

IMAGE DENOISING BASED ON THE SURFACE FITTING STRATEGY

ZHI-FENG PANG AND BAOLI SHI

Abstract. In this work, we propose a surface fitting strategy based on a two-step model to remove noise from digital images. In the first step, we minimize the total variation energy functional of an image by using the projection gradient method in order to obtain the dual variable as the smoothed normal vector. In the second step, we try to find a surface as the recovered image to fit the smoothed normal vector. Based on the projection gradient method and the variable splitting method, we propose an efficient numerical method to solve this two-step model and also give the convergence analysis of the proposed method. Some numerical comparisons are given to validate the effectiveness of our proposed model.

Key words. Two-step model, Staircase effect, Projection gradient method, Edge indicator function

1.. Introduction

The existence of noise is inevitable in the course of obtaining images. It may be introduced in many different ways, such as image formation processing, image recording and image transmission. These random distortions make it difficult to carry out any required picture processing. Therefore, noise removal is an important and challenging problem in image restoration.

In the view of mathematics, the denoising problem can be expressed as follows: Assume that $g : \Omega \subset R^2 \rightarrow R$ denotes a noisy image and u denotes the desired clean image, it follows that

$$g = u + \eta,$$

where η is the additive noise. The aim is to recover the true image u from g .

In practice we want to preserve image edges and features while removing noise for the image denoising problem. Many researchers have devoted their efforts to this study, see [4, 15]. Wherein many variational models have been proposed to eliminate noise and to preserve edges and the small scale characteristics at the same time. The total variation (TV) minimization, as a classical variational model, was first introduced by Rudin, Osher and Fatemi (called the ROF model) in [27] as the following form

$$(1.1) \quad \min_u \int_{\Omega} |Du| + \frac{\lambda}{2} \|u - g\|_{L^2(\Omega)}^2,$$

where the regularization parameter $\lambda > 0$. It has been demonstrated to be very successful in edge-preserving for image restoration problem, see [10, 14, 18, 19, 20, 22, 27, 28] and references therein. However, the ROF model has the undesirable staircase effect since the smooth regions of the restored image are transformed into the piecewise constants. To overcome this deficiency, some higher-order PDEs [14, 21, 23, 29] have been proposed during the last few years. It has been proved

Received by the editors September 7, 2011 and, in revised form, January 25, 2013.
2000 *Mathematics Subject Classification.* 94A08.

that the higher-order PDEs can effectively alleviate the undesirable staircase effect such as the model [21] with the form

$$(1.2) \quad \min_u \int_{\Omega} |D^2 u| + \frac{\zeta}{2} \|u - g\|_{L^2(\Omega)}^2,$$

where the tuning parameter $\zeta > 0$. However, the higher-order PDEs are generally complicated to be implemented and can not preserve the edges well.

Recently, a two-step method has been proposed by Lysaker-Osher-Tai (called the LOT model) [22] to reduce staircase effect while preserving important features, such as edges and textures. This two-step model can be denoted as the following two steps:

- In the first step, they minimized the total variation of the unit normal vector \mathbf{n} by the following scheme

$$(1.3) \quad \min_{|\mathbf{n}|=1} \int_{\Omega} |\nabla \mathbf{n}| dx + \frac{\alpha}{2} \|\mathbf{n} - \mathbf{n}_0\|_{L^2(\Omega)}^2,$$

where $\mathbf{n}_0 = \frac{\nabla g}{\|\nabla g\|_2}$ and the regularization parameter $\alpha > 0$.

- In the second step, in order to find a surface to fit the above smoothed normal vector \mathbf{n} , they considered

$$(1.4) \quad \min_v \int_{\Omega} (|\nabla v| - (\nabla v)^T \cdot \mathbf{n}) dx + \frac{\beta}{2} \|v - g\|_{L^2(\Omega)}^2,$$

where the regularization parameter $\beta > 0$.

