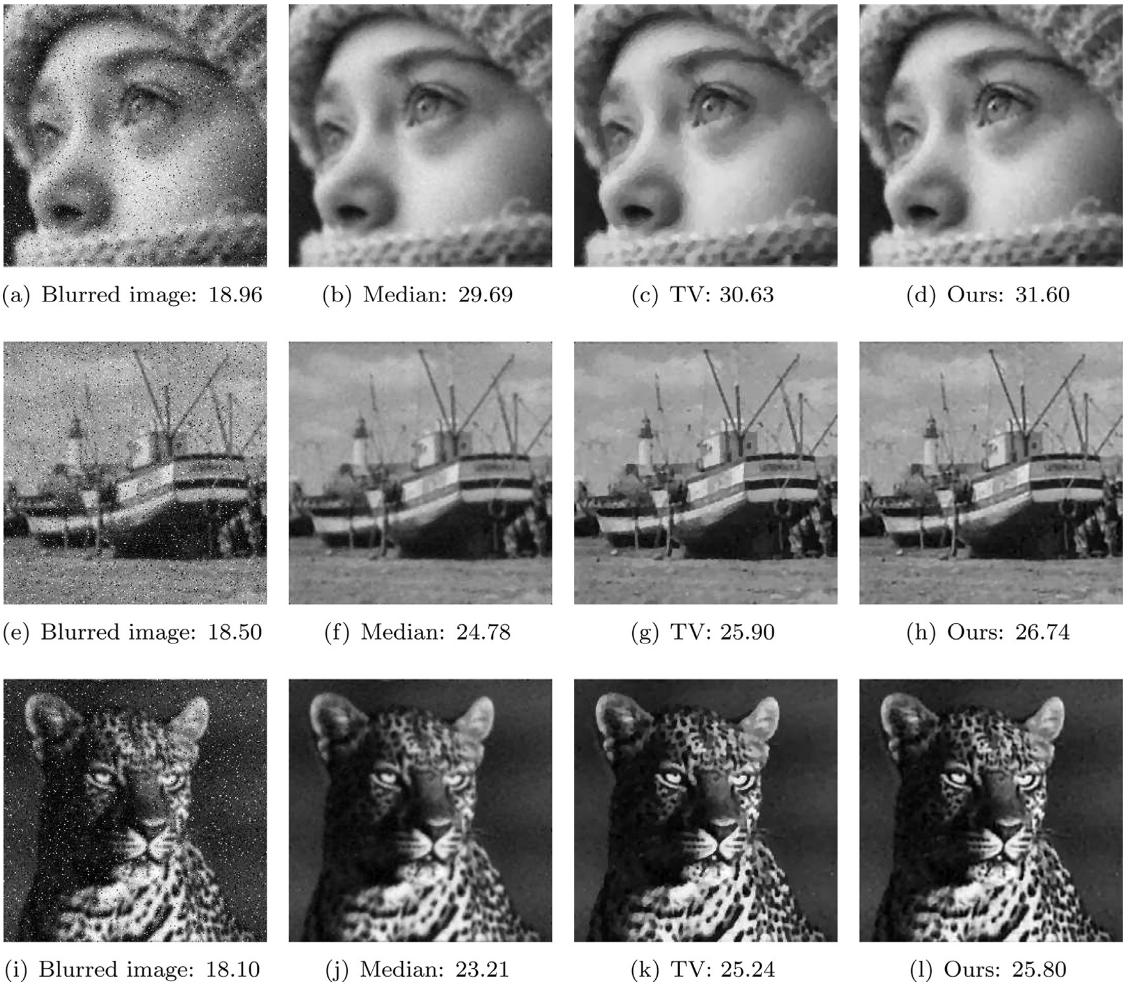



**Fig. 10.** Comparison of restored images from different methods for deblurring and denoising the images degraded by a Gaussian blur and corrupted by Cauchy noise. The number under the images represents the PSNR (dB) of images. First column: blurred and noisy images ($\xi = 0.02$); second column: restored images by the median filter; third column: restored images by TV method ($\lambda = 1.3$ (Beauty); 2.2 (Boat); 2.1 (Leopard), $\mu = 6.25$, and $\gamma = \frac{\sqrt{2}}{10}$); fourth column: restored images by our method ($\lambda = 10$ (Beauty); 14 (Boat); 13 (Leopard), $\mu = 6.25$, and $\gamma = \frac{\sqrt{2}}{10}$).

