

A Novel Variational Model with Strict Convexity for Multiplicative Noise Removal

Xuegang Hu^{a,b}, Yuefang Lou^b, Tianqiao Hu^c

^a Research Center of Systems Theory and Applications, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
Email: huxg@cqupt.edu.cn

^b College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
Email: yuefanglou@aliyun.com

^c College of Mathematics, Sichuan University, Chengdu 610064, China
Email: hutianqiao@hotmail.com

Abstract—In this paper, a novel variational model with strict convexity for removing multiplicative noise from images is proposed and studied. Firstly, by applying maximum likelihood estimation method and the Bayesian formulation, the variational model is derived. Then, we use an alternating minimization algorithm to find out the minimizer of the objective function, and prove the existence of the minimizer for the underlying variational problem in theory. Finally, Our experimental results show that the quality of images denoised by the proposed method is quite good, and the proposed model is superior to the existing key models in preventing the images from stair-casing, and in restoring more texture details of the denoised image.

Index Terms—Image Denosing, Multiplicative Noise, Partial Differential Equation, Convex Function, Variational method

I. INTRODUCTION

DUE to the imperfection of image acquisition systems and transmission channels, images are often corrupted by noise. This degradation leads to a significant reduction of image quality and then makes more difficult to perform high level vision tasks such as recognition, 3-D reconstruction, or scene interpretation [1]. Image denoising plays an important role in the areas of image processing. A variety of methods have been proposed for removing noises from images over the last decades [2]–[4]. The additive noise is the most widely considered noise in the literature, which has been extensively studied over the last decades, and tends to be quite comprehensive and mature [5]–[7]. In many real world image processing applications, multiplicative noises are commonly found, for example in laser images, microscope images, synthetic aperture radar (SAR) images and medical ultrasonic images [8], [9]. How to remove the multiplicative noise in the corrupted images is becoming the hot research issue in recent years [10]–[13].

Manuscript received August 9, 2013; revised September 01, 2013; accepted September 2, 2013. © 2005 IEEE.

This work was supported partly by the National Natural Science Foundation of China (Grant No. 11071266), and partly by the Science Foundation Project of Chongqing Municipal Education Commission (Grant No. KJ100505).

In a multiplicative noise model, a recorded image g , defined on $\Omega \subset \mathbb{R}^2$, is the multiplication of an original image u and a noise n [14]:

$$g = un$$

Several variational approaches for multiplicative noise removal problems are available in the literature [8–12]. The variational approach dealing with multiplicative noise was firstly proposed by Rudin, Lions and Osher in [15] (called the RLO model). The RLO model is written as follows:

$$\min J(u) := \min \left\{ \int_{\Omega} |Du| dx dy + \lambda_1 \int_{\Omega} \frac{g}{u} dx dy + \lambda_2 \int_{\Omega} \left(\frac{g}{u} - 1 \right)^2 dx dy \right\}, \quad (1)$$

where the first term is a regularizer, the latter two terms are data fitting terms, λ_1 and λ_2 are the weighted parameters.

When the multiplicative noise is out of the Gauss distribution, the above RLO model is no longer available. In the past few years, many researchers have developed various denoising models by using PDE methods versus the Gamma noise which is more complex.

Under the assumption that the multiplicative noise follows the Gamma distribution with mean one, Aubert and Aujol used the maximum a posteriori (MAP) regularization approach and derived the denoising model in [11] (called the AA model) with Bayesian rules and variational method as follows:

$$\min E(u) := \min \left\{ \int_{\Omega} |Du| dx dy + \lambda \int_{\Omega} \left(\frac{g}{u} + \log u \right) dx dy \right\}, \quad (2)$$

where λ is the weighted parameter and the last term is the fitting term. Obviously, the energy functional is not convex for all u . Furthermore, for the logarithm function appears in the last term, we must force each entry of u to be positive in the optimization process. Therefore, in practice, some special processing is used to overcome the

difficulty. However, the efficiency of numerical methods for solving (2) is low and the computed solutions by some optimization methods are not necessary to be a global optimal solution of (2). Therefore, the quality of the restored image may not be good, for example, some texture details will be blurred because the “stair-casing effect” appears in the recovered image.

Based on the AA model, the authors applied a logarithmic transformation method to propose a new denoising model (called the JY model) for multiplicative noise in [13] and it eased the “stair-casing effect” to some extent; Motivated by the works of the JY model. Wang, Feng and Huo presented an anisotropy TV denoising model in [16] (called the WFH model) with iterative heavy weights. Despite the protection from losing edge information of the image, the WFH model generates some artificial trace that will blur the observed image.

In [12], Huang, Ng and Wen considered a logarithmic transformation $z = \log u$ in the fitting term of the AA model (2), added a fitting term and proposed the following model (the HNW model):

$$\min E(u) := \min \left\{ \int_{\Omega} (z + ge^{-z}) dx dy + \lambda_1 \int_{\Omega} (z - w)^2 dx dy + \lambda_2 \int_{\Omega} |Dw| dx dy \right\}, \tag{3}$$

where λ_1 and λ_2 are positive regularization parameters. It is clear that when u contains an edge, z also contains an edge at the same location. Moreover, because the second derivative with respect to z of (3) is constantly positive, the function is strictly convex for all z . The process tactfully solved the ill-posed situation of the AA model, and also proved theoretically that the iterative sequence derived from the alternating minimization algorithm converges to the optimal solution of (3). The experimental results also show the quality of the denoised images is quite good.

To the best of our knowledge, however, multiplicative noise removal is still a challenging problem. This is because the noise arises in important applications but up to now, there is no entirely satisfactory methods for handling it. Most of the existing models are effective in denoising, but the problem of reducing or avoiding stair-casing effect has not yet been well solved as the TV regularization yields images containing numerous regions with constant values and consequently the textures and fine details are removed.

How to alleviate or even eliminate the “stair-casing effect” while keeping the advantage of the existing models is one of the problems needing to be solved urgently. And the key is to explore a new regularization term and construct new variational models. Motivated by the thoughts of [12] and [17], in this paper we employ a new regularization term to construct a modified variational model. It is able to remove the Gamma noise effectively and avoid the image blurring. In addition, the proposed model can preserve the edge features and texture details better, which smoothes and improves the visual effect.

The rest of this paper is organized as follows. Section II presents some preliminaries to be used in this paper; In section III, the new model is derived by using Bayesian rules through the analysis of the mathematical characteristics of Gamma noise; Section IV gives the numerical minimization method by the alternating optimization algorithm; Section V proves that the iteration sequence converges to the optimal solution; Section VI shows experimental results to demonstrate the quality of the denoised images and the efficiency of our proposed method; In final section, a brief conclusion is given.

