

Research Article

Image Variational Denoising Using Gradient Fidelity on Curvelet Shrinkage

Liang Xiao,^{1,2} Li-Li Huang,^{1,3} and Badrinath Roysam²

¹ School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China

² Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180-3590, USA

³ Department of Information and Computing Science, Guangxi University of Technology, Liuzhou 545000, China

Correspondence should be addressed to Liang Xiao, xiaoliang@mail.njust.edu.cn

Received 27 December 2009; Revised 20 March 2010; Accepted 7 June 2010

Academic Editor: Ling Shao

Copyright © 2010 Liang Xiao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A new variational image model is presented for image restoration using a combination of the curvelet shrinkage method and the total variation (TV) functional. In order to suppress the staircasing effect and curvelet-like artifacts, we use the multiscale curvelet shrinkage to compute an initial estimated image, and then we propose a new gradient fidelity term, which is designed to force the gradients of desired image to be close to the curvelet approximation gradients. Then, we introduce the Euler-Lagrange equation and make an investigation on the mathematical properties. To improve the ability of preserving the details of edges and texture, the spatial-varying parameters are adaptively estimated in the iterative process of the gradient descent flow algorithm. Numerical experiments demonstrate that our proposed method has good performance in alleviating both the staircasing effect and curvelet-like artifacts, while preserving fine details.

1. Introduction

Image denoising is a very important preprocessing in many computer vision tasks. The tools for attacking this problem come from computational harmonic analysis (CHA), variational approaches, and partial differential equations (PDEs) [1]. The major concern in these image denoising models is to preserve important image features, such as edges and texture, while removing noise.

In the direction of multi-scale geometrical analysis (MGA), the shrinkage algorithms based on the CHA tools, such as contourlets [2] and curvelets [3–5], are very important in image denoising because they are simple and have efficient computational complexity and promising properties for singularity analysis. Therefore, the pseudo-Gibbs artifacts caused by the shrinkage methods based on Fourier transform and wavelets attempt to be overcome by the methods based on MGA at least partially. However, there are still some curve-like artifacts in MGA-based shrinkage methods [6].

Algorithms designed by variational and PDE models are free from the above lacks of MGA but cost heavy

computational burden that is not suitable for time critical application. In addition, the PDE-based algorithms tend to produce a staircasing effect [7], although they can achieve a good trade-off between noise removal and edge preservation. For instance, the total variation (TV) minimizing [8] method has some undesirable drawbacks such as the staircasing effect and loss of texture, although it can reduce pseudo-Gibbs oscillations effectively. Similar problem can be found in many other nonlinear diffusion models such as the Perona-Malik model [9] and the mean curvature motion model [10]. In this paper, we will focus on the hybrid variational denoising method. Specifically, we will emphasize on the improvement on the TV model, and we will propose a novel gradient fidelity term based on the curvelet shrinkage algorithm.

1.1. Related Works and Analysis. To begin with, we will review some related works of the variational methods. To cope with the ill-posed nature of denoising, the variational methods often use regularization technique. Let u_0 denote the observed raw image data and u is the original good image;

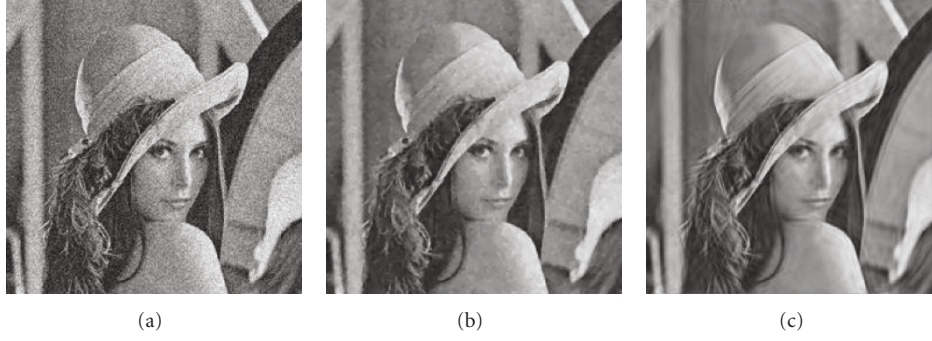


FIGURE 1: (a) The noisy “Lena” image with standard deviation 35; (b) the denoised “Lena” image by the TV algorithm; (c) the denoised “Lena” image by the curvelet hard shrinkage algorithm.

then the regularization functional-based denoising is given by

$$\hat{u} = \operatorname{argmin}_u \{\lambda E_{\text{data}}(u, u_0) + E_{\text{smooth}}(u)\}, \quad (1)$$

where the first term $E_{\text{data}}(u, u_0)$ is the image fidelity term, which penalizes the inconsistency between the under-estimated recovery image and the acquired noisy image, while the second term $E_{\text{smooth}}(u)$ is the regularization term which imposes some priori constraints on the original image and to a great degree determines the quality of the recovery image, and λ is the regularization parameter which balances trade-off between the image fidelity term $E_{\text{data}}(u, u_0)$ and the regularization term $E_{\text{smooth}}(u)$.

A classical model is the minimizing total variational (TV) functional [8]. The TV model seeks the minimal energy of an energy functional comprised of the TV norm of image u and the fidelity of this image to the noisy image u_0 :

$$\hat{u} = \operatorname{argmin}_{u \in \text{BV}(\Omega)} E(u) = \int_{\Omega} \left\{ |\nabla u| + \frac{1}{2} \lambda (u - u_0)^2 \right\} dx dy. \quad (2)$$

Here, Ω denotes the image domain and $\text{BV}(\Omega)$ is the space of functions of $L^1(\Omega)$ such that the TV norm is $\text{TV}(u) = \int_{\Omega} |\nabla u| dx dy < \infty$. The gradient descent evolution equation is

$$\frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda(u_0 - u). \quad (3)$$

In this formulation, λ can be considered as a Lagrange multiplier, computed by

$$\lambda = \frac{1}{|\Omega| \sigma^2} \int_{\Omega} \operatorname{div} \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon^2}} \right) (u - u_0) dx. \quad (4)$$

Although the TV model can reduce oscillations and regularize the geometry of level sets without penalizing discontinuities, it possesses some properties which may be undesirable under some circumstances [11], such as staircasing and loss of texture (see Figures 1(a)–1(c)).

Currently, there are three approaches which can partially overcome these drawbacks. One approach to preventing staircasing is to introduce higher-order derivatives into the energy. In [12], an image u is decomposed into two parts: $u = u_1 + u_2$. The u_1 component is measured using the total variation norm, while the second component u_2 is measured using a higher-order norm. More precisely, one solves the following variational problem that now involves two unknowns:

$$\begin{aligned} \min_{u_1, u_2} E(u_1, u_2) \\ = \int_{\Omega} \left\{ |\nabla u_1| + \alpha H(\nabla u_2) + \lambda(u_1 + u_2 - u_0)^2 \right\} dx dy. \end{aligned} \quad (5)$$

Here $H(\nabla u_2)$ could be some higher-order norm, for example, $H(\nabla u_2) = |\nabla^2 u_2|$. More complex higher-order norms were brought to variational method in order to alleviate the staircasing effect [13].

The second approach overcoming the staircasing effect is to adopt the new data fidelity term. Gilboa et al. proposed an adaptive fidelity term to better preserve fine-scale features [14]. Zhu and Xia in [7] introduced the gradient fidelity term defined by

$$\begin{aligned} E(u, u_0) = \int_{\Omega} \alpha (u - u_0)^2 dx dy \\ + \int_{\Omega} \beta \|\nabla u - \nabla(G_{\sigma} \otimes u_0)\|^2 dx dy, \end{aligned} \quad (6)$$

where G_{σ} is the Gaussian kernel with scale σ , and the symbol “ \otimes ” denotes the convolution operator. Their studies show that this gradient fidelity term can alleviate the staircasing effect. However, classical Gaussian filtering technique is the uniform smoothing in all direction of images and fine details are easily destroyed with these filters. Hence, the gradient of smoothed image is unreliable near the edges, and the gradient fidelity cannot preserve the gradient and thereby image edges.

The third approach is the combination of the variational model and the MGA tools. In [14], the TV model has been combined with wavelet to reduce the pseudo-Gibbs artifacts resulted from wavelet shrinkage [15]. Nonlinear diffusion

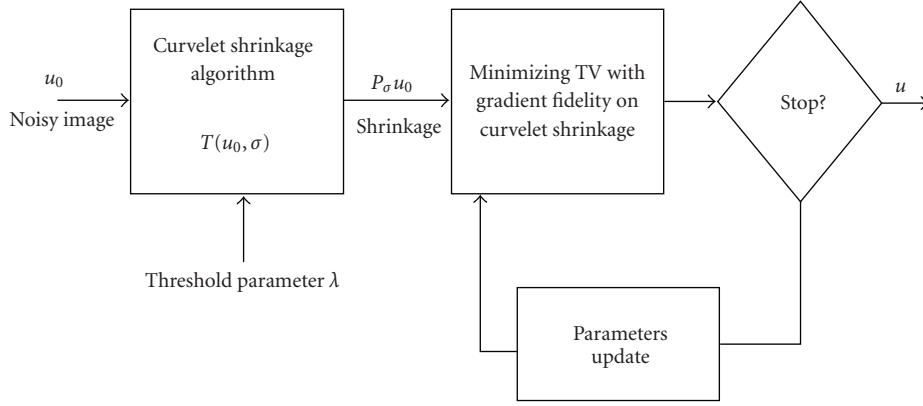


FIGURE 2: Illustrate the proposed self-optimizing image denoising approach.

has been combined with wavelet shrinkage to improve the rotation invariance [16]. The author in [17] presented a hybrid denoising methods in which the complex ridgelet shrinkage was combined with total variation minimization [6]. From their reports, the combination of MGA and PDE methods can improve the visual quality of the restored image and provides a good way to take full advantages of both methods.

1.2. Main Contribution and Paper's Organization. In this paper, we add a new gradient fidelity term to the TV model to some second-order nonlinear diffusion PDEs for avoiding the staircasing effect and curvelet-like artifacts. This new gradient fidelity term provides a good mechanism to combine curvelet shrinkage algorithm and the TV regularization.

This paper is organized as follows. In Section 2, we introduce the curvelet transform. In Section 3, we propose a new hybrid model for image smoothing. In this model, we have two main contributions. We propose a new hybrid fidelity term, in which the gradient of multi-scale curvelet shrinkage image is used as a feature fidelity term in order to suppress the staircasing effect and curvelet-like artifacts. Secondly, we propose an adaptive gradient descent flow algorithm, in which the spatial-varying parameters are adaptively estimated to improve the ability of preserving the details of edges and texture of the desired image. In Section 4, we give numerical experiments and analysis.

The pipeline of our proposed method is illustrated in Figure 2. There are three core modules in our method. In the first module, we apply the curvelet shrinkage algorithm to obtain a good initial restored image $P_\sigma u_0$. The second module is the minimizing TV with the new gradient fidelity, and this module is a global optimizing process which is guided by our proposed general objective functional. The third one is the parameters adjustment module, which provides an adaptive process to compute the value for the system's parameters. The rationale behind the proposed method is that high visual quality image restoration scheme is expected to be a blind process to filter out noise, preserve the edge, and alleviate other artifacts.