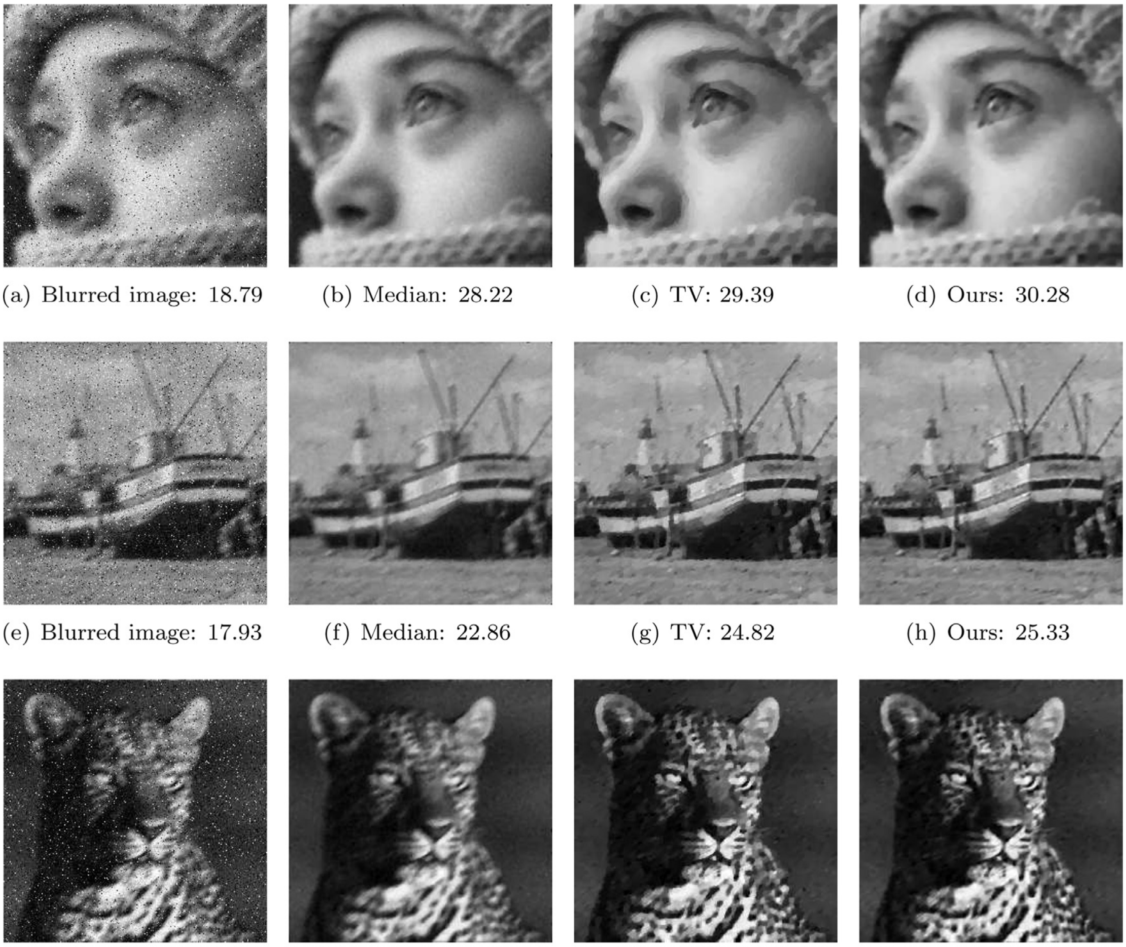
**Table 5**
The PSNR (dB) and SSIM values for images degraded by a Gaussian blur and Cauchy noise($\xi = 0.02$) and restored images by different methods.

| Image | PSNR | | | | SSIM | | | |
|---|---|---|---|---|---|---|---|---|
| | Blurred | Median | TV | Ours | Blurred | Median | TV | Ours |
| Beauty | 18.96 | 29.69 | 30.63 | **31.60** | 0.2656 | 0.9005 | 0.9093 | **0.9248** |
| Baby | 18.85 | 29.05 | 30.88 | **31.40** | 0.2688 | 0.8522 | 0.8805 | **0.8937** |
| Boat | 18.50 | 24.78 | 25.90 | **26.74** | 0.3061 | 0.7199 | 0.7586 | **0.7687** |
| Lena | 18.64 | 26.09 | 27.52 | **28.13** | 0.3185 | 0.7935 | 0.8240 | **0.8297** |
| Leopard | 18.10 | 23.21 | 25.24 | **25.80** | 0.3593 | 0.7245 | 0.7756 | **0.7825** |
| Tulips | 18.34 | 24.25 | 26.01 | **26.30** | 0.3951 | 0.7553 | 0.7899 | **0.8041** |
| House | 18.13 | 23.30 | 24.99 | **25.30** | 0.3086 | 0.6986 | 0.7206 | **0.7510** |