II. PRELIMINARIES

In order to facilitate the mathematical study later in this paper and for convenience of readers, we present here some definitions, and give two important theorems to be used in this paper. Throughout this paper, we will use the following classical notations and distributional spaces. For more details we refer the reader to the references [20]–[22], [24].

Definition 2.1. If for all functions $\varphi = (\varphi_1, \varphi_2) \in C_0^1(\Omega)^2$, $\|\varphi\|_{L^\infty(\Omega)} \leq 1$, the formula

$$\int_{\Omega} u \operatorname{div} \varphi dx = - \int_{\Omega} Du \cdot \varphi dx,$$

holds, then $Du = (D_1u, D_2u)$ is called the distribution gradient of u , and u is called a bounded variation function. Also, the total variation of Du on Ω is defined as follows

$$\int_{\Omega} |Du| dx := \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi dx : \varphi = (\varphi_1, \varphi_2) \in C_0^1(\Omega)^2, |\varphi|_{L^\infty(\Omega)} \leq 1 \right\}.$$

Remark 1. From Green’s formula, we have $\int_{\Omega} u \operatorname{div} \varphi dx = - \int_{\Omega} \nabla u \cdot \varphi dx$, for any $u \in C^1(\Omega)$. Therefore, $\int_{\Omega} |Du| dx = \int_{\Omega} |\nabla u| dx$. That is, if the classical derivative of u exists, the distribution gradient of u is its classical gradient.

Definition 2.2(see [20]) Assume that X is a closed set, $\Gamma_0(X)$ is the set of all the convex lower semicontinuous functions. For any $\varphi \in \Gamma_0(X)$, its proximal operator is defined by

$$\operatorname{prox}_{\varphi}(y) = \arg \min_x \frac{1}{2} \|y - x\|_2^2 + \varphi(x).$$

Definition 2.3(see [20]) The operator P defined in R^2 is called nonexpansive if the following inequality

$$\|P(x_1) - P(x_2)\|_2 \leq \|x_1 - x_2\|_2.$$

holds for any $x_1, x_2 \in R^{n^2}$, where $\|\cdot\|_2$ is the Euclidean norm in R^{n^2} .

Definition 2.4(see [22]) If there exists some nonexpansive operator A and some $\alpha \in (0, 1)$ such that $P = (1 - \alpha)I + \alpha A$, then P is called α -averaged nonexpansive.

Definition 2.5(see [24]) For any sequence $\{x_k\}$ in X , if there exists $\lim_{k \rightarrow \infty} \varphi(x_k) = \infty$ when $\lim_{k \rightarrow \infty} \|x_k\|_2 = \infty$, then we call the φ is coercive in X .

Let $W^{1,p}(\Omega)$ be the standard notation for the Sobolev space. Finally, we give two important theorems to be used later.

Theorem 2.1(see [21]) Assume that $f(x, u, Du)$ is coercive, and let

$$F(u) = \int_{\Omega} f(x, u, Du) dx dy.$$

Then f is convex for Du if and only if

(1) the functional $F(u)$ is lower semicontinuous in $W^{1,p}(\Omega)$ ($p \geq 1$);

(2) the problem $\inf_{\Omega} F(u)$ has solutions. Moreover, there exists a unique solution when f is strictly convex for both Du and u .

Theorem 2.2(see [21]) For any $\varphi \in \Gamma_0(X)$, if there exist its proximal operators, $\text{prox}_{\varphi}x$ and $\text{prox}_{\varphi}y$, then

$$\|\text{prox}_{\varphi}x - \text{prox}_{\varphi}y\| \leq \langle \text{prox}_{\varphi}x - \text{prox}_{\varphi}y, x - y \rangle.$$

III. A NEW VARIATIONAL MODEL FOR REMOVAL OF MULTIPLICATIVE NOISE

Consider the gamma distribution with density function

$$p(x) = \begin{cases} \frac{K^K x^{K-1}}{(K-1)!}, & x \geq 0, \\ 0, & x < 0. \end{cases} \quad (4)$$

where K is an integer. Obviously, the mean value and the variance of the Gamma distribution are 1 and $1/K$, respectively. We are applying the Bayesian rule to establish a proper fidelity term. The aim of this paper is to find the best approximation of the original image u , we denote it by \hat{u} . According to the maximum likelihood estimation and Bayesian rule, we have

$$\begin{aligned} \hat{u} &= \arg \max_u Pr(u|g) = \arg \max_u \frac{Pr(g|u)Pr(u)}{Pr(g)} \\ &= \arg \max_u (Pr(g)Pr(u)). \end{aligned} \quad (5)$$

Noticing $-\log(Pr(g|u)Pr(u)) = -(\log Pr(g|u) + \log Pr(u))$, and combining with Gamma density function and Gibbs formula, we have derived

$$-\log Pr(g|u) = \sum_{i,j=1}^n \left\{ K \left(\log u(i, j) + \frac{g(i, j)}{u(i, j)} \right) \right\}.$$

Thus, Eq.(5) is equivalent to

$$\hat{u} = \arg \min_u \sum_{i,j=1}^n \left\{ \left(\log u(i, j) + \frac{g(i, j)}{u(i, j)} \right) + \lambda \phi(i, j) \right\}, \quad (6)$$

where λ is the regularization parameter, ϕ is a given function. The continuous form of (6) provides the fidelity term $\int_{\Omega} (\log u + \frac{g}{u})$, regularizing term $\int_{\Omega} \phi(|Du|)$. Referring to [14], the fidelity term is transformed to be $\int_{\Omega} (z + ge^{-z})$.

In the variational model (3), $\int_{\Omega} |Dw|$ is the regularization term, $\int_{\Omega} (z + ge^{-z})$ and $\int_{\Omega} (z - w)^2$ are fitting terms. The general form of the regularization terms in the TV denoising models is $\int_{\Omega} |Dw|^p$. It can be proved that the model is ensured to be well-posed when $p \geq 1$. In

general, larger the value of p , stronger the edge penalty; The variational model is ill-posed when $0 < p < 1$, but the edge is preserved well. p is equal to 1 in the regularization terms of the main variational models at present, such as the AA model and the HNW model, etc.

In order to inherit the advantages of the HNW model, and overcome its drawbacks as much as possible, inspired by [18], this paper uses the new regularization term $\int_{\Omega} |Dw| \log(e + |Dw|)$ to propose the unconstrained TV denoising functional as follows:

$$\begin{aligned} &\min_{z,w} J(z, w) \\ &= \min_{z,w} \left\{ \int_{\Omega} (z + ge^{-z}) + \lambda_1 \int_{\Omega} (z - w)^2 \right. \\ &\quad \left. + \lambda_2 \int_{\Omega} |Dw| \log(e + |Dw|) \right\}. \end{aligned} \quad (7)$$

The first term on the right hand side of our model (7) is called the loyalty term which ensures recovering image u to retain the main features from the virtual image $\log g$. The second term is the coordination term which measures the influence between the fitting term and the regularization term to the model. The final term is the regularization term which ensures the smooth of the denoising image w , and removes the noise frequently.