2. Curvelet Transform

In the next, we review the basic principles of curvelets which were originally proposed by Candès et al. in [5]. Let $W(r)$ ($r \in (1/2, 2)$) and $V(t)$ ($t \in (-1, 1)$) be a pair of smooth, non-negative real-valued functions; here $W(r)$ is called “radial window” and $V(t)$ is called “angular window”. Both of them need to satisfy the admissibly conditions:

$$\begin{aligned} \sum_{j=-\infty}^{+\infty} W^2(2^j r) &= 1, \quad r \in \left(\frac{3}{4}, \frac{3}{2}\right), \\ \sum_{l=-\infty}^{+\infty} V^2(t-l) &= 1, \quad t \in \left(-\frac{1}{2}, 1\right). \end{aligned} \quad (7)$$

Now, for each $j \geq j_0$, let the window U_j in Fourier domain be given by

$$U_j(r, \theta) = 2^{-3j/4} W(2^{-j} r) V\left(\frac{2^{\lfloor j/2 \rfloor} \theta}{(2\pi)}\right), \quad (8)$$

where $\lfloor j/2 \rfloor$ is the integer part of $j/2$ and (r, θ) denotes the polar coordinate; thus the support of U_j is a polar “wedge” which is determined by “radial window” and “angular window”. Let U_j be the Fourier transform of $\varphi_j(x)$, that is, $\hat{\varphi}_j(x) = U_j(\omega)$. We may think of $\varphi_j(x)$ as a “mother” curvelet in the sense that all curvelets at scale 2^{-j} are obtained by rotation and translations of $\varphi_j(x)$. Let R_θ be the rotation matrix by θ radians and R_θ^{-1} its inverse, then curvelets are indexed by three parameters: a scale 2^{-j} ($j \geq j_0$), an equispaced sequence of orientation $\theta_{j,l} = 2\pi \cdot 2^{-\lfloor j/2 \rfloor} \cdot l$ ($l = 1, 2, \dots, 0 \leq \theta_l \leq 2\pi$), and the position $x_k^{(j,l)} = R_{\theta_{j,l}}^{-1}(k_1 2^{-j}, k_2 2^{-j})$ ($k = (k_1, k_2) \in \mathbb{Z}^2$). With these parameters, the curvelets are defined by

$$\varphi_{j,l,k}(x) = \varphi_j\left(R_{\theta_{j,l}}\left(x - x_k^{(j,l)}\right)\right). \quad (9)$$

A curvelet coefficient is then simply the inner product between an element $u \in L^2(\mathbb{R}^2)$ and a curvelet $\varphi_{j,l,k}$, that is,

$$c_{j,l,k} = \langle u(x), \varphi_{j,l,k} \rangle = \int_{\mathbb{R}^2} u(x) \overline{\varphi_{j,l,k}(x)} dx. \quad (10)$$

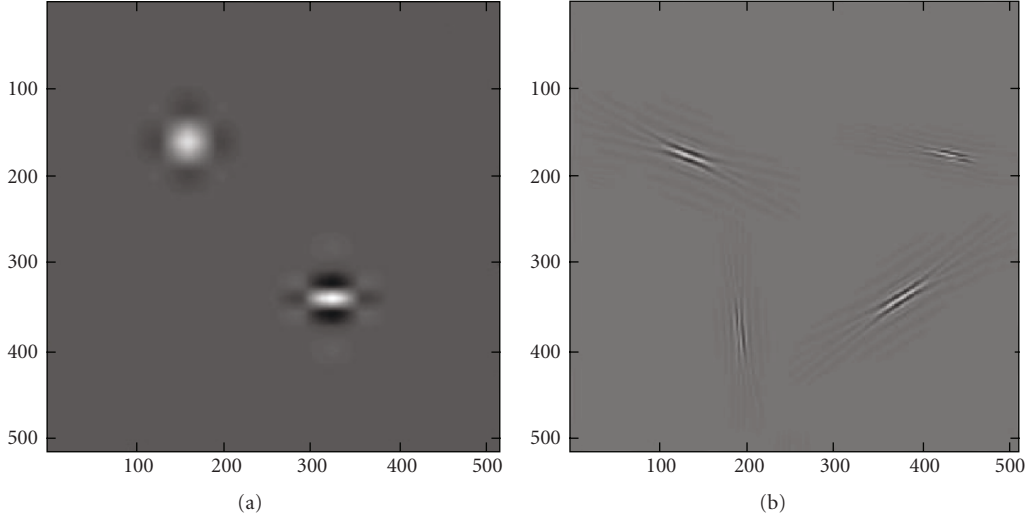


FIGURE 3: The elements of wavelets (a) and curvelets on various scales, directions and translations in the spatial domain (b).

Let $\mu = (j, l, k)$ be the collection of the triple index. The family of curvelet functions forms a tight frame of $L^2(R^2)$. That means that each function $u \in L^2(R^2)$ has a representation:

$$u(x) = \sum_{\mu} \langle u(x), \varphi_{\mu} \rangle \varphi_{\mu}, \quad (11)$$

where $\langle u, \varphi_{\mu} \rangle$ denotes the L^2 -scalar product of u and φ_{μ} . The coefficients $c_{\mu}(u) = \langle u(x), \varphi_{\mu} \rangle$ are called coefficients of function u . In this paper, we apply the second-generation curvelet transform, and the digital implementations can be outlined roughly as three steps [5]: apply 2D FFT, product with frequency windows, and apply 2D inverse FFT for each window. The forward and inverse curvelet transforms have the same computational cost of $O(N^2 \log N)$ for an $N \times N$ data [11]. More details on curvelets and recent applications can be found in recent reviewer papers [3–6, 18, 19]. Figure 3 shows the elements of curvelets in comparison with wavelets. Note that the tensor-product 2D wavelets are not strictly isotropic but have three directions, while curvelets have almost arbitrary directional selectivity.

3. Combination TV Minimization with Gradient Fidelity on Curvelet Shrinkage

3.1. The Proposed Model. We start from the following assumed additive noise degradation model:

$$u_0 = u + v, \quad (12)$$

where u_0 denotes the observed raw image data, u is the original good image, and v is additive measurement noise. The goal of image denoising is to recover u from the observed image data u_0 . The shrinkage algorithm on some multi-scale frame $\{\varphi_{\mu} : \mu \in \Lambda\}$ can be written as follows:

$$C_{\mu}u_0 = C_{\mu}u + C_{\mu}v, \quad (13)$$

where C_{μ} is the corresponding MGA operator, that is, $C_{\mu}(u) = \langle u, \varphi_{\mu} \rangle$, for all $\mu \in \Lambda$, and “ Λ ” is a set of indices. The rationale is that the noise $C_{\mu}v$ is nearly Gaussian. The principles of the shrinkage estimators which estimate the frame coefficients $\{C_{\mu}u\}$ from the observed coefficients $\{C_{\mu}u_0\}$ have been discussed in different frameworks such as Bayesian and variational regularization [20, 21].

Although traditional wavelets perform well only for representing point singularities, they become computationally inefficient for geometric features with line and surface singularities. To overcome this problem, we choose the curvelet as the tool of shrinkage algorithm. In general, the shrinkage operators are considered to be in the form of a symmetric function $T : R \rightarrow R$; thus the coefficients are estimated by

$$\widetilde{C_{\mu}u} = T(C_{\mu}u_0), \quad \forall \mu \in \Lambda. \quad (14)$$

Let $\{\widetilde{\varphi}_{\mu} : \mu \in \Lambda\}$ denote the dual frame, and then a denoised image $P_{\lambda}u_0$ is generated by the reconstruction algorithm:

$$P_{\lambda}u_0 = \sum_{\mu \in \Lambda} T_{\lambda}(c_{\mu}(u_0)) \widetilde{\varphi}_{\mu}. \quad (15)$$

Following the wavelet shrinkage idea which was proposed by Donoho and Johnstone [22], the curvelet shrinkage operators $T_{\lambda}(\cdot)$ can be taken as a soft thresholding function defined by a fixed threshold λ , that is,

$$T_{\lambda}(x) = \begin{cases} x - \lambda, & x \geq \lambda, \\ 0, & |x| < \lambda, \\ x + \lambda, & x \leq -\lambda. \end{cases} \quad (16)$$

or a hard shrinkage function

$$T_{\lambda}(x) = \begin{cases} x, & |x| \geq \lambda, \\ 0, & |x| < \lambda. \end{cases} \quad (17)$$

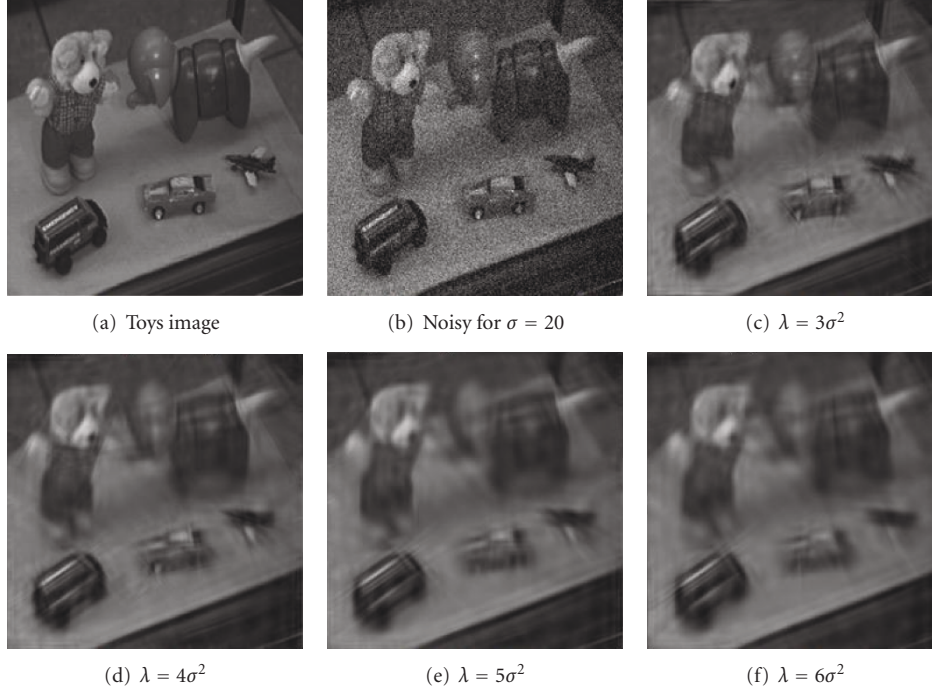


FIGURE 4: (a) Original “Toys” image. (b) Noisy “Toys” image for Gaussian noise with standard deviation $\sigma = 20$. (c)–(f) Denoising of “Toys” image shown in (b) where the curvelet transform is hard-thresholding according to (17) for different choices of λ .

The major problem with wavelet shrinkage methods, as discussed, is that shrinking large coefficients entails an erosion of the spiky image features, while shrinking small coefficients towards zero yields Gibbs-like oscillation in the vicinity of edges and loss of texture. As a new MGA tool, Curvelet shrinkage can suppress this pseudo-Gibbs and preserve the image edge; however, some shapes of curve-like artifacts are generated (see Figure 4).

In order to suppress the staircasing effect and curvelet-like artifacts, we propose a new objective functional:

$$\begin{aligned} E(u) &= \text{TV}(u) + E_{\text{data}}(u, u_0) \\ &= \int_{\Omega} |\nabla u| \, dx dy + \alpha(x, y) \int_{\Omega} (u - u_0)^2 dx dy \quad (18) \\ &\quad + \beta(x, y) \int_{\Omega} |\nabla u - \nabla(P_{\lambda} u_0)|^2 dx dy. \end{aligned}$$

In the cost functional (18), the term $\int_{\Omega} |\nabla u - \nabla(P_{\lambda} u_0)|^2 dx dy$ is called curvelet shrinkage-based gradient data fidelity term and is designed to force the gradient of u to be close to the gradient estimation $\nabla(P_{\lambda} u_0)$ and to alleviate the staircase effect. And the parameters $\alpha(x, y) > 0$ and $\beta(x, y) > 0$ control the weights of each term. For the sake of simplicity for description, we always let $\alpha := \alpha(x, y)$ and $\beta := \beta(x, y)$ in the following sections.