### *4.2. Image deblurring and denoising*

In this subsection, we consider restoring blurred images under Cauchy noise. Here, we consider two blur kernels: the Gaussian blur kernel with size 9 and standard deviation 1; the motion blur kernel with $len = 8$ and $theta = 30$. Then, Cauchy noise with $\xi = 0.02$ is added into the blurred images.

Figs. 10 and 11 show the restored images for deblurring and denoising. Tables 5 and 6 list the values of PSNR and SSIM by applying different methods. From Tables 5 and 6, we can see that the proposed method always obtain the highest PSNR



(a) Blurred image: 18.79 (b) Median: 28.22 (c) TV: 29.39 (d) Ours: 30.28

(e) Blurred image: 17.93 (f) Median: 22.86 (g) TV: 24.82 (h) Ours: 25.33

(i) Blurred image: 16.86 (j) Median: 20.44 (k) TV: 22.59 (l) Ours: 23.09

**Fig. 11.** Comparison of restored images from different methods for deblurring and denoising the images degraded by a motion blur and corrupted by Cauchy noise. The number under images represents the PSNR (dB) of images. First column: blurred and noisy images ($\xi = 0.02$); second column: restored images by the median filter; third column: restored images by TV method ($\lambda = 1.3$ (Beauty); 3.3 (Boat); 3.2 (Leopard), $\mu = 6.25$, and $\gamma = \frac{\sqrt{2}}{10}$); fourth column: restored images by our method ($\lambda = 8$ (Beauty); 17 (Boat); 16 (Leopard), $\mu = 6.25$, and $\gamma = \frac{\sqrt{2}}{10}$).

**Table 6**
The PSNR (dB) and SSIM values for images degraded by a motion blur and Cauchy noise($\xi = 0.02$) and restored images by different methods.

| Image | PSNR | | | | SSIM | | | |
|---|---|---|---|---|---|---|---|---|
| | Blurred | Median | TV | Ours | Blurred | Median | TV | Ours |
| Beauty | 18.79 | 28.22 | 29.39 | **30.28** | 0.2495 | 0.8704 | 0.8889 | **0.9232** |
| Baby | 18.37 | 25.88 | 28.81 | **28.96** | 0.2304 | 0.7866 | 0.8398 | **0.8546** |
| Boat | 17.93 | 22.86 | 24.82 | **25.33** | 0.2360 | 0.6230 | 0.6804 | **0.7051** |
| Lena | 18.03 | 23.41 | 25.86 | **26.09** | 0.2627 | 0.7058 | 0.7747 | **0.7914** |
| Leopard | 16.86 | 20.44 | 22.59 | **23.09** | 0.2604 | 0.6062 | 0.6939 | **0.7163** |
| Tulips | 17.53 | 21.78 | 24.41 | **24.54** | 0.3098 | 0.6239 | 0.7181 | **0.7306** |
| House | 17.32 | 21.21 | 23.18 | **23.77** | 0.2368 | 0.6023 | 0.6797 | **0.6879** |

and SSIM values. From Figs. 10 and 11, the recovered images by the median filter are oversmoothing, since the median filter does not deblur image. The TV method can recover edge and remove noise, but yield the staircase effects. This phenomenon is especially obvious from the zoom-in regions of restored images in Fig. 12, such as the eye of Beauty. Therefore, it is obvious that our method not only preserves the fine features, but also removes Cauchy noise effectively. Furthermore, our method balances well between the edge preserving and staircase effects reduction.

(a) (b) (c) (d)

**Fig. 12.** Zoomed vision of restored images degraded by a Gaussian blur (the first row) and a motion blur (the second row). (a) blurred and noisy images, (b) restored images by the median filter, (c) restored images by TV method, and (d) restored images by our method.