IV. THE ITERATIVE ALGORITHM

Inspired from the thought of [14], this paper uses an alternating minimization algorithm to solve the problem (7). Starting from the initial data $w^{(0)}$, we solve the following optimization problem

$$\begin{cases} z^{(m)} = R(w^{(m-1)}) := \arg \min_z \left\{ \int_{\Omega} (z + ge^{-z}) \right. \\ \quad \left. + \lambda_1 \int_{\Omega} (z - w^{(m-1)})^2 \right\}, \\ w^{(m)} = S(z^{(m)}) := \arg \min_w \left\{ \lambda_1 \int_{\Omega} (z^{(m)} - w)^2 \right. \\ \quad \left. + \lambda_2 \int_{\Omega} |Dw| \log(e + |Dw|) \right\}. \end{cases} \quad (8)$$

and get $z^{(1)}$ and $w^{(1)}$. In the same way, repeating the alternating iteration, we obtain the following sequence : $w^{(0)}, z^{(1)}, w^{(1)}, z^{(2)}, w^{(2)}, \dots, z^{(m)}, w^{(m)}, \dots$.

Firstly, in order to solve the first minimization problem of (8), we need to solve its discretization:

$$\begin{aligned} &\arg \min_z \left\{ \sum_{i,j=1}^n (z(i, j) + g(i, j)e^{-z(i, j)}) \right. \\ &\quad \left. + \lambda_1 \sum_{i,j=1}^n (z(i, j) - w^{(m-1)}(i, j))^2 \right\}. \end{aligned} \quad (9)$$

Here we denote

$$f(z(i, j)) := z(i, j) + g(i, j)e^{-z(i, j)} + \lambda_1 (z(i, j) - w^{(m-1)}(i, j))^2.$$

Obviously, the solution of (9) is the minimum of the function f . Since f is continuous and derivable within the specified range, this problem is equivalent to solving the regular system with n^2 equations:

$$\begin{aligned} 1 - g(i, j)e^{-z(i, j)} + 2\lambda_1 (z(i, j) - w^{(m-1)}(i, j)) &= 0, \\ i, j &= 1, 2, \dots, n. \end{aligned} \quad (10)$$

The function f is strictly convex for every $z(i, j)$, so the corresponding nonlinear equation has a unique solution and we use the Newton iteration method to find it. The iterative formula of (10) is as follows

$$z^{(m)}(i, j) = z^{(m-1)}(i, j) + \frac{f'(z_1(i, j))}{f''(z_1(i, j))}. \quad (11)$$

Secondly, Using the image $z^{(m)}$ generated by (11) in the previous step, we try to get $w^{(m)}$. Letting

$$F(x, y, w, w_x, w_y) = \lambda_1 \int_{\Omega} (z - w)^2 + \lambda_2 \int_{\Omega} |Dw| \log(e + |Dw|),$$

we get the corresponding Euler-Lagrange equation as follows

$$\frac{\partial}{\partial x} \left(\frac{\partial F}{\partial w_x} \right) + \frac{\partial}{\partial y} \left(\frac{\partial F}{\partial w_y} \right) - \frac{\partial F}{\partial w} = 0, \quad (12)$$

That is,

$$\lambda_2 \operatorname{div} \left(\frac{|Dw| + (e + |Dw|) \log(e + |Dw|)}{(e + |Dw|)|Dw|} Dw \right) + 2\lambda_1(z - w) = 0. \quad (13)$$

Let $t(x) = [x + (e + x) \log(e + x)] / [(e + x)x]$, (13) is simplified as

$$\lambda_2 \operatorname{div}(t(|Dw|)Dw) + 2\lambda_1(z - w) = 0. \quad (14)$$

In this paper, Dw is the derivative of w in distribution sense. If the derivative is continuous, then the gradient at the location (i, j) is $Dw(i, j) = (w(i, j)_x, w(i, j)_y)$, and $|Dw(i, j)| = \sqrt{w(i, j)_x^2 + w(i, j)_y^2}$, $i, j = 1, 2, \dots, n$, where

$$w(i, j)_x = \begin{cases} w(i + 1, j) - w(i, j), & i < n, \\ 0, & i = n. \end{cases}$$

and

$$w(i, j)_y = \begin{cases} w(i, j + 1) - w(i, j), & j < n, \\ 0, & j = n. \end{cases}$$

Now we use the finite differential method to obtain the approximate solution of the equation (14).

Step 1. Let step size $h = 1$, and take samples with constant interval. The pixel value at (i, j) is $w(i, j)$, which is marked as the target pixel. Let w, e, s, n denote its four adjacent pixels as in Figure 1, write $\Lambda = \{(i - 1, j), (i + 1, j), (i, j - 1), (i, j + 1)\}$.

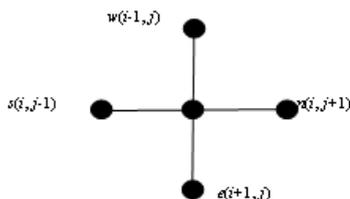


Figure 1. A target pixel and its neighbors.

Step 2. Complete the discretization processing of $\operatorname{div}(t(|Dw|)Dw)$. A small number ε is considered to

avoid a zero divisor, $(e + |Dw|)|Dw|$ in smooth regions, so $|Dw| = \sqrt{w_x^2 + w_y^2} + \varepsilon$, and $t(|Dw|) = t(|Dw|_\varepsilon)$. ε is too small to affect the denoising quality. Let $v = (v^1, v^2) = t(|Dw|_\varepsilon)Dw$, then the curl of $t(|Dw|)$ can be described as

$$\operatorname{div}(t(|Dw|_\varepsilon)Dw) = \frac{\partial v^1}{\partial x} + \frac{\partial v^2}{\partial y} \approx (v_e^1 - v_w^2) + (v_n^2 - v_s^1). \quad (15)$$

where

$$v_e^1 = t(|Dw_e|_\varepsilon) \left[\frac{\partial w}{\partial x} \right]_e \approx t(|Dw_e|_\varepsilon)[w(i + 1, j) - w(i, j)], \quad (16)$$

$$v_w^1 = t(|Dw_w|_\varepsilon) \left[\frac{\partial w}{\partial x} \right]_w \approx t(|Dw_w|_\varepsilon)[w(i + 1, j) - w(i, j)], \quad (17)$$

$$v_n^2 = t(|Dw_n|_\varepsilon) \left[\frac{\partial w}{\partial x} \right]_n \approx t(|Dw_n|_\varepsilon)[w(i + 1, j) - w(i, j)], \quad (18)$$

$$v_s^2 = t(|Dw_s|_\varepsilon) \left[\frac{\partial w}{\partial x} \right]_s \approx t(|Dw_s|_\varepsilon)[w(i + 1, j) - w(i, j)], \quad (19)$$

here

$$t(|Dw_e|_\varepsilon) = \frac{|Dw_e|_\varepsilon + (e + |Dw_e|_\varepsilon) \log(e + |Dw_e|_\varepsilon)}{(e + |Dw_e|_\varepsilon)|Dw_e|_\varepsilon},$$

and

$$|Dw_e|_\varepsilon \approx \sqrt{\varepsilon + (w_e)_x^2 + (w_e)_y^2}.$$

Substituting (16), (17), (18) and (19) into (15), we get

$$\operatorname{div}(t(|Dw|_\varepsilon)Dw) = \sum_{p \in \Lambda} [t(|Dw_p|_\varepsilon)(w(p) - w(i, j))]. \quad (20)$$