3.2. Basic Properties of Our Model. Let us denote

$$E(u) = \alpha \int_{\Omega} (u - u_0)^2 dx dy + \beta \int_{\Omega} |\nabla u - \nabla(P_{\lambda} u_0)|^2 dx dy. \quad (19)$$

Then, the cost function is a new hybrid data fidelity term, and its corresponding Euler equation is

$$\alpha(u_0 - u) + \beta(\Delta u - \Delta(P_{\lambda} u_0)) = 0. \quad (20)$$

Proposition 1. *The Euler equation (20) equals to produce a new image whose Fourier transform is described as follows: if $\alpha > 0$, $\beta > 0$, then*

$$F(u) = \frac{\alpha F(u_0) + \beta(w^2 + v^2)F(P_{\lambda} u_0)}{\alpha + \beta(w^2 + v^2)}. \quad (21)$$

Proof. Apply Fourier transform to the Euler equation (20) and we will get

$$\alpha(F(u_0) - F(u)) + \beta(F(\Delta u) - F(\Delta(P_{\lambda} u_0))) = 0. \quad (22)$$

According to the differential properties of the Fourier transform

$$\begin{aligned} F(\Delta u)_{(w,v)} &= F\left(\frac{\partial^2}{\partial x^2} u + \frac{\partial^2}{\partial y^2} u\right)_{(w,v)} \\ &= -(w^2 + v^2)F(u)_{(w,v)}, \end{aligned} \quad (23)$$

we have

$$\alpha(F(u_0) - F(u)) + \beta((w^2 + v^2)(F(P_{\lambda} u_0) - F(u))) = 0. \quad (24)$$

If $\alpha > 0$, $\beta > 0$, then we get

$$F(u) = \frac{\alpha F(u_0) + \beta(w^2 + v^2)F(P_{\lambda} u_0)}{\alpha + \beta(w^2 + v^2)}, \quad (25)$$

where w and v are parameters in the frequency domain. \square

Proposition 1 tells us that the Euler equation (20) equals to compute a new image whose Fourier frequency spectrum is the interpolation of $F(u_0)$ and $F(P_\lambda u_0)$. The weight coefficients of $F(u_0)$ and $F(P_\lambda u_0)$ are $\alpha/(\alpha + \beta(w^2 + v^2))$ and $\beta(w^2 + v^2)/(\alpha + \beta(w^2 + v^2))$, respectively.

Proposition 2. *The energy functional $E_{\text{data}}(u, u_0)$ is convex.*

Proof. For all $0 \leq \lambda \leq 1$, for all u_1, u_2 , on one hand we have the following conclusion:

$$(\lambda u_1 + (1 - \lambda)u_2 - u_0)^2 \leq \lambda(u_1 - u_0)^2 + (1 - \lambda)(u_2 - u_0)^2. \quad (26)$$

On the other hand, we have the following conclusion:

$$\begin{aligned} & \|\lambda \nabla u_1 + (1 - \lambda) \nabla u_2 - \nabla(P_\lambda u_0)\|^2 \\ &= \|\lambda(\nabla u_1 - \nabla(P_\lambda u_0)) + (1 - \lambda)(\nabla u_2 - \nabla(P_\lambda u_0))\|^2 \\ &= \lambda^2 \|\nabla u_1 - \nabla(P_\lambda u_0)\|^2 + (1 - \lambda)^2 \|\nabla u_2 - \nabla(P_\lambda u_0)\|^2 \\ &\quad + 2\lambda(1 - \lambda) \langle \nabla u_1 - \nabla(P_\lambda u_0), \nabla u_2 - \nabla(P_\lambda u_0) \rangle \\ &\leq \lambda^2 \|\nabla u_1 - \nabla(P_\lambda u_0)\|^2 + (1 - \lambda)^2 \|\nabla u_2 - \nabla(P_\lambda u_0)\|^2 \\ &\quad + 2\lambda(1 - \lambda) \|\nabla u_1 - \nabla(P_\lambda u_0)\| \cdot \|\nabla u_2 - \nabla(P_\lambda u_0)\| \\ &= [\lambda \|\nabla u_1 - \nabla(P_\lambda u_0)\| + (1 - \lambda) \|\nabla u_2 - \nabla(P_\lambda u_0)\|]^2. \end{aligned} \quad (27)$$

Then, we have

$$\begin{aligned} & \|\lambda \nabla u_1 + (1 - \lambda) \nabla u_2 - \nabla(P_\lambda u_0)\| \\ &\leq \lambda \|\nabla u_1 - \nabla(P_\lambda u_0)\| + (1 - \lambda) \|\nabla u_2 - \nabla(P_\lambda u_0)\|. \end{aligned} \quad (28)$$

According to (26) and (28), we get the $E_{\text{data}}(\lambda u_1 + (1 - \lambda)u_2, u_0) \leq \lambda E_{\text{data}}(u_1, u_0) + (1 - \lambda)E_{\text{data}}(u_2, u_0)$. \square

From Proposition 2, the convexity of the energy functional can guarantee the global optimizing and the existence of the unique solution, while Proposition 1 shows us that the solution has some special form in Fourier domain. Combining Propositions 1 and 2 together, we can remark that the unique solution of (19) is

$$\tilde{u} = F^{-1} \left(\frac{\alpha F(u_0) + \beta(w^2 + v^2)F(P_\lambda u_0)}{\alpha + \beta(w^2 + v^2)} \right). \quad (29)$$

Then, we can prove the following existence and uniqueness theorem.

Theorem 1. *Let $u_0 \in L^\infty(\Omega)$ be a positive, bounded function with $\inf_\Omega u_0 > 0$; then the minimizing problem of energy functional $E(u) = TV(u) + E_{\text{data}}(u, u_0)$ in (18) admits a unique solution $u \in BV(\Omega)$ satisfying*

$$\inf_\Omega (u_0) \leq u \leq \sup_\Omega (u_0). \quad (30)$$

Proof. Using the lower semicontinuity and compactness of $BV(\Omega)$ and the convexity of $E_{\text{data}}(u, u_0)$, the proof can be made following the same procedure of [23, 24] (for detail, see appendix in [24]). \square

3.3. Adaptive Parameters Estimation. For solving the minimizing energy functional $E(u)$, it often transforms the optimizing problem into the Euler-Lagrange equation. Using the standard computation of Calculus of Variation of $E(u)$ with respect to u , we can get its Euler-Lagrange equation:

$$\begin{aligned} & -\operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) + \alpha(u - u_0) + \beta(\Delta(P_\lambda u_0) - \Delta u) = 0, \\ & \frac{\partial u}{\partial \vec{n}} \Big|_{\partial\Omega} = 0, \end{aligned} \quad (31)$$

where \vec{n} is the outward unit normal vector on the boundary $\partial\Omega$, and $\vec{n} = \nabla u / |\nabla u|$. For a convenient numerical simulation of (31), we apply the gradient descent flow and get the evolution equation:

$$\begin{aligned} & \frac{\partial u}{\partial t} = \operatorname{div} \left(\frac{\nabla u}{|\nabla u|} \right) + \alpha(u_0 - u) + \beta(\Delta u - \Delta(P_\lambda u_0)), \\ & u(x, 0) = u_0(x). \end{aligned} \quad (32)$$

There are three parameters λ, α, β involved in the iterative procedure. For the threshold parameter λ in the curvelet coefficients shrinkage, the common one is to choose $\lambda = k\sigma^2$, where σ denotes the standard deviation of Gaussian white noise. The Monte-Carlo simulations can calculate an approximation value σ^2 of the individual variance. In our experiments, we use the following hard-thresholding rule for estimating the unknown curvelet coefficients:

$$T_\lambda(x) = \begin{cases} x, & |x| \geq k\sigma^2, \\ 0, & |x| < k\sigma^2. \end{cases} \quad (33)$$

Here, we actually chose a scale-dependent value for k : we have $k = 4$ for the first scale (the finest scale) and $k = 3$ for the others.

For the parameters α and β , they are very important to balance trade-off between the image fidelity term and the regularization term. An important prior fact is that the Gaussian distributed noise has the following restriction condition:

$$\int_\Omega (u_0 - u)^2 = |\Omega| \sigma^2. \quad (34)$$

Therefore, we merely multiply the first equation of (32) by $(u_0 - u)$ and integrate by parts over Ω ; if the steady state has been reached, the left side of the first equation of (32) vanishes; thus we have

$$\begin{aligned} & \int_\Omega \operatorname{div} \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon^2}} \right) (u_0 - u) + \alpha \int_\Omega (u_0 - u)^2 \\ & + \beta \int_\Omega (\Delta u - \Delta(P_\lambda u_0))(u_0 - u) = 0. \end{aligned} \quad (35)$$

Obviously, the above equation is not sufficient to estimate the values of α and β simultaneously. This implies that we should introduce another prior knowledge. Borrowing the idea of

spatial varying data fidelity from Gilboa [14], we compute the parameter by the formula:

$$\alpha \approx \frac{(u - u_0) \operatorname{div} \left(\nabla u / \left(\sqrt{|\nabla u|^2 + \varepsilon^2} \right) \right)}{S(x, y)}, \quad (36)$$

where $S(x, y) \approx \sigma^4 / P_R(x, y)$, and $P_R(x, y)$ is the local power of the residue $R = u_0 - u$. The local power of the residue is given by

$$P_R(x, y) = \frac{1}{|\Omega|} \int_{\Omega} (R(\bar{x}, \bar{y}) - \eta(R))^2 w_{x,y}(\bar{x}, \bar{y}) d\bar{x} d\bar{y}, \quad (37)$$

where $w_{x,y}$ is a normalized and radially symmetric Gaussian Function window, and $\eta(R)$ is the expected value. After we compute the value of α , we can estimate the value of β using (38), that is,

$$\beta \approx \frac{\int_{\Omega} \operatorname{div} \left(\nabla u / \left(\sqrt{|\nabla u|^2 + \varepsilon^2} \right) \right) (u_0 - u) dx + \alpha |\Omega| \sigma^2}{\int_{\Omega} (\Delta(P_{\lambda} u_0) - \Delta u) (u_0 - u) dx}. \quad (38)$$

3.4. Description of Proposed Algorithm. To discretize equation (32), the finite difference scheme in [8] is used. Denote the space step by $h = 1$ and the time step by τ . Thus we have

$$\begin{aligned} D_x^{\pm}(u_{i,j}) &= \pm [u_{i\pm 1,j} - u_{i,j}], \\ D_y^{\pm}(u_{i,j}) &= \pm [u_{i,j\pm 1} - u_{i,j}], \\ |D_x(u_{i,j})|_{\varepsilon} &= \sqrt{(D_x^+(u_{i,j}))^2 + (m[D_y^+(u_{i,j}), D_y^-(u_{i,j})])^2} + \varepsilon, \\ |D_y(u_{i,j})|_{\varepsilon} &= \sqrt{(D_y^+(u_{i,j}))^2 + (m[D_x^+(u_{i,j}), D_x^-(u_{i,j})])^2} + \varepsilon, \end{aligned} \quad (39)$$

where $m[a, b] = (\operatorname{sign}(a) + \operatorname{sign}(b)/2) \cdot \min(|a|, |b|)$ and $\varepsilon > 0$ is the regularized parameter chosen near 0.