**Table 7**
PSNR (dB), SSIM, and CPU-time (in seconds) for the inner iteration number N on images Cameraman and Parrot corrupted by Cauchy noise with the noise level $\xi = 0.02$.

| Images | $N$ | 1 | 3 | 5 | 10 | 20 | 50 | 100 | 200 |
|---|---|---|---|---|---|---|---|---|---|
| Cameraman | PSNR | 28.76 | 28.91 | 28.92 | 28.93 | 28.93 | 28.93 | 28.93 | 28.93 |
| | SSIM | 0.8400 | 0.8738 | 0.8770 | 0.8781 | 0.8781 | 0.8781 | 0.8781 | 0.8781 |
| | Time | 2.57 | 3.25 | 3.91 | 5.50 | 8.76 | 18.62 | 34.90 | 67.62 |
| Parrot | PSNR | 29.61 | 29.84 | 29.85 | 29.85 | 29.85 | 29.85 | 29.85 | 29.85 |
| | SSIM | 0.8590 | 0.8858 | 0.8878 | 0.8883 | 0.8883 | 0.8883 | 0.8883 | 0.8883 |
| | Time | 2.76 | 3.34 | 3.97 | 5.54 | 8.85 | 18.95 | 35.26 | 68.16 |

## 5. Discussion

### 5.1. Experimental results analysis

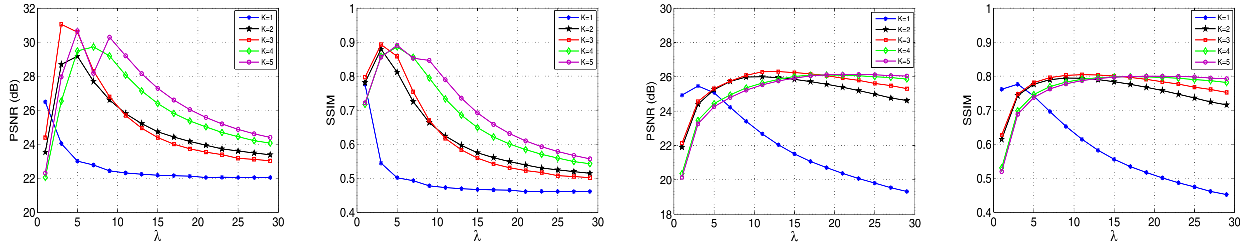
This paper proposes a new convex optimization model for restoring the blurred images corrupted by Cauchy noise. Our objective is to reduce staircase effects produced by using the total variation. The experiments on different noise levels and blur kernels show the potential of the proposed method for removing Cauchy noise, and the restored images indicate that our method can alleviate staircase effects efficiently.

From the gradient images illustrated in Fig. 1, it can be observed that the gradient of the image is not only sparse, but group sparse. Inspired by this, we utilize the sparsity and group sparsity of the gradient to reduce staircase effects. Moreover, from the denoised images in Figs. 4 and 6, we find that our method gives the best visual quality. The median filter works quite well if the noise level is low, but it is not able to preserve most details. In addition, although the TV method can remove noise efficiently, it produces staircases effects. The reason why our method performs better is because the gradient of the image is group sparse. Specially, the results in Tables 1, 2, 5, and 6 show that the proposed method achieves the highest PSNR and SSIM values in all experiments than those of the comparing methods.

In summary, our method can apply to various Cauchy noise removal and deblurring problems.

### 5.2. Analysis of the parameters

1) parameters: $N$, $K$, and $\lambda$. We study the settings of the inner iteration $N$ in the MM, the group size $K$ in OGS-TV, and the regularization parameter $\lambda$. The images Cameraman and Parrot are used to study the choice of $N$. They are corrupted by Cauchy noise with the noise level $\xi = 0.02$. In order to check how the inner iteration number $N$ impacts the performance of the proposed method, we fix the group size $K = 3$ for the experiments. We set $\lambda = 4$ for images Cameraman and Parrot. The results are shown in Table 7. From these tables, we conclude that when the iteration number $N \geq 10$, the values of PSNR and SSIM tend to be stable, while costing more time. Therefore, we set $N = 10$ in our work.




(a) (b) (c) (d)

**Fig. 13.** The PSNR and SSIM values on images Peppers and Tulips, as related to the group size K and the regularization parameter λ. (a) and (b): the PSNR and SSIM values of the testing image Peppers with respect to K and λ, respectively; (c) and (d): the PSNR and SSIM values of the testing image Tulips with respect to K and λ, respectively.