Based on (20), the Euler-Lagrange equation (14) can be viewed as

$$\lambda_2 \sum_{p \in \Lambda} [t(|Dw_p|_\varepsilon)(w(p) - w(i, j))] + 2\lambda_1(w - z^{(m)}) = 0. \quad (21)$$

Here we quote the gradient descent method to find $w^{(m)}$.

Letting

$$w.t = \lambda_2 \sum_{p \in \Lambda} [t(|Dw_p|_\varepsilon)(w(p) - w(i, j))] + 2\lambda_1(w - z^{(m)}). \quad (22)$$

we obtain the iterative formula

$$w^{(m)} = w^{(m-1)} + dt \times w.t, \quad (23)$$

where dt is a constant, we take $dt = 0.12$ in this paper.

Step 3. The iterative process of the proposed method stops when the relative difference between $w^{(m)}$ and $w^{(m+1)}$ satisfies the following inequality:

$$\frac{\|w^{(m+1)} - w^{(m)}\|}{\|w^{(m)}\|} \leq 10^{-4}.$$

V. CONVERGENCE ANALYSIS

From (8), we can obtain the following relationship:

$$w^{(m)} = S(R(w^{(m-1)})) = T(w^{(m-1)}), \quad (24)$$

which generates the sequence $\{w^{(m)}\}$. The main aim of this section is to show the convergence of the sequence $\{w^{(m)}\}$.

Theorem 5.1 From any initial value $w^{(0)}$, the iterative sequence $\{w^{(m)}\}$ converges to the optimal solution of the model (7).

Proof. 1) The operator R is nonexpansive, S is 1/2-averaged nonexpansive, and T is also nonexpansive.

By [15], R is nonexpansive. Letting $M = 2S - I$, we have

$$\begin{aligned} & \|M(z_1) - M(z_2)\|_2^2 \\ &= \|(2S - I)(z_1) - (2S - I)(z_2)\|_2^2 \\ &= \|2(S(z_1) - S(z_2)) - (z_1 - z_2)\|_2^2 \\ &= 4\|S(z_1) - S(z_2)\|_2^2 + \|z_1 - z_2\|_2^2 \\ &\quad - 4\langle S(z_1) - S(z_2), z_1 - z_2 \rangle. \end{aligned}$$

Utilizing the second problem of (8), we get

$$\begin{aligned} S(z^{(m)}) &= \arg \min 2\lambda_1 \left[\frac{1}{2} \|z - w^{(m-1)}\|_2^2 \right. \\ &\quad \left. + \frac{\lambda_2}{2\lambda_1} \int_{\Omega} |Dw| \log(e + |Dw|) \right]. \end{aligned} \tag{25}$$

Let $\varphi = \frac{\lambda_2}{2\lambda_1} \int_{\Omega} |Dw| \log(e + |Dw|)$. Through Definition 2.2, S is the proximal operator of φ , and $\varphi \in \Gamma_0(X)$. Thus, we can obtain the following inequality by Theorem 2.2:

$$\|S(z_1) - S(z_2)\|_2^2 \leq \langle S(z_1) - S(z_2), z_1 - z_2 \rangle. \tag{26}$$

Noticing (5.2) and (26), we have

$$\|M(z_1) - M(z_2)\|_2^2 \leq \|z_1 - z_2\|_2^2. \tag{27}$$

So the operator $M = 2S - I$ is nonexpansive. Since the operator S satisfied the equation $S = (1 - 1/2)I + 1/2M$, it is obvious that S is 1/2-averaged nonexpansive. Similarly, we also can get that the operator T is nonexpansive.

2) The series $\sum_{m=1}^{\infty} \|w^{(m-1)} - w^{(m)}\|_2^2$ converges.

Letting

$$J_1(z, w) = \int_{\Omega} \|z - w\|_2^2,$$

and

$$J_2(w) = \int_{\Omega} |Dw| \log(e + |Dw|),$$

we obtain from (7)

$$\begin{aligned} & J(z^{(m)}, w^{(m)}) - J(z^{(m)}, w^{(m+1)}) \\ &= \lambda_1 \left[J_1(z^{(m)}, w^{(m)}) - J_1(z^{(m)}, w^{(m+1)}) \right] \\ &\quad + \lambda_2 \left[J_2(w^{(m)}) - J_2(w^{(m+1)}) \right]. \end{aligned} \tag{28}$$

The Taylor series expansion in the second variable of $J_1(z, w)$ is given

$$\begin{aligned} & J_1(z^{(m)}, w^{(m)}) \\ &= J_1(z^{(m)}, w^{(m+1)}) + (w^{(m)} - w^{(m+1)})^t \frac{\partial J_1}{\partial w}(z^{(m)}, w^{(m+1)}) \\ &\quad + \frac{1}{2} (w^{(m+1)} - w^{(m)})^t \frac{\partial^2 J_1}{\partial w^2}(z^{(m)}, w^{(m+1)}) (w^{(m+1)} - w^{(m)}). \end{aligned} \tag{29}$$

Here we note that J_1 is quadratic in w and x^t denotes a transpose of x . While J_2 is a convex function, so we have

$$\begin{aligned} & J_2(w^{(m)}) \\ &\geq J_2(w^{(m+1)}) + (w^{(m)} - w^{(m+1)})^t \frac{\partial J_2}{\partial w}(w^{(m+1)}). \end{aligned} \tag{30}$$

Combining (28), (29) and (30), we get

$$\begin{aligned} & J(z^{(m)}, w^{(m)}) - J(z^{(m)}, w^{(m+1)}) \\ &\geq (w^{(m+1)} - w^{(m)})^t \left[\lambda_1 \frac{\partial J_1}{\partial w}(z^{(m)}, w^{(m+1)}) \right. \\ &\quad \left. + \lambda_2 \frac{\partial J_2}{\partial w}(w^{(m+1)}) \right] \\ &\quad + \frac{\lambda_1}{2} \|w^{(m+1)} - w^{(m)}\|_2^2. \end{aligned} \tag{31}$$