The numerical algorithm for (32) is given in the following (the subscripts i, j are omitted):

$$\begin{aligned} \frac{u^{(n+1)} - u^{(n)}}{\tau} &= \left(D_x^- \left(\frac{D_x^+ u^{(n)}}{|D_x u^{(n)}|_{\varepsilon}} \right) + D_y^- \left(\frac{D_y^+ u^{(n)}}{|D_y u^{(n)}|_{\varepsilon}} \right) \right) \\ &+ \alpha (u_0 - u^{(n)}) + \beta (\Delta u^{(n)} - \Delta(P_{\lambda} u_0)) \end{aligned} \quad (40)$$

with boundary conditions

$$\begin{aligned} u_{0,j}^{(n)} &= u_{1,j}^{(n)}, & u_{N,j}^{(n)} &= u_{N-1,j}^{(n)}, \\ u_{i,0}^{(n)} &= u_{i,1}^{(n)}, & u_{i,N}^{(n)} &= u_{i,N-1}^{(n)}, \end{aligned} \quad (41)$$

for $i, j = 1, 2, \dots, N-1$. The parameters are chosen like this: $\tau = 0.02$, $\varepsilon = 1$, while the parameters λ, α, β are computed

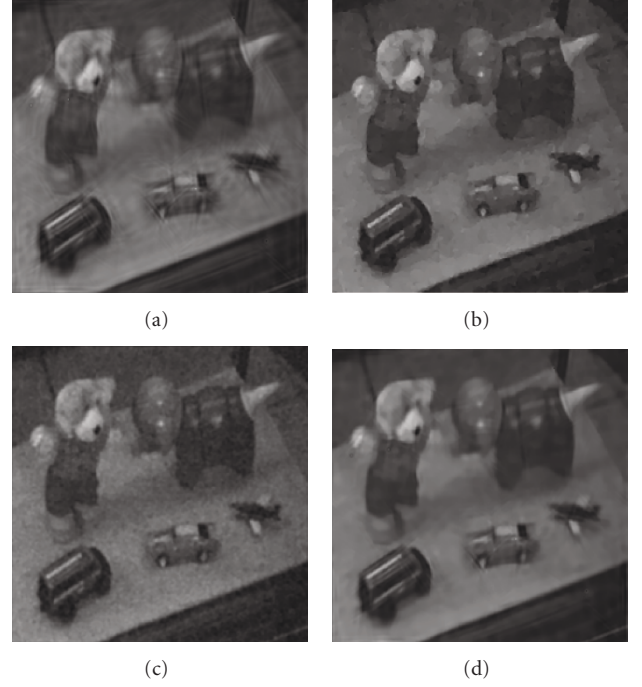


FIGURE 5: The restored “Toys” images (a) by the curvelet shrinkage method (SNR = 13.33, MSSIM = 0.85); (b) by using the “TV” model (SNR = 14.21, MSSIM = 0.85); (c) by using the “TVGF” model (SNR = 13.28, MSSIM = 0.72); (d) by using our proposed model (SNR = 14.37, MSSIM = 0.88).

dynamically during the iterative process according to the formulae (33), (36), and (38).

In summary, according to the gradient descent flow and the discussion of parameters choice, we now present a sketch of the proposed algorithm (the pipeline is shown in Figure 2).

Initialization. $u^{(0)} = u_0$, $\alpha^{(0)} > 0$, $\beta^{(0)} > 0$, Iterative-Steps

Curvelet Shrinkage.

- (1) Apply curvelet transform (the FDCT [5]) to noisy image $u^{(0)}$ and obtain the discrete coefficients $c_{j,l,k}$.
- (2) Use robust method to estimate an approximation value σ , and then use the shrinkage operator in (33) to obtain the estimated coefficients $\tilde{c}_{j,l,k}$.
- (3) Apply the inverse curvelet transform and obtain the initial restored image $P_{\lambda} u_0$.

Iteration. **While** $n < \text{Iterative-Steps}$ **Do**

- (1) Compute $u^{(n+1)}$ according to (40).
- (2) Update the parameter $\alpha^{(n+1)}$ according to (36).
- (3) Update the parameter $\beta^{(n+1)}$ according to (38).

End Do

Output: u^* .

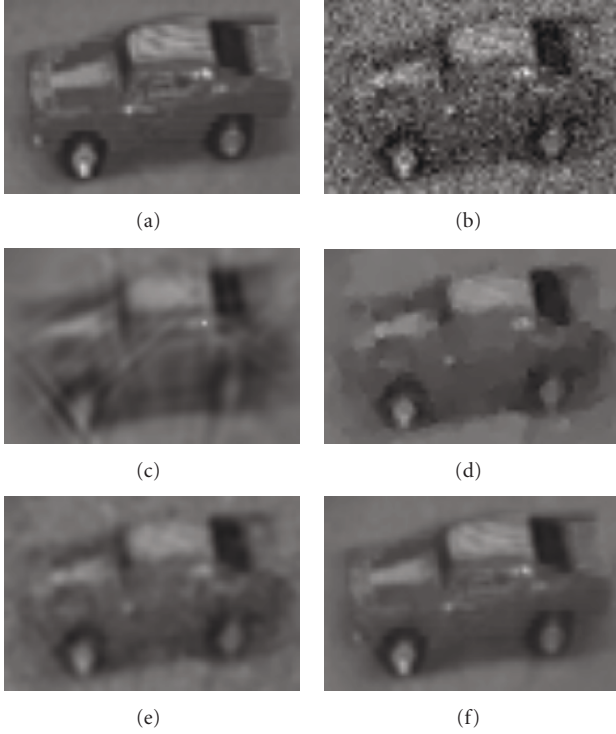


FIGURE 6: The subimages of the original, noisy, and restored “Toys” in Figure 5: (a) the original image; (b) the noisy image; (c) the restored image by the curvelet Shrinkage method; (d) the restored image by using the “TV” model; (e) the restored image by using the “TVGF” model; (f) the restored image by using our proposed model.

3.5. Analysis of Staircase and Curve-Like Effect Alleviation. The essential idea of denoising is to obtain the cartoon part of the image u_C , preserve more details of the edge and texture parts u_T , and filter out the noise part u_n . In classical TV algorithm and Curvelet threshold algorithm, the staircase effects and curve-like artifacts are often generated in the restored cartoon part u_C , respectively. Our model provides a similar ways to force gradient to be close to an approximation. However, our model provides a better mechanism to alleviate the staircase effects and curve-like artifacts.

Firstly, the “TVGF” model in [7] uses the Gaussian filtering to approximation. However, because Gaussian filter is uniform smoothing in all directions of an image, it will smooth the image too much to preserve edges. Consequently, their gradient fidelity term cannot maintain the variation of intensities well. Differing from the TVGF model, our model takes full advantage of curvelet transform. The curvelets allow an almost optimal sparse representation of object with C^2 -singularities. For a smooth object u with discontinuities along C^2 -continuous curves, the best m -term approximation u_m by curvelet thresholding obeys $\|u - u_m\| \leq Cm^{-2}(\log m)^3$, while for the wavelets the decay rate is only m^{-1} .

Secondly, from the regularization theory, the gradient fidelity term $\int_{\Omega} \beta \|\nabla u - \nabla(P_{\lambda}u_0)\|^2 dx dy$ works as Tikhonov regularization in Sobolev functional space $W^{1,2}(\Omega) = \{u \in$

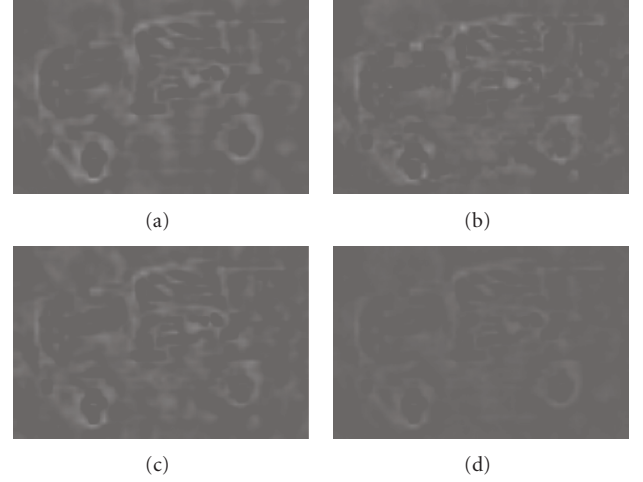


FIGURE 7: The difference images between the original “Toys” image and restored “Toys” images in Figure 6: (a) by the curvelet Shrinkage method; (b) by using the “TV” model; (c) by using the “TVGF” model; (d) by using our proposed model.

$L^2(\Omega); \nabla u \in L^2(\Omega) \times L^2(\Omega)\}$. The problem of $\inf(E(u), u \in W^{1,2}(\Omega))$ admits a unique solution characterized by the Euler-Lagrange equation $(\Delta u - \Delta(P_{\lambda}u_0)) = 0$. Moreover, the function $u - P_{\lambda}u_0$ is called harmonic (subharmonic, superharmonic) Ω if it satisfies $\Delta(u - P_{\lambda}u_0) = (\geq, \leq)0$. Using the mean value theorems [25], for any ball $B_R(x, y) \subset \Omega$, we have

$$\begin{aligned} \Delta(u - P_{\lambda}u_0) &= (\geq, \leq)0 \\ \implies u(x, y) &= (\leq, \geq)P_{\lambda}u_0 \\ &+ \frac{1}{\pi R^2} \int_B (u - P_{\lambda}u_0) dx' dy'. \end{aligned} \quad (42)$$

However, in [7], the gradient fidelity term is chosen as

$$\begin{aligned} \Delta(u - G_{\sigma} \otimes u_0) &= (\geq, \leq)0 \\ \implies u(x, y) &= (\leq, \geq)G_{\sigma} \otimes u_0 \\ &+ \frac{1}{\pi R^2} \int_B (u - G_{\sigma} \otimes u_0) dx' dy'. \end{aligned} \quad (43)$$

Comparing the above two results, we can understand the difference between two gradient fidelity term smoothing mechanism. We remark that the gradient fidelity term in [7] will tend to produce more edge blurring effect and remove more texture components with the increasing of the scale parameter σ of Gaussian kernel, although it helps to alleviate the staircase effect and produces some smoother results. However, our model tends to produce the curvelet shrinkage image and can remain the curve singularities in images; thus it will obtain good edge preserving performance.