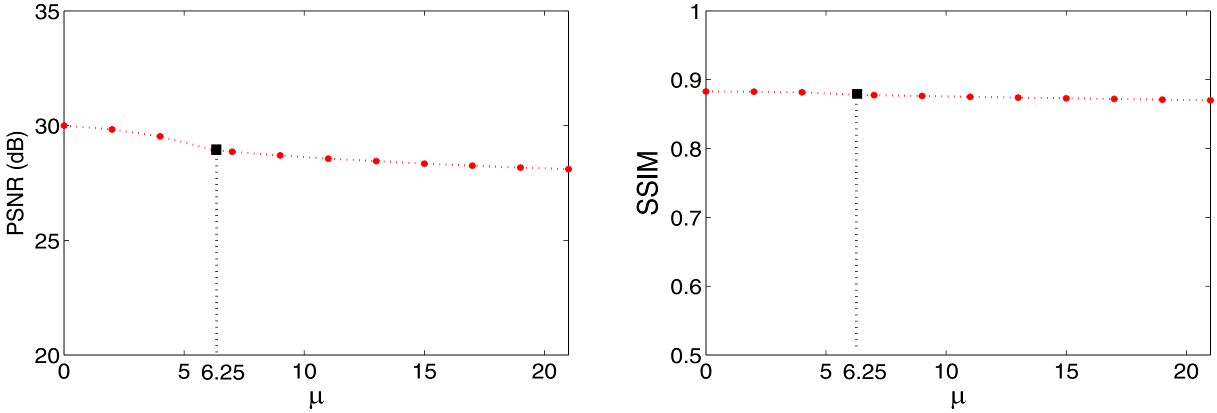



**Fig. 14.** The PSNR and SSIM values with respect to the parameter $\mu$.

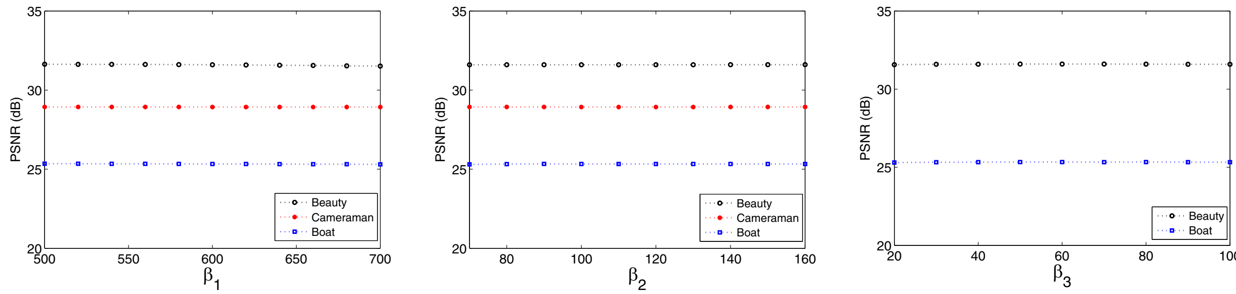



**Fig. 15.** The PSNR values with respect to penalty parameters $\beta_1$, $\beta_2$, and $\beta_3$, respectively.

Then, we fix $N = 10$ and do more experiments to find a good group size $K$. We carry out experiments to compute the PSNR and SSIM values with respect to $K$ and λ. We test two images: Peppers degraded by Cauchy noise with the noise level $\xi = 0.02$; Tulips blurred by the Gaussian blur kernel with size 9 and standard deviation 1 and corrupted Cauchy noise with the noise level $\xi = 0.02$. This test is also used to examine the effect of group size $K$ and regularization parameter λ. The results are shown in Fig. 13. From Fig. 13, it can be seen that the maximum PSNR and SSIM values are obtained when the group size $K = 3$. Then we empirically choose $K = 3$ in the experiments. It can also be observed that the performance of the proposed method achieves the best around $\lambda = 3$ on Peppers and $\lambda = 13$ on Tulips. Since our experiments involve different noise levels and blur kernels, we empirically set the parameter $\lambda \in [2, 20]$.