$w^{(m+1)}$ is the minimizer of $J(z^{(m)}, w)$, so

$$\frac{\partial J}{\partial w}(z^{(m)}, w^{(m+1)}) = 0,$$

and

$$\frac{\partial J_1}{\partial w}(z^{(m)}, w^{(m+1)}) + \frac{\partial J_2}{\partial w}(w^{(m+1)}) = 0,$$

which can be substituted into (31), we obtain

$$\begin{aligned} & J(z^{(m)}, w^{(m)}) - J(z^{(m)}, w^{(m+1)}) \\ &\geq \frac{\lambda_1}{2} \|w^{(m+1)} - w^{(m)}\|_2^2. \end{aligned}$$

The energy in the iteration is decreasing, that is to say, $J(z^{(m+1)}, w^{(m+1)}) \leq J(z^{(m)}, w^{(m+1)})$, so the following inequality

$$\begin{aligned} & J(z^{(m)}, w^{(m)}) - J(z^{(m+1)}, w^{(m+1)}) \\ &\geq \frac{\lambda_1}{2} \|w^{(m+1)} - w^{(m)}\|_2^2. \end{aligned}$$

holds. Hence, the series $\sum_{m=1}^{\infty} \|w^{(m-1)} - w^{(m)}\|_2^2$ is bounded and convergent.

3) The function $J(z, w)$ in (7) is coercive.

Denote L_h, L_v as the one-sided difference matrix on the horizontal direction and the vertical direction, respectively:

$$L = \begin{pmatrix} L_h \\ L_v \end{pmatrix},$$

it is easy to prove that the matrix L is not a full-rank

matrix. The lower bound of the discrete TV is given by

$$\begin{aligned}
 & \int_{\Omega} |Dw| \log(e + |Dw|) \\
 &= \sum_{1 \leq j, k \leq n} |\nabla w| \log(e + |\nabla w|_{j,k}) \\
 &= \sum_{1 \leq j, k \leq n} \sqrt{[(\nabla w)_{j,k}^x]^2 + [(\nabla w)_{j,k}^y]^2} \\
 & \quad \log(e + \sqrt{[(\nabla w)_{j,k}^x]^2 + [(\nabla w)_{j,k}^y]^2}) \\
 &\geq \sum_{1 \leq j, k \leq n} \sqrt{[(\nabla w)_{j,k}^x]^2 + [(\nabla w)_{j,k}^y]^2} \\
 &\geq \frac{\sqrt{2}}{2} \sum_{1 \leq j, k \leq n} (|(\nabla w)_{j,k}^x| + |(\nabla w)_{j,k}^y|) \\
 &= \frac{\sqrt{2}}{2} \|L_w\|_1.
 \end{aligned}$$

Referring to Lemma 3.8 in [14] and Definition 2.5, we can get that $J(z, w)$ is coercive.

4) The set of fixed points of T is nonempty.

The third step have showed that $J(z, w)$ is coercive, which assures that the set of the minimizers of $J(z, w)$ is nonempty. If (z, w) is the minimizer of $J(z, w)$, then we have the following equation

$$w = S(z) = S(R(w)) = T(w). \quad (32)$$

(32) shows that w is the fixed point of the function T . Namely, the set of fixed points of T is nonempty.

5) From any initial value $w^{(0)}$, the iterative sequence $\{w^{(m)}\}$ converges to the optimal solution of the model (7).

Since the fidelity term in (7) is strictly convex, the proposed objective function is also strictly convex. It is obvious that $J(z, w)$ is differential with regard to its first variable z , so the set of fixed points are just minimizers of $J(z, w)$. The fixed point set of T is nonempty. Moreover, from Theorem 2.1, the strict convexity of the function J assures T has a unique fixed point. If we denote the fixed point as w_0 , then $w^{(m)} = T^{(m)}(w_0)$.

We have proved T is nonexpansive in the first step, consequently, both $T^{(m)}(w_0)$ and $\|T^{(m)}(w) - w_0\|_2$ are also nonexpansive. By Definition 2.4, we obtain

$$\begin{aligned}
 & \|T^{(m+1)}(w) - w_0\|_2 \\
 &= \|T(T^{(m)}(w)) - T(w_0)\|_2 \leq \|T^{(m)}(w) - w_0\|_2, \\
 & m = 0, 1, 2, \dots
 \end{aligned}$$

Hence, there exists a nonnegative limit $d(w_0) = \lim_{m \rightarrow \infty} \|T^{(m)}(w) - w_0\|$. As long as we prove that $d(w_0)$ is equal to zero, the conclusion is gotten. Here we apply contradiction method to prove it. Suppose there is a subsequence $\{T^{(m_i)}(w)\}$ in $\{T^{(m)}(w)\}$, whose limit is w' , and $w' \neq w_0$. Because of the asymptotic regularity of T , we get $\lim_{m_i \rightarrow \infty} [(I - T)(T^{(m_i)}(w))] = 0$, then $(I - T)(w) = 0$, i.e., $T(w') = w'$, w' is also the fixed point of T , which is apparently a contradiction. As a result, the unique limit of $\{T^{(m)}(w)\}$ is w_0 .

Letting $w = \lim_{m \rightarrow \infty} w^{(m)}$, we have $R(w) = \lim_{m \rightarrow \infty} R(w^{(m)})$. Since the operator R is nonexpansive, we have

$$\begin{aligned}
 & \|z^{(m)} - R(w)\|_2 \\
 &= \|R(w^{(m-1)}) - R(w)\|_2 \leq \|w^{(m-1)} - w\|_2.
 \end{aligned}$$

It implies that $\lim_{m \rightarrow \infty} (w^{(m-1)} - w) = 0$, and, $\lim_{m \rightarrow \infty} z^{(m)} = R(w)$, the sequence $\{z^{(m)}\}$ also converges to a unique fixed point, so finishes up the proof. ■

VI. NUMERICAL EXPERIMENTS

In this section, numerical results are presented to demonstrate the performance of our proposed algorithm. The results are compared with those models, such as the AA model, the RLO model and the HNW model. For this purpose, it is sufficient and is also more convenient to use the synthetical and commonly-used test images. In our experiments we used two gray original images and two color original images. The two gray original images are the the synthetic image (named as ‘‘Coin’’) and the Lena image (As shown in Figure 2(a) and Figure 2(b)), respectively. The two color images are the original flower image and the original Lena image (As shown in Figure 2(c) and Figure 2(d)), respectively. For the geometry structure, the gray coin image is very simple, and the flower image is slightly more complicated. the Lena images (both gray and color) have nice mixture of details, flat regions, shading area and texture. The four images serve the purpose of our experiments.