In addition, another rationale behind our proposed model is that the spatially varying fidelity parameters $\alpha(x, y)$ and $\beta(x, y)$ are incorporated into our model. In our proposed

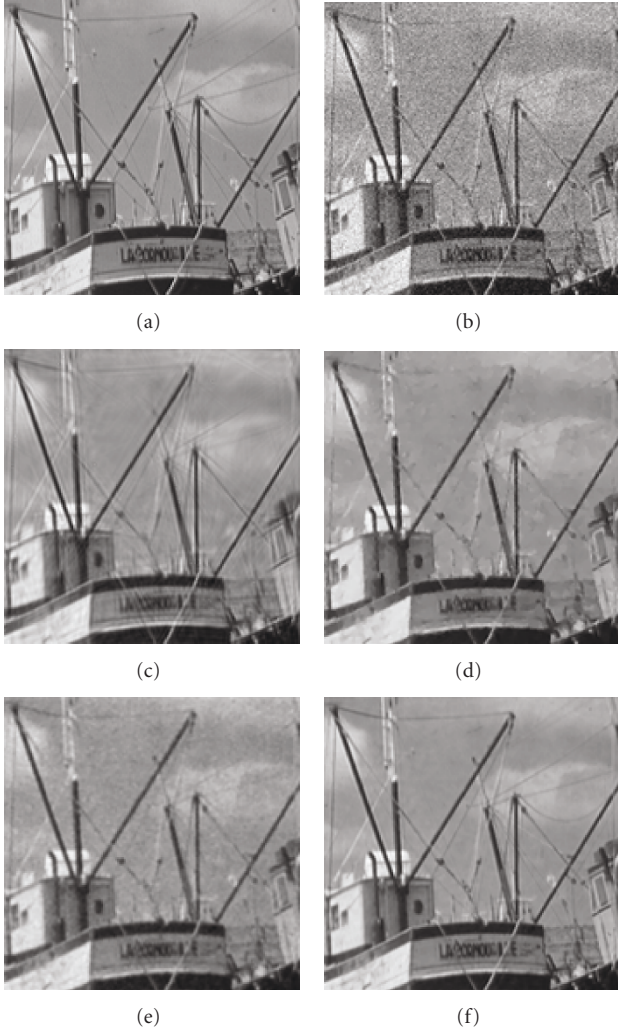


FIGURE 8: The original, noisy, and restored “Boat” images: (a) the original image; (b) the noisy image (SNR = 7.36, MSSIM = 0.43); (c) the restored image by the curvelet Shrinkage method (SNR = 13.51, MSSIM = 0.76); (d) the restored image by using the “TV” model (SNR = 14.36, MSSIM = 0.78); (e) the restored image by using the “TVGF” model (SNR = 13.25, MSSIM = 0.76); (f) the restored image by using our proposed model (SNR = 14.67, MSSIM = 0.79).

algorithm, as described in (36), we use the measure $S(x, y) \approx \sigma^4/P_R(x, y)$; here $P_R(x, y)$ is the local power of the residue $R = u_0 - u$. In the flat area where $u_R \approx u_n$ (basic cartoon model without textures or fine-scale details), the local power of the residue is almost constant $P_R(x, y) \approx \sigma^2$ and hence $S(x, y) \approx \sigma^2$. We get a high-quality denoising process $u \approx u_C = u_{\text{orig}}$ so that the noise, the staircase effect, and the curve-like artifacts are smoothed. In the texture area, since the noise is uncorrelated with the signal, thus the total power of the residue can be approximated as $P_{NC}(x, y) + P_n(x, y)$, the sum of local powers of the noncartoon part and the noise, respectively. Therefore, textured regions are characterized by high local power of the residue. Thus, our algorithm will reduce the level of filtering so that it will preserve the detailed structure of such regions.

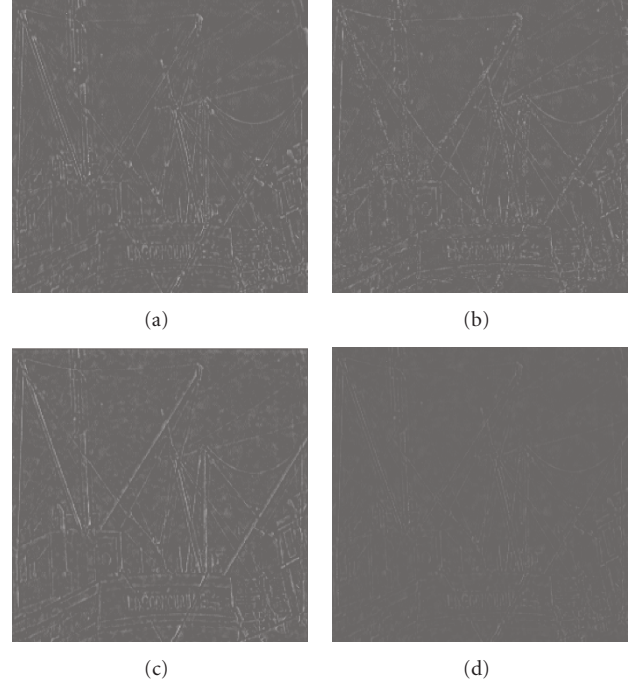


FIGURE 9: The difference images between the original “Boat” and restored “Boat” images in Figure 8: (a) by the curvelet Shrinkage method; (b) by using the “TV” model; (c) by using the “TVGF” model; (d) by using our proposed model.

4. Experimental Results and Analysis

In this section, experimental results are presented to demonstrate the capability of our proposed model. The results are compared with those obtained by using the curvelet shrinkage method [26], the “TV” model (2) proposed by Rudin et al. [8], and the “TVGF” model proposed by Zhu and Xia [7].

In the curvelet shrinkage method, denoising is achieved by hard-thresholding of the curvelet coefficients. We select the thresholding at $3\sigma_{j,l}$ for all but the finest scale where it is set at $4\sigma_{j,l}$; here $\sigma_{j,l}$ is the noise level of a coefficient at scale j and angle l . In our experiments, we actually use a robust estimator to estimate noise level using the following formula: $\sigma_{j,l} = \text{MED}(\text{abs}(c(j, l) - \text{MED}(c(j, l))))/0.6745$, here, $c(j, l)$ represents the corresponding curvelet coefficients at scale j and angle l , and $\text{MED}(\cdot)$ represents the medium operator to calculate the medium value for a sequence coefficients.

The solution of the “TV” model in (2) is obtained by using the following explicit iterative:

$$\frac{u^{(n+1)} - u^{(n)}}{\tau} = \left(D_x^- \left(\frac{D_x^+ u^{(n)}}{|D_x u^{(n)}|_\varepsilon} \right) + D_y^- \left(\frac{D_y^+ u^{(n)}}{|D_y u^{(n)}|_\varepsilon} \right) \right) + \lambda(u_0 - u^{(n)}). \quad (44)$$

Here τ and ε are set to be 0.02 and 0.01, respectively. The regularization parameter λ is dynamically updated to satisfy the noise variance constraints according to (4).

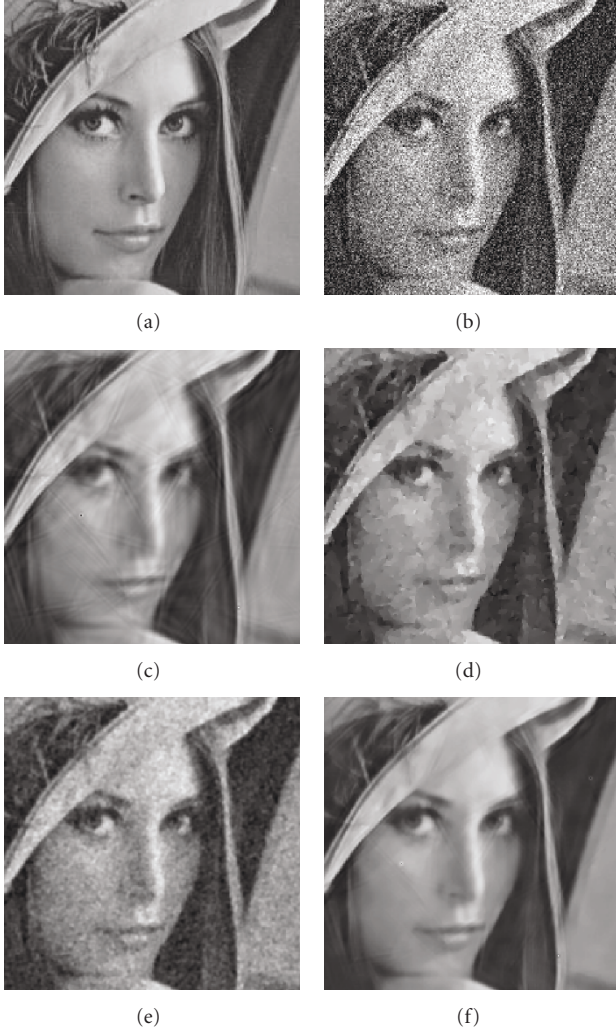


FIGURE 10: The subimages of the original, noisy, and restored “Lena”: (a) the original image; (b) the noisy image (SNR = 1.54, MSSIM = 0.15); (c) the restored image by the curvelet shrinkage method (SNR = 13.72, MSSIM = 0.79); (d) the restored image by using the “TV” model (SNR = 12.94, MSSIM = 0.71); (e) the restored image by using the “TVGF” model (SNR = 12.01, MSSIM = 0.61); (f) the restored image by using the proposed model (SNR = 14.36, MSSIM = 0.82).

The solution of the “TVGF” model is obtained by using the following explicit iterative:

$$\begin{aligned} \frac{u^{(n+1)} - u^{(n)}}{\tau} = & \left(D_x^- \left(\frac{D_x^+ u^{(n)}}{|D_x u^{(n)}|_\varepsilon} \right) + D_y^- \left(\frac{D_y^+ u^{(n)}}{|D_y u^{(n)}|_\varepsilon} \right) \right) \\ & + \beta \left(\Delta u^{(n)} - \Delta(G_\sigma \otimes u_0) \right) \\ & + \alpha (u_0 - u^{(n)}). \end{aligned} \quad (45)$$

Here the size and standard deviation σ of Gaussian lowpass filter G_σ are set to be 7 and 1, respectively. The evolution step

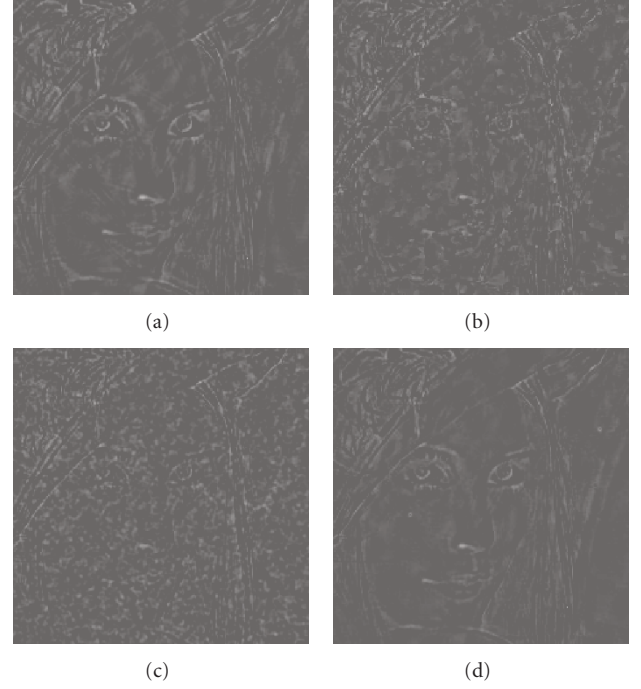


FIGURE 11: The difference images between the original “Lena” and restored “Lena” images in Figure 10: (a) by the curvelet Shrinkage method; (b) by using the “TV” model; (c) by using the “TVGF” model; (d) by using our proposed model.

length τ is set to be 0.02. The weight β is set to be 0.5 as suggested in [7], and α is dynamically updated according to

$$\begin{aligned} \alpha = & \frac{1}{|\Omega| \sigma^2} \\ & \times \left[\int_{\Omega} \operatorname{div} \left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \varepsilon}} \right) (u - u_0) dx \right. \\ & \left. + \beta \int_{\Omega} (\Delta(G_\sigma \otimes u_0) - \Delta u)(u_0 - u) dx \right], \end{aligned} \quad (46)$$

where ε is set to be 0.01, and σ^2 is the noise variance.

4.1. Image Quality Assessment. For the following experiments, we compute the quality of restored images by the signal-to-noise ratio (SNR) to compare the performance of different algorithms. Because of the limitation of SNR on capturing the subjective appearance of the results, the mean structure similarity (MSSIM) index as defined in [27] is used to measure the performance of the different methods. As shown by theoretical and experimental analysis [27], the MSSIM index intends to measure the perceptual quality of the images.