2) parameters: $\mu$, $\beta_1$, $\beta_2$, and $\beta_3$. In Fig. 14, we plot the PSNR and SSIM values for our algorithm against different values of $\mu$. The test image is Cameraman corrupted by Cauchy noise with the noise level $\xi = 0.02$ (with $\lambda = 3$). The point at $\mu = 6.25$ marked by the solid black point in Fig. 14 refers to the PSNR value 28.93 dB and the SSIM value 0.8781. We can see that when $\mu \geq 6.25$, i.e., our model is convex, the PSNR and SSIM curves tend to be stable. It indicates that our method is robust with respect to $\mu$. Therefore, we choose $\mu$ such that the convexity condition is just satisfied, i.e., $8\mu\gamma^2 = 1$.

In Fig. 15, we plot the PSNR values for our algorithm against different values of parameters $\beta_1$, $\beta_2$, and $\beta_3$. We test three images: Cameraman corrupted by Cauchy noise with the noise level $\xi = 0.02$ (with $\lambda = 3$); Beauty blurred by the

Gaussian blur kernel with size 9 and standard deviation 1 and corrupted Cauchy noise with the noise level $\xi = 0.02$ (with $\lambda = 10$); Boat blurred by the motion blur kernel with $len = 8$ and $theta = 30$ and corrupted Cauchy noise with the noise level $\xi = 0.02$ (with $\lambda = 17$). We remark that, in denoising case, $H$ is the identity operator, the degraded $f$ is only corrupted by Cauchy noise, so there are only three parameters $\lambda$, $\beta_1$, and $\beta_2$ in (23). Thus, in the third column of Fig. 15, there are two PSNR curves with respect to the parameter $\beta_3$. Fig. 15 shows that the PSNR curves of our method are very flat and our method is stable for a wide range of $\beta_1$, $\beta_2$, and $\beta_3$. That is to say, our method is robust with respect to $\beta_1$, $\beta_2$, and $\beta_3$. Therefore, we empirically set $\beta_1 = 600$, $\beta_2 = 100$, and $\beta_3 = 50$ for all our experiments.

## 6. Conclusion

In this paper, we propose a new convex variational model based on the overlapping group sparsity total variation regularizer for restoring the blurred images corrupted by Cauchy noise. We employ the ADMM to solve the proposed convex model. In addition, we discuss the uniqueness of the solution of our model and the convergence analysis of the proposed algorithm. Numerical experiments show that our proposed method outperforms other competitive methods in terms of PSNR and SSIM values. The comparison between the reconstructed images shows that the proposed method can reduce the staircase effects efficiently while preserving sharp edges.

For TV method and our method, the level of Cauchy noise and the blur kernel have to be known, but it may be different in real degraded images. So in future work, we plan to improve the OGS-TV-based method by estimating the noise level [37] and the blur kernel.

## Acknowledgments

The authors would like to thank the anonymous referees and editor for their valuable remarks, questions, and comments that enabled the authors to improve this paper. This research is supported by the National Science Foundation of China (61772003, 61876203), Science Strength Promotion Programme of UESTC, the Fundamental Research Funds for the Central Universities (ZYGX2016J132), Science Strength Promotion Programme of UESTC.