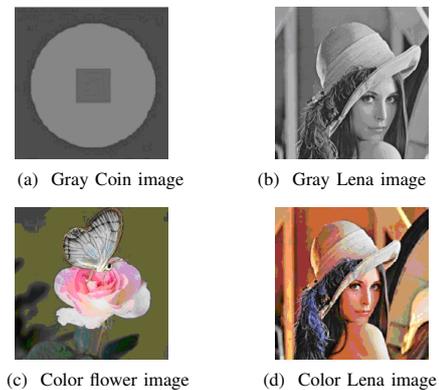


Figure 2. The four original test images.

In the tests, each pixel of the original image is degraded by the gamma noise with the mean value of one (see (4)), and the noise level is controlled by the value of K in the experiments. Obviously the pictures are more noisy with the decrease of the parameter K . The original gray coin image in Figure 2(a) is distorted by the gamma noise with $K = 20$ and $K = 5$ respectively, and the noisy images are shown in Figure 3(a) and Figure 3(b), respectively. The original Lena image in Figure 2(b) is distorted by the gamma noise with $K = 20$ and $K = 5$ respectively and the noisy images are shown in Figure 3(c) and Figure 3(d) respectively.

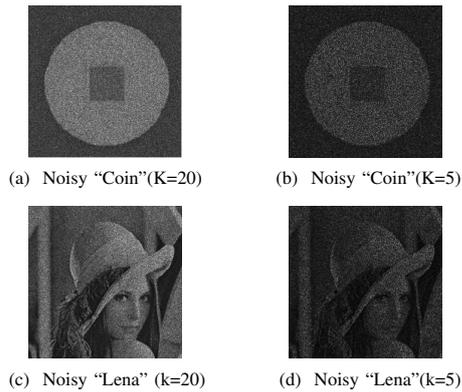


Figure 3. The four noisy images.

The scalar parameters $\lambda_1 = 0.004$ and $\lambda_2 = 0.01$ in our model because our model has the best performance under this case, while the scalar parameters in the AA model, the RLO model and the HNW model are chosen based on [11], [15] and [12], respectively. The solutions of the three models are also computed by discretizing their corresponding gradient descent flow equations with finite difference algorithm respectively. All of the algorithms were implemented in MATLAB2010.

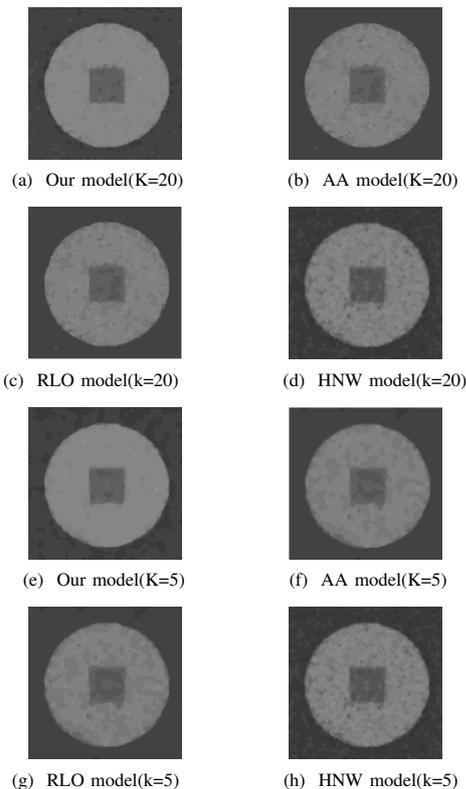


Figure 4. The restoration of "Cion" by different models.

In Figure 4, we show how our model and the other three models behave for the degraded coin image with a gamma noise $K = 20$ and $K = 5$ respectively. From Figure 4(a) to Figure 4(d) (or Figure 4(e) to Figure 4(h)), the images are respectively the ones restored by our model, the AA model, the RLO model and the HNW model. The first

four images in Figure 4 are all from the same degraded coin image in Figure 3(a) with the gamma noise $K = 20$, while the other four images in Figure 4 are all restored from the same degraded coin image in Figure 3(b) with the gamma noise $K = 5$. So does Figure 5.

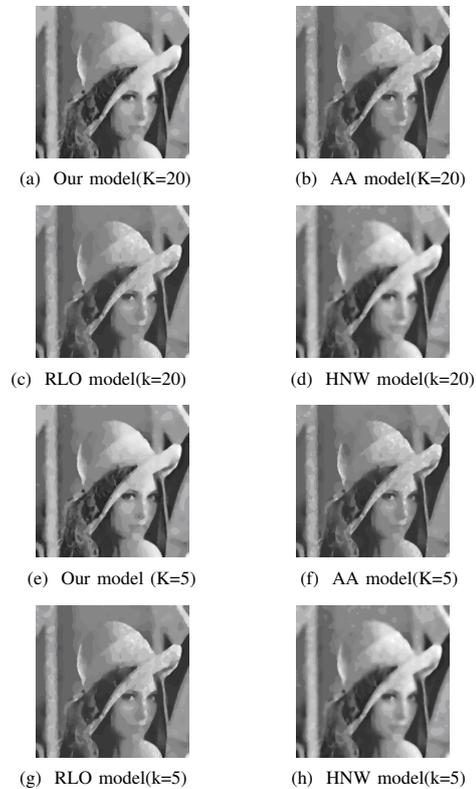


Figure 5. The restoration of "Lena" by different models.

Obviously, the quality of the images restored by our model is the best in the four models. For the Lena image, our experimental results have shown that the quality of the images restored by our model is excellent and our method has the best performance in noise removal. Especially, our model is superior to the other three models for preserving the textures and fine details of the images, but the images restored by one of the other three models contain numerous regions with constant values (named as "stair-casing effect"), as shown in Figure 5. In fact, from Table I to Table IV, we can also get the same conclusion.

In addition to visual examination, we can use a signal-to-noise ratio (abbreviated *SNR*), Peak Signal-to-Noise Ratio (abbreviated *PSNR*), and a relative Error (abbreviated *ReErr*) of the images to assess the quality of the restored images. These three indicators are measures used in science and engineering to quantify how much a signal has been corrupted by noise. For more details, we refer readers to [17], [20], [23]. The larger the value of *SNR* (or *PSNR*), the better the quality of the restored images. Smaller the value of *ReErr*, the better the quality of the restored images. In Table I-III, we compare their restoration results in *SNRs*, *PSNRs* and *ReErrs*. We observe from Table I-III that every index of the restored images by the proposed method is better than that of the

corresponding restored image by one of the others.

TABLE I.
COMPARISON OF THE VALUES OF THE STORED "COIN"
USING THE FOUR METHODS ($K = 20$).