SNR is defined by

$$\text{SNR} = 10 \log_{10} \left\{ \frac{\|u - \mu_u\|^2}{\|u^* - u\|^2} \right\}. \quad (47)$$

TABLE 1: The SNR, MSSIM of the restored “Lena” images by using four models. The numbers in the bracket under the “MSSIM” column refer to the total iteration number of the algorithm.

Noise standard deviation	Noisy image		Curvelet shrinkage method		“TV” Model		“TVGF” Model		Proposed model	
	SNR (dB)	MSSIM	SNR (dB)	MSSIM	SNR (dB)	MSSIM	SNR (dB)	MSSIM	SNR (dB)	MSSIM
20	7.59	0.34	16.62	0.86	16.76	0.85 (1923)	15.68	0.82 (969)	17.26	0.88 (486)
25	5.63	0.27	15.70	0.84	15.68	0.82 (2290)	14.75	0.77 (1115)	16.36	0.86 (615)
30	4.04	0.22	14.95	0.82	14.81	0.80 (2991)	13.79	0.71 (1135)	15.58	0.85 (763)
35	2.70	0.18	14.29	0.81	13.92	0.76 (3000)	12.88	0.67 (1142)	14.92	0.84 (898)
40	1.54	0.15	13.72	0.79	12.94	0.71 (3000)	12.01	0.61 (1151)	14.36	0.82 (1033)

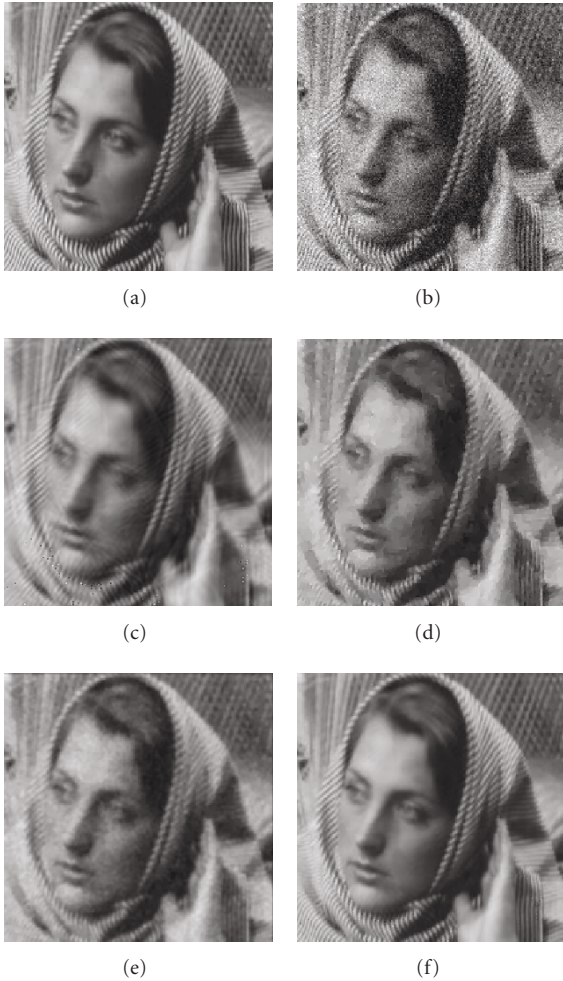


FIGURE 12: The detail of the original, noisy, and restored “Barbara” images: (a) the original image; (b) the noisy image (SNR = 8.73, MSSIM = 0.48); (c) the restored image by the curvelet shrinkage method (SNR = 12.05, MSSIM = 0.77); (d) the restored image by using the “TV” model (SNR = 12.74, MSSIM = 0.77); (e) the restored image by using the “TVGF” model (SNR = 11.39, MSSIM = 0.72); (f) the restored image by using our proposed model (SNR = 13.15, MSSIM = 0.81).

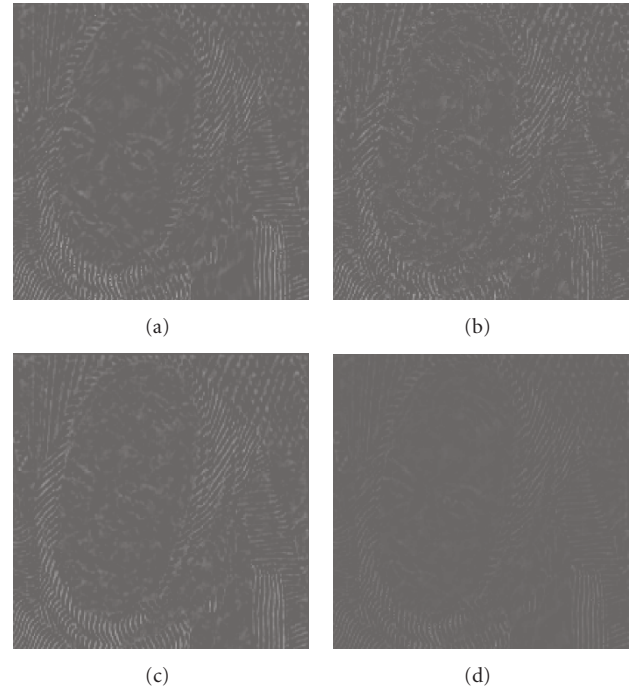


FIGURE 13: The difference images between the original “Barbara” and restored “Barbara” images in Figure 12: (a) by the curvelet Shrinkage method; (b) by using the “TV” model; (c) by using the “TVGF” model; (d) by using our proposed model.

MSSIM is given by

$$\text{MSSIM}(u, u^*) = \frac{1}{M} \sum_{j=1}^M \text{SSIM}(u_j, u_j^*), \quad (48)$$

where u and u^* are the original and the restored images, respectively; μ_{u^*} is the mean of the image u^* ; u_j and u_j^* are the image contents at the j th local window, respectively; M is the number of local windows in the image;

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (49)$$

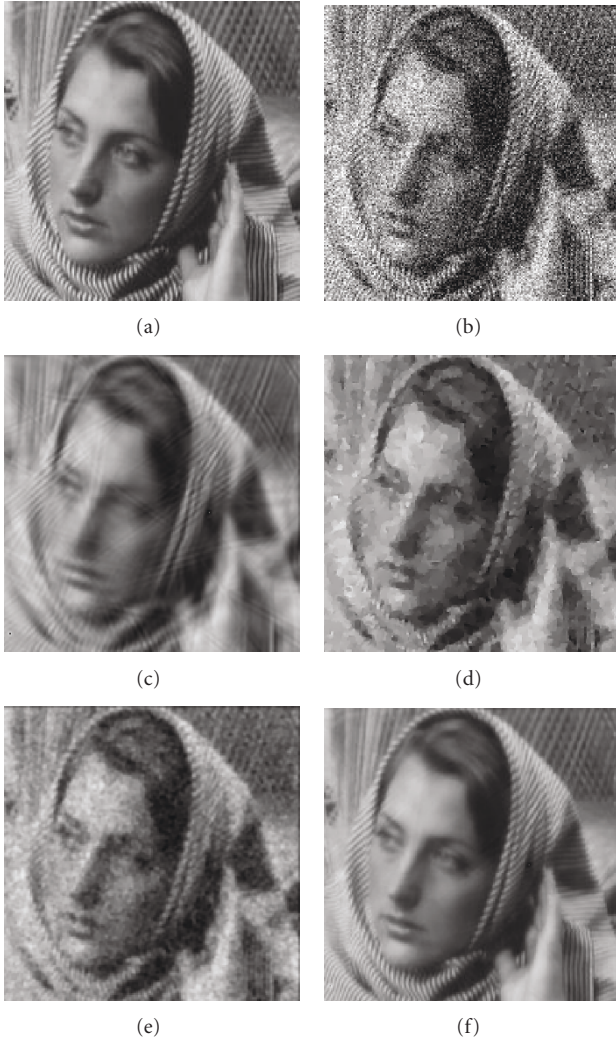


FIGURE 14: The subimages of the original, noisy, and restored “Barbara” images: (a) the original image; (b) the noisy image (SNR = 2.71, MSSIM = 0.26); (c) the restored image by Curvelet Shrinkage method (SNR = 10.45, MSSIM = 0.68); (d) the restored image by using the “TV” model (SNR = 10.12, MSSIM = 0.61); (e) the restored image by using the “TVGF” model (SNR = 9.93, MSSIM = 0.56); (f) the restored image by using our proposed model (SNR = 10.81, MSSIM = 0.70).

where μ_x is the mean of the image x ; σ_x is the standard deviation of the image x ; σ_{xy} is the covariance of the image x and image y ; C_1, C_2 are the constants.

In order to evaluate the performance of alleviation of staircase effect and curve-like artifacts, the difference image between the restored image and the original clean image is used to visually judge image quality.

The stopping criterion of both the proposed method, “TV” method, and the “TVGF” method is that the MSSIM reached maximum or the total iteration number reached the maximal iteration number 3000.

4.1.1. Qualitative and Quantitative Results. Firstly, we take “Toys” image (see Figure 4(a)) as the test image, and we add

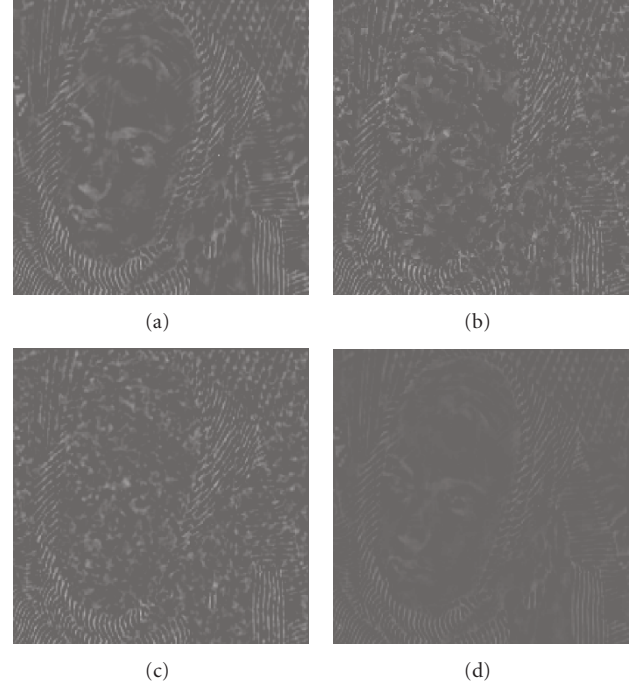


FIGURE 15: The difference images between the original “Barbara” and restored “Barbara” images in Figure 14: (a) by the curvelet Shrinkage method; (b) by using the “TV” model; (c) by using the “TVGF” model; (d) by using our proposed model.

Gaussian noise with standard deviation $\sigma = 20$ (the noisy image is shown in Figure 4(b)). The SNR value of noisy Toys image is 3.99 dB while the value of MSSIM is 0.30. The results of the curvelet shrinkage, the TV algorithm, the TVGF algorithm, and our proposed algorithm are displayed in Figures 5(a)–5(d), respectively. All the algorithm can improve the value of SNR and MSSIM greatly, the value of SNR obtained by our algorithm reaches 14.37 dB, while those obtained by the TVGF model, the curvelet algorithm and the TV model is just 13.28 dB, 13.33 dB, and 14.21 dB, respectively. Similarly, the improvement of MSSIM obtained by our algorithm is the largest among all the algorithms. From this kind of cartoon image, our proposed algorithm does a good job at restoring faint geometrical structures of the image. In order to make better judgement on the visual difference of different restored images of different algorithms, we display the subimages which are cropped from Figures 5(a)–5(d). As illustrated in Figures 6(a)–6(f), the restored image by our proposed image (Figure 6(f)) is more nature than any other restored images as shown in Figures 6(c), 6(d), and 6(e). Figure 7 shows the difference image between “Toys” image and restored “Toys” images in Figure 6. From Figure 6, we can see that our algorithm’s restored image has less staircase and curve-like distortion comparing with other algorithm’s result.