## References


[1] M.V. Afonso, J.M. Bioucas-Dias, M.A.T. Figueiredo, An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems, IEEE Trans. Image Process. 20 (3) (2011) 681–695.
[2] M. Benning, C. Brune, M. Burger, J. Müller, Higher-order TV methods enhancement via Bregman iteration, J. Sci. Comput. 54 (2) (2013) 269–310.
[3] D.P. Bertsekas, A. Nedic, A.E. Ozdaglar, Convex analysis and Optimization, Athena Scientific, 2003.
[4] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. 3 (1) (2011) 1–122.
[5] M. Burger, L. He, C.B. Schönlieb, Cahn-Hilliard inpainting and a generalization for grayvalue images, SIAM J. Imaging Sci. 2 (4) (2009) 1129–1167.
[6] J.-F. Cai, S. Osher, Z. Shen, Split bregman methods and frame based image restoration, Multiscale Model. Simul. 8 (2) (2010) 337–369.
[7] A. Chambolle, P.L. Lions, Image recovery via total variation minimization and related problems, Numer. Math. 76 (2) (1997) 167–188.
[8] A. Chambolle, T. Pock, A first-order primal-dual algorithm for convex problems with applications to imaging, J. Math. Imag. Vis. 40 (1) (2011) 120–145.
[9] R.H. Chan, Y. 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.
[10] R.H. Chan, M. Tao, X. Yuan, Constrained total variation deblurring models and fast algorithms based on alternating direction method of multipliers, SIAM J. Imag. Sci. 6 (1) (2013) 680–697.
[11] T. Chan, A. Marquina, P. Mulet, High-order total variation-based image restoration, SIAM J. Sci. Comput. 22 (2) (2000) 503–516.
[12] 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.
[13] Y. Chang, S. Kadaba, P. Doerschuk, S. Gelfand, Image restoration using recursive Markov random field models driven by Cauchy distributed noise, IEEE Signal Process. Lett. 8 (3) (2001) 65–66.
[14] D.-L. Chen, Y.-Q. Chen, D.-Y. Xue, Fractional-order total variation image denoising based on proximity algorithm, Appl. Math. Comput. 257 (2015) 537–545.
[15] P.L. Combettes, J.C. Pesquet, Proximal splitting methods in signal processing, Heinz H Bauschke 49 (2009) 185–212.
[16] 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.
[17] L.-J. Deng, W.-H. Guo, T.-Z. Huang, Single-image super-resolution via an iterative reproducing kernel hilbert space method, IEEE Trans. Circ. Syst. Video Technol. 26 (11) (2016) 2001–2014.
[18] D.C. Dobson, F. Santosa, Recovery of blocky images from noisy and blurred data, SIAM J. Appl. Math. 56 (4) (1996) 1181–1198.
[19] 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.
[20] J. Eckstein, D.P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math. Prog. 55 (1) (1992) 293–318.
[21] M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, IEEE Trans. Image Process. 15 (12) (2006) 3736–3745.
[22] P. Escande, P. Weiss, W.-X. Zhang, A variational model for multiplicative structured noise removal, J. Math. Imag. Vis. 57 (1) (2017) 43–55.
[23] E. Esser, X.-Q. Zhang, T.F. Chan, A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science, SIAM J. Imag. Sci. 3 (4) (2010) 1015–1046.
[24] R. Fergus, B. Singh, A. Hertzmann, S.T. Roweis, W.T. Freeman, Removing camera shake from a single photograph, ACM Trans. Graph. 25 (3) (2006) 787–794.
[25] M. Figueiredo, J.M. Bioucas-Dias, R.D. Nowak, Majorization-minimization algorithms for wavelet-based image restoration, IEEE Trans. Image Process. 16 (12) (2007) 2980–2991.
[26] T. Goldstein, B. O'Donoghue, S. Setzer, R. Baraniuk, Fast alternating direction optimization methods, SIAM J. Imag. Sci. 7 (3) (2014) 1588–1623.
[27] T. Goldstein, S. Osher, The split Bregman method for L1-regularized problems, SIAM J. Imag. Sci. 2 (2) (2009) 323–343.
[28] J.G. Gonzalez, G.R. Arce, Optimality of the myriad filter in practical impulsive-noise environments, IEEE Trans. Signal Process. 49 (2) (2001) 438–441.