	Our model	AA model	RLO model	HNW model
SNR	22.8184	18.8203	18.8381	19.4729
PSNR	70.5635	62.6710	62.7175	63.8269
ReErr	0.0723	0.1145	0.1143	0.1092

TABLE II.
COMPARISON OF THE VALUES OF THE RESTORED "LENA"
USING THE FOUR METHODS ($K = 20$).

	Our model	AA model	RLO model	HNW model
SNR	22.0574	19.4902	20.2772	21.0580
PSNR	64.5340	59.4585	60.9216	61.4133
ReErr	0.0789	0.1060	0.0969	0.0893

TABLE III.
COMPARISON OF THE VALUES OF THE RESTORED "LENA"
USING THE FOUR METHODS ($K = 5$).

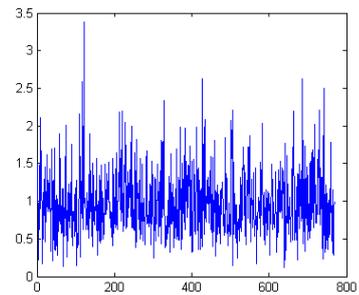
	Our model	AA model	RLO model	HNW model
SNR	13.0624	8.6885	8.0598	11.3559
PSNR	46.6965	39.1902	38.1840	43.6160
ReErr	0.2223	0.3678	0.3954	0.2705

Furthermore, in order to emphasis on the comparison, we show the 101st-103th lines of the original, noisy, and restored Lena images by the AA model, the HNW model and the proposed model in Figure 6-7. In Figure 7, the blue solid lines are the line in Figure 6(b), i.e., the 101st-103th lines of the original Lena image, while the red dotted lines is the 101st-103th lines of the restored Lena images by the corresponding method. It is clear from the figures that the performance of the proposed method is the best in the four models.

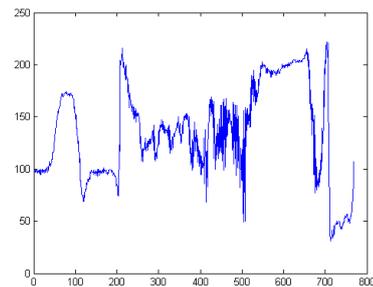
In addition to the quality of the restored gray images, we also find that the proposed algorithm is quite efficient for color images denoising. For the original flower image and the color Lena image (shown in Figure 2(c) and Figure 2(d) respectively), they are distorted by the gamma noise with $K = 20$ and $K = 10$ respectively, and the noisy images are shown in Figure 8(a) and Figure 9(a), respectively. The two degraded images are restored by the HNW model and the proposed model respectively, shown in Figure 8 and Figure 9. As we expect, there is little "stair-casing effect" and less blurring in the restored image by our model, which is smoother and the texture information is preserved better.

Table IV records the *ReErrs* of the results by the HNW model and the proposed model respectively when the value of K is 200, 100, 50, 20 and 10. More precisely, the data in Table IV illustrates that whatever the intensity of the Gamma noise is, all the results of color images restored by the proposed method turn out to approach the original color images more closely than those by the HNW model.

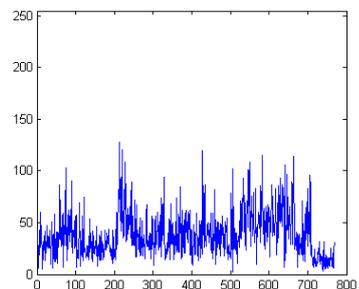
Now, we discuss the denoising efficiency. In [13], Huang, Ng and Wen had found that the HNW model is more efficient than the AA model. Here we compare the



(a) Gamma noise ($K = 5$)



(b) The original Lena images



(c) The noised Lena images

Figure 6. The 101th-103th lines of Gamma noise ($K=5$), the original and noised Lena images.

number of iterations and the computational time required by the proposed method and the HNW model, as showed in Table V. According to it, we find that the efficiency of the proposed method, as well as the denoising quality by the proposed method, also surpasses that of the HNW model.

VII. CONCLUSIONS

In this paper, we have studied a variational method for multiplicative noise removal problems. The proposed method is based on a strictly convexity. The alternating iterative optimization algorithm is implemented, and we have proved that the iteration sequence converges to the optimal solution. Simulation experiments show that the quality of the images restored by our model is excellent. Moreover, the proposed method restrains the "stair-casing

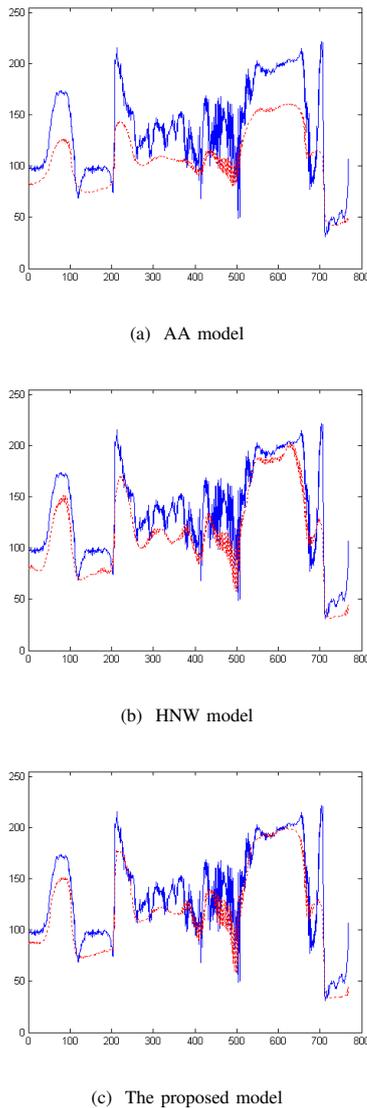


Figure 7. The 101th-103th lines of the restored Lena images by different models.

TABLE IV.
REERRS OF THE RESTORATION RESULTS.

K	200	100	50	20	10
Flower(HNW)	0.0626	0.0675	0.0810	0.1081	0.1116
Flower(Ours)	0.0545	0.0612	0.0781	0.1048	0.1108
Lena(HNW)	0.0810	0.0817	0.1065	0.1259	0.2132
Lena(Ours)	0.0687	0.0795	0.1013	0.1113	0.1716

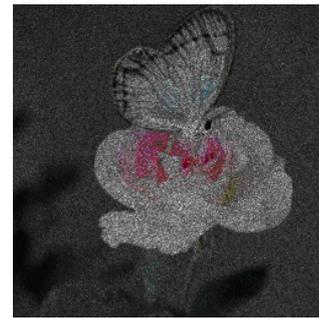
effect” of restored images so effectively that the results are much closer to the original images.