In the second experiments, we take “Boat” as the test image. The “Boat” is a kind of image which contains many linear edges and curve singularities. The restored results are depicted in Figure 8, and the difference images are shown

TABLE 2: The SNR, MSSIM of the restored “Barbara” images by using four models. The numbers in the bracket under the “MSSIM” column refer to the total iteration number of the algorithm.

Noise standard deviation	Noisy image		Curvelet shrinkage method		“TV” Model		“TVGF” Model		Proposed model	
	SNR (dB)	MSSIM	SNR (dB)	MSSIM	SNR (dB)	MSSIM	SNR (dB)	MSSIM	SNR (dB)	MSSIM
20	8.73	0.48	12.05	0.77	12.74	0.77 (2032)	11.39	0.72 (724)	13.15	0.81 (427)
25	6.79	0.40	11.42	0.74	11.83	0.72 (2634)	11.05	0.68 (836)	12.26	0.78 (567)
30	5.21	0.34	11.02	0.71	11.17	0.69 (3000)	10.69	0.64 (894)	11.64	0.75 (687)
35	3.91	0.30	10.72	0.70	10.63	0.65 (3000)	10.32	0.60 (936)	11.19	0.73 (783)
40	2.71	0.26	10.45	0.68	10.12	0.61 (3000)	9.93	0.56 (953)	10.81	0.70 (928)

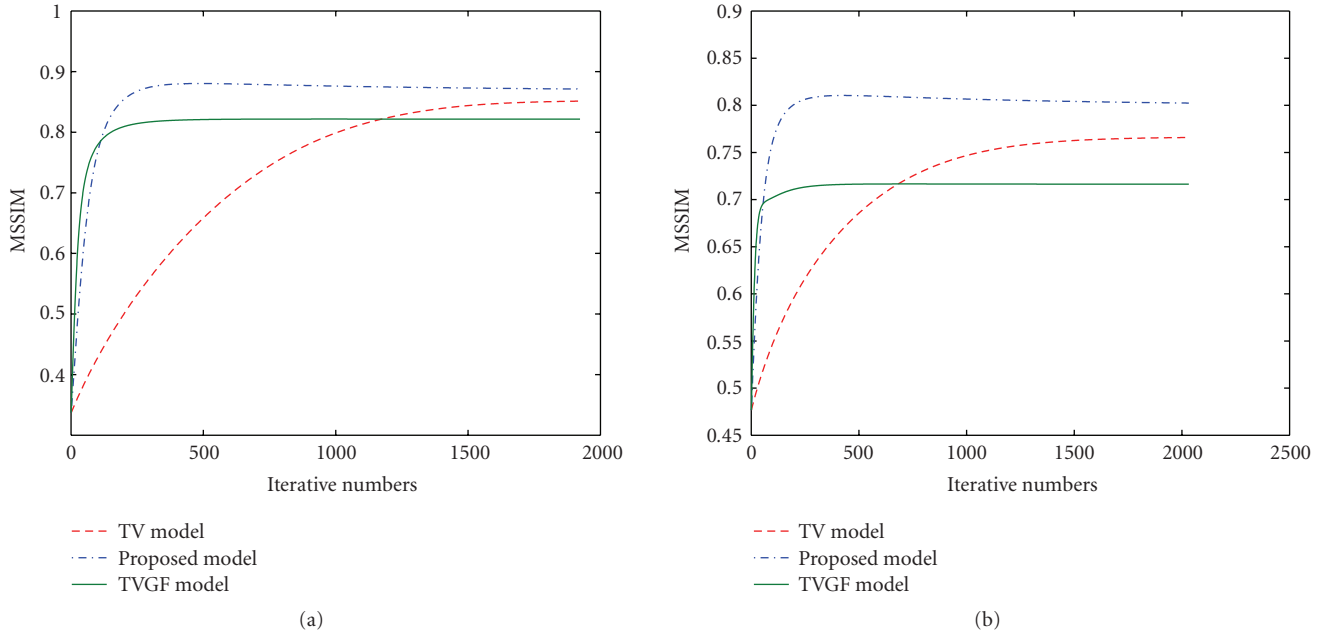


FIGURE 16: Comparison of the MSSIM results of the different images which is corrupted by some Gaussian noise of standard deviation 20: (a) “Lena”; (b) “Barbara”.

in Figure 9. As expected, our hybrid method is less prone to the staircase effect and curve-like artifacts and takes benefits from the curvelet transform for capturing efficiently the geometrical content of image.

In the third experiments, we carry experiments on “Lena” with different noise level and the standard derivation of Gaussian noise is 20, 25, 30, 35, and 40. Figures 10(a)–10(f) display the results for the noisy “Lena” image when the standard deviation of noise is 40. In this case, the value of SNR and MSSIM of the noisy image is 1.54 dB and 0.15, the value of SNR obtained by our algorithm reaches 14.36 dB, while that obtained by the TVGF model, the curvelet algorithm, and the TV model is just 12.01 dB, 13.72 dB and 12.94 dB, respectively. As we know, the “Lena” image is an international standard test image which not only

contains the strong vertical and curve edge but also have hair texture. From the result of curvelet shrinkage (Figure 10(c)), the hair texture and the strong vertical and curve edge are restored very good; however some annoying curve-like artifacts are generated. The TV algorithm can preserve the edge, but it lose the hair texture; moreover, the staircase effect is very obvious (Figure 10(d)). The TVGF algorithm can reduce the staircase effect; however, the edges and textures in “Lena” image look a bit blurred. Compared with the other algorithms, our hybrid algorithm obtains a sharper image with better hair detail, and the restored image is almost free of artifacts. The values of MSSIM of four algorithms also show that our algorithm has the best image perceptual quality. From Table 1, we can find that our algorithm can obtain the largest value of MSSIM. We just display two group

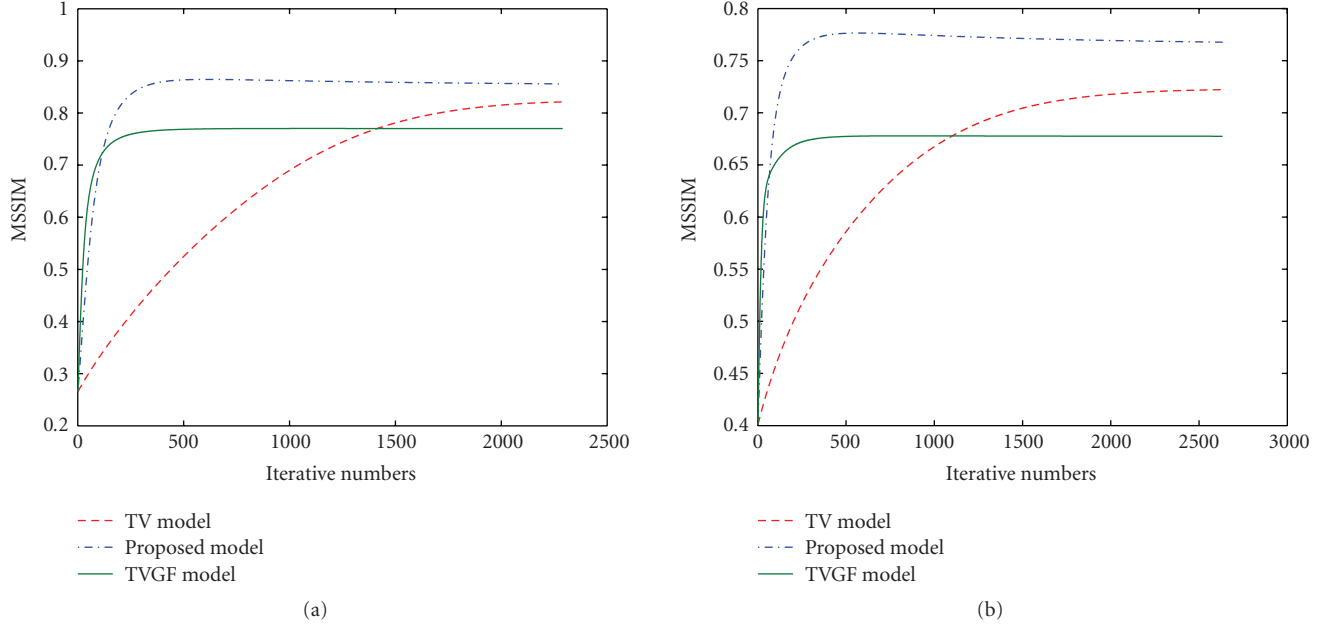


FIGURE 17: Comparison of the MSSIM results of the different images which is corrupted by some Gaussian noise of standard deviation 25: (a) “Lena”; (b) “Barbara”.

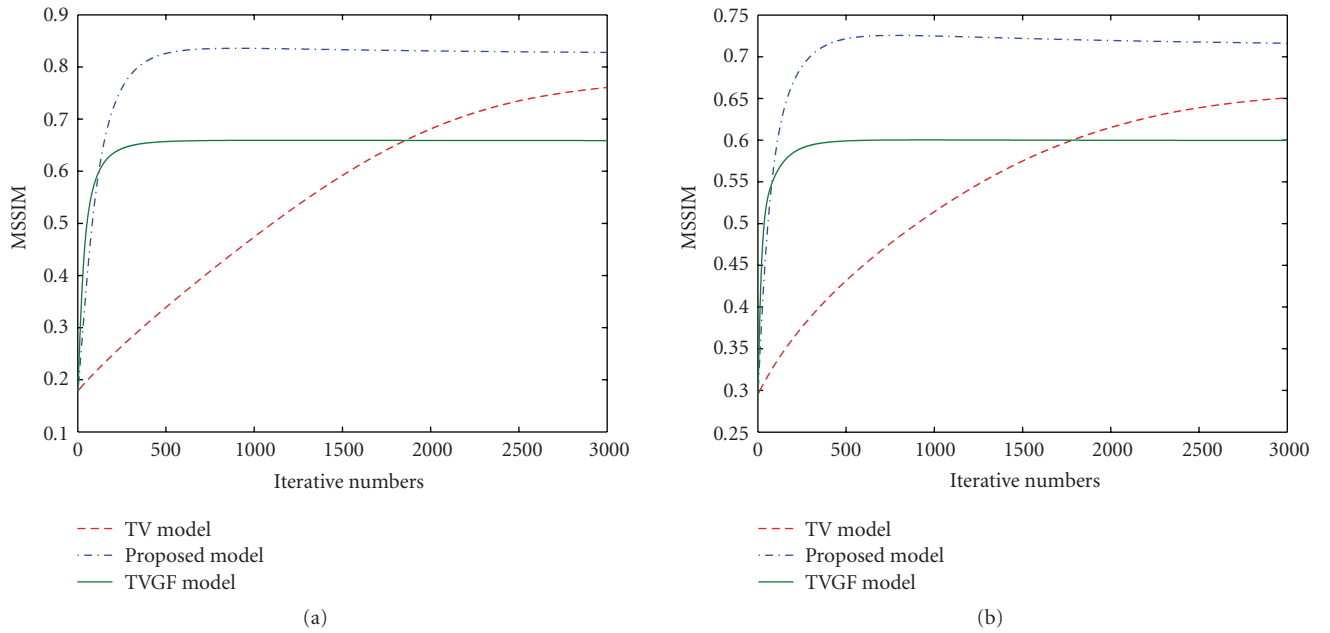


FIGURE 18: Comparison of the MSSIM results of the different images which is corrupted by some Gaussian noise of standard deviation 35: (a) “Lena”; (b) “Barbara”.

results (Figures 10 and 11) for the face part of “Lena” image of various algorithms in the cases of Gaussian noise with standard deviation $\sigma = 40$, respectively.