[29] B.-S. He, H. Yang, Some convergence properties of a method of multipliers for linearly constrained monotone variational inequalities, Oper. Res. Lett. 23 (3) (1998) 151–161.
[30] B.-S. He, X. Yuan, On the O(1/n) convergence rate of the Douglas-Rachford alternating direction method, SIAM J. Numer. Anal. 50 (2) (2012) 700–709.
[31] M. Idan, J.L. Speyer, Cauchy estimation for linear scalar systems, IEEE Trans. Autom. Control 55 (6) (2010) 1329–1342.
[32] T.-Y. Ji, T.-Z. Huang, X.-L. Zhao, T.-H. Ma, L.-J. Deng, A non-convex tensor rank approximation for tensor completion, Appl. Math. Modell. 48 (2017) 410–422.
[33] T.-Y. Ji, T.-Z. Huang, X.-L. Zhao, T.-H. Ma, G. Liu, Tensor completion using total variation and low-rank matrix factorization, Inf. Sci. 326 (2016) 243–257.
[34] T.-X. Jiang, T.-Z. Huang, X.-L. Zhao, T.-Y. Ji, L.-J. Deng, Matrix factorization for low-rank tensor completion using framelet prior, Inf. Sci. 436 (2018) 403–417.
[35] D. Krishnan, R. Fergus, Fast image deconvolution using hyper-Laplacian priors, in: Neural Information Processing Systems (2009).
[36] E.E. Kuruoglu, W.J. Fitzgerald, P.J.W. Rayner, Near optimal detection of signals in impulsive noise modeled with a symmetric $\alpha$-stable distribution, IEEE Commun. Lett. 2 (10) (1998) 282–284.
[37] F. Laus, F. Pierre, G. Steidl, Nonlocal myriad filters for cauchy noise removal, J. Math. Imaging Vis. 60 (8) (2018) 1324–1354.
[38] F. Li, X.-G. Lv, A Decoupled method for image inpainting with patch-based low rank regulariztion, Appl. Math. Comput. 314 (2017) 334–348.
[39] J. Liu, T.-Z. Huang, I.W. Selesnick, X.-G. Lv, P.Y. Chen, Image restoration using total variation with overlapping group sparsity, Inf. Sci. 295 (2015) 232–246.
[40] F. Luisier, T. Blu, M. Unser, A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding, IEEE Trans. Image Process. 16 (3) (2007) 593–606.
[41] M. Lysaker, A. Lundervold, X.-C. Tai, Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time, IEEE Trans. Image Process. 12 (12) (2003) 1579–1590.
[42] M. Lysaker, X.-C. Tai, Iterative image restoration combining total variation minimization and a second-order functional, Int. J. Comput. Vis. 66 (1) (2006) 5–18.
[43] T.-H. Ma, T.-Z. Huang, X.-L. Zhao, Y.-F. Lou, Image deblurring with an inaccurate blur kernel using a group-based low-rank image prior, Inf. Sci. 408 (2017) 213–233.
[44] 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.
[45] M. Nikolova, A variational approach to remove outliers and impulse noise, IEEE Trans. Image Process. 20 (1) (2004) 99–120.
[46] J.P. Nolan, Stable Distributions-Models for Heavy Tailed Data, Birkhäuser Boston, Cambridge http://fs2.american.edu/jpnolan/www/stable/stable.html.
[47] J. Oliveira, J. Bioucas-Dias, M. Figueiredo, Adaptive total variation image deblurring: a majorization minimization approach, Signal Process. 89 (9) (2009) 1683–1693.
[48] T. Pander, New polynomial approach to myriad filter computation, Signal Process. 90 (6) (2010) 1991–2001.
[49] Y. Peng, J. Chen, X. Xu, F. Pu, SAR images statistical modeling and classification based on the mixture of $\alpha$-stable distributions, Remote Sens. 5 (5) (2013) 2145–2163.
[50] J. Portilla, V. Strela, M.J. Wainwright, E.P. Simoncelli, Image denoising using scale mixtures of Gaussians in the wavelet domain, IEEE Trans. Image Process. 12 (11) (2003) 1338–1351.
[51] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Phys. D Nonlinear Phenom. 60 (1–4) (1992) 259–268.
[52] F. Sciacchitano, Y.-Q. Dong, T.-Y. Zeng, Variational approach for restoring blurred images with Cauchy noise, SIAM J. Imag. Sci. 8 (3) (2015) 1894–1922.
[53] I.W. Selesnick, P.Y. Chen, Total variation denoising with overlapping group sparsity, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2013, pp. 5696–5700.
[54] G. Steidl, A note on the dual treatment of higher-order regularization functionals, Computing 76 (1) (2006) 135–148.
[55] M.F. Tappen, B.C. Russell, W.T. Freeman, Exploiting the sparse derivative prior for super-resolution and image demosaicing, in: Proceedings of the IEEE Workshop on Statistical and Computational Theories of Vision, 2003.
[56] T. Wan, N. Canagarajah, A. Achim, Segmentation of noisy colour images using Cauchy distribution in the complex wavelet domain, IET Image Process. 5 (2) (2011) 159–170.
[57] S. Wang, W.-H. Guo, T.-Z. Huang, G. Raskutti, Image inpainting using reproducing kernel Hilbert space and Heaviside functions, J. Comput. Appl. Math. 311 (2017) 551–564.
[58] S. Wang, T.-Z. Huang, X.-L. Zhao, J.-J. Mei, J. Huang, Speckle noise removal in ultrasound images by first- and second-order total variation, Numer. Algorithms 78 (2) (2018) 513–533.
[59] 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.
[60] 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.
[61] 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.
[62] K. Zhang, W.-M. Zuo, Y.-J. Chen, D.-Y. Meng, L. Zhang, Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising, IEEE Trans. Image Process. 26 (2017) 3142–3155.
[63] 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.
[64] 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) B1304–B1320.