ACKNOWLEDGMENT

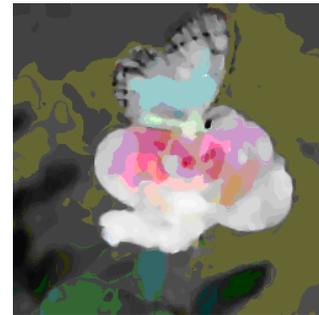
We are grateful to the referees for their valuable comments and suggestions which have led to significant improvement of the paper.

REFERENCES

[1] C. L. Huo, J. Cheng, H. Q. Lu, et al, “Objective-level Change Detection Based on Multiscale Fusion”, *Acta Automatica Sinica*, 2008, 34(3), pp. 251-257.



(a) The degraded image



(b) The restored by HNW model



(c) The restored by our model

Figure 8. Denoising results of “Flower” image by the HNW model and our model ($K = 20$).

[2] L. Wu, Y. A. Hu, “A Survey of Automatic Road Extraction from Remote Sensing Images”, *Acta Automatica Sinica*, 2010, 36(7),pp. 912-922.

[3] Q. Q. Ruan, “Digital Image Processing”. *Electronic Industry Press*, 2007.

[4] J.Peng, Y.Ma, “Integer wavelet image denoising method based on principle component analysis”, *Journal of Software*, 2012, 7(5), pp.982-989.

[5] L. I. Rudin, S. Osher, “E Fatemi. Nonlinear total variation based noise removal algorithms”, *Physica D: Nonlinear Phenomena*, 1992, 60(1),pp. 259-268.

[6] 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 Transactions on Image Processing*, 2003, 12(12),pp. 1579-1590.

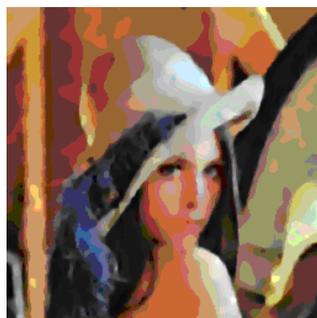
[7] S. Osher, M. Burger, D. Goldfarb, et al, “An iterative regularization method for total variation-based image restora-



(a) The degraded image



(b) The restored by HNW model



(c) The restored by our model

Figure 9. Denoising results of Lena image by HNW model and our model ($K = 10$).

tion”, *Multiscale Modeling and Simulation*, 2005, 4(2), pp. 460-489.

[8] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, “Statistics of speckle in ultrasound B-scans”, *IEEE Trans. Sonics Ultrason*, 1983, 30, pp. 156-163.

[9] W. B. Wang, X. D. Zhang, X. L. Wang, “Speckle Suppression Method in SAR image Based on Curvelet Domain BivaShrink Model”, *Journal of Software*, 2013, 8(4), pp.947-954.

[10] W. K. Allard, “Total variation regularization for image denoising, I. Geometric theory”. *SIAM Journal on Mathematical Analysis*, 2007, 39(4), pp. 1150-1190.

[11] G. Aubert, J. F. Aujol, “A variational approach to removing multiplicative noise”, *SIAM Journal on Applied Mathematics*, 2008, 68(4): 925-946.

[12] Y. M. Huang, M. K. Ng, Y. W. Wen, “A new total variation method for multiplicative noise removal”, *SIAM Journal on Imaging Sciences*, 2009, 2(1), pp. 20-40.

TABLE V.
ITERATION NUMBER AND COMPUTATIONAL TIME

Images	Iteration number		Computational time (s)	
	HNW	Ours	HNW	Ours
Fig.6.2(a)	158	103	12.9531	10.0000
Fig.6.2(b)	321	231	23.5469	20.2344
Fig.6.2(c)	208	194	21.5625	12.8903
Fig.6.2(d)	377	313	32.0156	23.5469
Fig.6.7(a)	221	218	40.3906	33.125
Fig.6.8(a)	507	496	75.7188	69.0781

[13] Z. Jin, X. Yang, “A variational model to remove the multiplicative noise in ultrasound images”, *Journal of Mathematical Imaging and Vision*, 2011, 39(1), pp. 62-74.

[14] J. Shi, S. Osher, “A nonlinear inverse scale space method for a convex multiplicative noise model”, *SIAM Journal on Imaging Sciences*, 2008, 1(3), pp. 294-321.

[15] L. Rudin, P. L. Lions, S. Osher, “Multiplicative denoising and deblurring: theory and algorithms”, *Geometric Level Set Methods in Imaging, Vision, and Graphics*, Springer New York, 2003, pp. 103-119.

[16] X. D. Wang, X. C. Feng, L. G. Huo, “Iteratively Reweighted Anisotropic-TV Based Multiplicative Noise Removal Model”, *Acta Automatica Sinica*, 2012, 38(3), pp. 444-451.

[17] X. G. Hu, L. T. Zhang, W. Jiang, “Improved variational model to remove multiplicative noise based on partial differential equation (in Chinese)”, *Journal of Computer Applications*, 2012, 32(7), pp. 1879-1881.

[18] H. Y. Zhang, Q. C. Peng, “PDE Deduction and Its Numerical Realization in Variational Image Restoration”, *Computer Engineering and Science*, 2006, 28(6), pp. 44-46.

[19] J. Shen, T. F. Chan, “Mathematical models for local nontexture inpaintings”, *SIAM Journal on Applied Mathematics*, 2002, 62(3), pp. 1019-1043.

[20] J. J. Moreau, “Fonctions convexes duales et points proximaux dans un espace hilbertien (in French)”, *CR Acad. Sci. Paris*, 1962, 255, pp. 2897-2899.

[21] P. L. Combettes, V. R. Wajs, “Signal recovery by proximal forward-backward splitting”, *Multiscale Modeling and Simulation*, 2005, 4(4), pp. 1168-1200.

[22] Z. Opial, “Weak convergence of the sequence of successive approximations for nonexpansive mappings”, *Bull. Amer. Math. Soc.*, 1967, 73(4), pp. 591-597.

[23] K. Fang, L. M. Shen, X. Y. Xu, et al, “Convexity Conditions for Parameterized Surfaces”, *Journal of Software*, 2012, 7(6), pp. 1219-1226.

[24] D. P. Bertsekas, A. Nedić, A. E. Ozdaglar, “Convex analysis and optimization”, *Belmont: Athena Scientific*, 2003.

Xuegang Hu received his B.S. degree in applied mathematics from China West Normal University, China, in 1988, and his M.S. and Ph.D. degree in applied mathematics from Sichuan University, China, in 1995 and 2006, respectively. He is a professor at Chongqing University of Posts and Telecommunications, China. His research interests include digital image processing and analysis, partial differential equations and their applications.

Yuefang Lou is currently working towards her M.S. degree in computer software and theory at Chongqing University of Posts and Telecommunications, China. Her current research interest includes digital image processing and analysis.

Tianqiao Hu is currently a college student and working towards her B.S. degree at Sichuan University, China. Her current research interest includes image processing and applied cryptography.