In the fourth experiments, we take “Barbara” as the test image. The “Barbara” is a kind of image which contains many texture details. The distinguished characteristic of the Barbara image is that it has abundant textures. We display two group results (Figures 12, 13, 14, and 15) for the face

part of “Barbara” image of various algorithms in the cases of Gaussian noise with standard deviation $\sigma = 20$ and $\sigma = 40$, respectively. From these experiments, the visual quality of restored image of our algorithm is better than any other algorithms. Of course, our hybrid algorithm can only restore the regular texture partially, some tiny textures cannot be restored, but the whole image looks more nature and more harmonious.

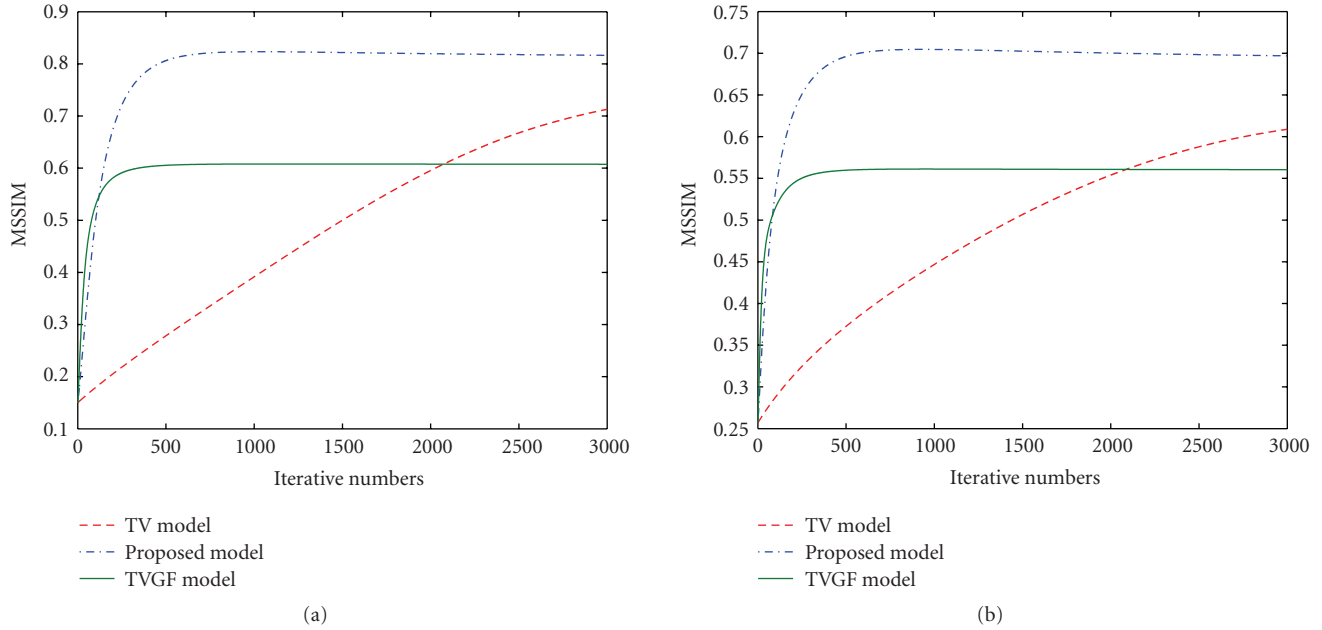


FIGURE 19: Comparison of the MSSIM results of the different images which is corrupted by some Gaussian noise of standard deviation 40: (a) "Lena"; (b) "Barbara".

In addition, we use the SNR and MSSIM index to evaluate the quality of restored images of various algorithms under different noise level. For fair comparison, all the results are the images whose values of MSSIM reach maximum in their iterative process, respectively. In order to systematically evaluate the performance of the various algorithms from different noise levels, the denoising performance results are tabulated in Tables 1 and 2 for the "Lena" image and the "Barbara" image, respectively. By inspection of these tables, both the SNR improvement and the MSSIM improvement by our model are larger than the other three models. It is interesting to point out that the performance of the TVGF algorithm is not good compared with three other algorithms. This phenomenon has been discussed and analyzed by the research of literature [28] and the reason is the blurring caused by Gaussian kernel convolution. Of course, the TVGF algorithm actually reduces the staircase effect to some extent. Differing with Gaussian kernel convolution, our proposed algorithm takes advantage of the anisotropic of the curvelets; therefore there is no any edge blurring. The MSSIM improvement shows that our algorithm has better "subjective" or perceptual quality. Especially, the important improvement becomes more salient as noise level increase. Also we can see, both the SNR improvement and the MSSIM improvement obtained by our algorithm become more salient than those obtained by other algorithms for more complex images with more textures, for instance, the Barbara image. Figures 16, 17, 18, and 19 show the amplitude of the MSSIM improvement during the iterative process of the TV model, the TVGF model, and our proposed model. Tables 1 and 2 also list the total iteration number of each algorithm. For example, in Table 1, when the noise standard deviation is 40, the "TV" model needs 3000 iteration steps

to reach the maximum MSSIM (0.71), the "TVGF" model needs 1151 iteration steps to reach the maximum MSSIM (0.61), the "TVGF" model needs 1151 iteration steps to reach the maximum MSSIM (0.61), while our model can obtain the good quality restored image with the value of MSSIM being 0.82 using 1033 iteration steps. These experiments show that our algorithm can more quickly reach the best restored image with maximum MSSIM, and our algorithm obtains the best performance comparing with the other algorithms.

5. Conclusion

In this paper, a curvelet shrinkage fidelity-based total variation regularization is proposed for discontinuity-preserving denoising. We propose a new gradient fidelity term, which is designed to force the gradients of desired image to be close to the curvelet approximation gradients. To improve the ability of preserving the details of edges and texture, the spatial-varying parameters are adaptively estimated in the iterative process of the gradient descent flow algorithm. We carry out many numerical experiments to compare the performance of various algorithms. The SNR and MSSIM improvements demonstrate that our proposed method has the best performance than the TV algorithm, the curvelet shrinkage, and the TVGF algorithm. Our future work will extend this new gradient fidelity term to other PDE-based methods.

Acknowledgments

The authors would like to express their gratitude to the anonymous referees for making helpful and constructive

suggestions. The authors also thank financial support from the National Natural Science Foundation of China (60802039) and the National 863 High Technology Development Project (2007AA12Z142), Specialized Research Fund for the Doctoral Program of Higher Education (20070288050), and NUST Research Funding under Grant no. 2010ZDJH07.

References

- [1] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490–530, 2005.
- [2] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091–2106, 2005.
- [3] E. Candès and D. Donoho, "Curvelets—a surprisingly effective nonadaptive representation for objects with edges," in *Curves and Surface Fitting: Saint-Malo 1999*, A. Cohen, C. Rabut, and L. Schumaker, Eds., pp. 105–120, Vanderbilt University Press, Nashville, Tenn, USA, 2000.
- [4] E. J. Candès and D. L. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities," *Communications on Pure and Applied Mathematics*, vol. 57, no. 2, pp. 219–266, 2004.
- [5] E. Candès, L. Demanet, D. Donoho, and L. Ying, "Fast discrete curvelet transforms," *Multiscale Modeling and Simulation*, vol. 5, no. 3, pp. 861–899, 2006.
- [6] J. Ma and G. Plonka, "Combined curvelet shrinkage and nonlinear anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2198–2206, 2007.
- [7] L. Zhu and D. Xia, "Staircase effect alleviation by coupling gradient fidelity term," *Image and Vision Computing*, vol. 26, no. 8, pp. 1163–1170, 2008.
- [8] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [9] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
- [10] L. Alvarez, P. Lions, and J. Morel, "Image selective smoothing and edge detection by nonlinear diffusion. II," *SIAM Journal on Numerical Analysis*, vol. 29, no. 3, pp. 845–866, 1992.
- [11] T. Chan, S. Esedoglu, F. Park, and A. Yip, "Recent developments in total variation image restoration," in *Handbook of Mathematical Models in Computer Vision*, N. Paragios, Y. Chen, and O. Faugeras, Eds., Springer, Berlin, Germany, 2004.
- [12] M. Lysaker and X.-C. Tai, "Iterative image restoration combining total variation minimization and a second-order functional," *International Journal of Computer Vision*, vol. 66, no. 1, pp. 5–18, 2006.
- [13] F. Li, C. Shen, J. Fan, and C. Shen, "Image restoration combining a total variational filter and a fourth-order filter," *Journal of Visual Communication and Image Representation*, vol. 18, no. 4, pp. 322–330, 2007.
- [14] G. Gilboa, Y. Y. Zeevi, and N. Sochen, "Texture preserving variational denoising using an adaptive fidelity term," in *Proceedings of the 2nd IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision (VLSM '03)*, pp. 137–144, Nice, France, 2003.
- [15] J. Ma and M. Fenn, "Combined complex ridgelet shrinkage and total variation minimization," *SIAM Journal of Scientific Computing*, vol. 28, no. 3, pp. 984–1000, 2006.
- [16] G. Plonka and G. Steidl, "A multiscale wavelet-inspired scheme for nonlinear diffusion," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 4, no. 1, pp. 1–21, 2006.
- [17] J. Ma and M. Fenn, "Combined complex ridgelet shrinkage and total variation minimization," *SIAM Journal of Scientific Computing*, vol. 28, no. 3, pp. 984–1000, 2006.
- [18] L. Ying, L. Demanet, and E. Candès, "3D discrete curvelet transform," in *Wavelets XI*, Proceedings of SPIE, pp. 1–11, San Diego, Calif, USA, August 2005.
- [19] L. Demanet and L. Ying, "Curvelets and wave atoms for mirror-extended images," in *Wavelets XII*, Proceedings of SPIE, San Diego, Calif, USA, August 2007.
- [20] A. Achim, P. Tsakalides, and A. Bezerianos, "SAR image denoising via Bayesian wavelet shrinkage based on heavy-tailed modeling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 8, pp. 1773–1784, 2003.
- [21] S. Durand and M. Nikolova, "Denoising of frame coefficients using ℓ^1 data-fidelity term and edge-preserving regularization," *Multiscale Modeling & Simulation*, vol. 6, no. 2, pp. 547–576, 2007.
- [22] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.
- [23] P. Kornprobst, R. Deriche, and G. Aubert, "Image sequence analysis via partial differential equations," *Journal of Mathematical Imaging and Vision*, vol. 11, no. 1, pp. 5–26, 1999.
- [24] L. Xiao, L.-L. Huang, and Z.-H. Wei, "A weberized total variation regularization-based image multiplicative noise removal algorithm," *EURASIP Journal on Advances in Signal Processing*, vol. 2010, Article ID 490384, 15 pages, 2010.
- [25] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equation of Second Order*, Springer, Berlin, Germany, 2003.
- [26] J.-L. Starck, E. J. Candès, and D. L. Donoho, "The curvelet transform for image denoising," *IEEE Transactions on Image Processing*, vol. 11, no. 6, pp. 670–684, 2002.
- [27] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [28] F. F. Dong and Z. Liu, "A new gradient fidelity for avoiding staircasing effect," *Journal of Computer Science and Technology*, vol. 24, no. 6, pp. 1162–1170, 2009.