In the view of the numerical implementation, the proposed algorithms of the LOT model in [22] are complicated and slow because of computing three nonlinear second-order PDEs. Additionally, when the information about noise is not known, this model can not preserve edges or textures well. Therefore, an improved LOT model was suggested in [18]. By letting $\mathbf{n} = (\cos \theta, \sin \theta)^T$ and $\mathbf{n}_0 = (\cos \theta_0, \sin \theta_0)^T$, they used the relationship $|\nabla \mathbf{n}| = |\nabla \theta|$ to transform the minimization problem (1.3) into the following form

$$\min_{\theta} \int_{\Omega} |\nabla \theta| dx + \alpha \int_{\Omega} (1 - \cos(\theta - \theta_0)) dx.$$

Based on the fact that

$$(1.5) \quad 1 - \cos(\theta - \theta_0) = 2 \sin^2 \left(\frac{\theta - \theta_0}{2} \right) \sim \frac{(\theta - \theta_0)^2}{2},$$

i.e. $1 - \cos(\theta - \theta_0)$ is the equivalent infinitesimal quantity of $\frac{(\theta - \theta_0)^2}{2}$, they then solved the following problem

$$(1.6) \quad \min_{\theta} \int_{\Omega} |\nabla \theta| dx + \frac{\alpha}{2} \|\theta - \theta_0\|_{L^2(\Omega)}^2$$

in the first step, where θ and θ_0 are the polar angle of \mathbf{n} and \mathbf{n}_0 respectively. But only when $(\theta - \theta_0) \rightarrow 0$, the formula (1.5) comes into the existence. Actually, the unit normal vector \mathbf{n} of the restoration image u and \mathbf{n}_0 defined in the problem (1.3) are impossible to be very close each other. Thus, the scheme substituting $1 - \cos(\theta - \theta_0)$ by $\frac{(\theta - \theta_0)^2}{2}$ is not convincing in [18]. In the second step, they introduced an edge indicator function $I(x)$ and adopted the L^1 norm for the image fidelity term. Then this step is given by

$$(1.7) \quad \min_v \int_{\Omega} I(x) (|\nabla v| - (\nabla v)^T \cdot \mathbf{n}) dx + \gamma \|v - g\|_{L^1(\Omega)},$$

where $\gamma > 0$ is a constant, $\mathbf{n} := (n_1, n_2)^T = (\cos \theta, \sin \theta)^T$ obtained from the first step (1.6) and $(\nabla v)^T \cdot \mathbf{n} = v_x n_1 + v_y n_2$.

Our model is closely related to the work in [18, 22, 25, 26], its key is based on the observation that the dual variable generated by using the duality method to solve the ROF model [10] can be regarded as the unitization for the gradient of the restoration image. In detail, in the first step, we minimize the ROF model by using the projection gradient method in order to obtain the dual variable as the smoothed normal vector. This course is very easy to be implemented since there is only one variable to be considered as did in [26]. Furthermore, this step is easier to identify whether the choosing regularization parameter is suitable or not. Moreover, unlike the second step in [26], here we try to find the restoration image by using L^1 norm fitting term and the edge indicator function $I(x)$. This notion is inspired by the observation that the edge indicator function tends to zero at steep gradients like edges and goes to one at the smooth regions. Thus the edges can be found and retained during restoration process. Simultaneously, adopting the L^1 norm can also keep contrast better [1, 2, 3, 12, 24]. Here we give the convergence analysis of our proposed algorithm under some assumptions. Some numerical examples are given to validate the effectiveness of our proposed model and algorithm.

This paper is organized as follows. In subsection 2.1, we introduce the two-step model for noise removal. In subsection 2.2, an algorithm and some related results of our proposed model are given to show how we obtain the restoration image. In section 3, some numerical comparisons are given to illustrate the effectiveness of our proposed model. Some concluding remarks are given in section 4.

2.. The Two-step Model

In this section, we first introduce the two-step model in subsection 2.1 and then give an efficient numerical method to solve it in subsection 2.2. The convergence analysis is also arranged in subsection 2.2.

2.1.. The two-step model.

Definition 2.1. Let $\Omega \subset \mathbb{R}^2$ be a bounded open subset and $u \in L^1(\Omega)$. Then the weighted total variation of u is

$$TV_{\zeta(x)}(u) = \int_{\Omega} \zeta(x) |Du| = \sup_{\mathbf{p}(x) \in \phi_{\zeta(x)}} \left\{ \int_{\Omega} u(x) \operatorname{div} \mathbf{p}(x) dx \right\},$$

where

$$\phi_{\zeta(x)} := \{ \mathbf{p} = (p^1, p^2)^T \in C_c^1(\Omega; \mathbb{R}^2) \mid \|\mathbf{p}\|_{\infty} \leq \zeta \}.$$

The space BV can now be defined as

$$BV(\Omega) = \{ u \in L^1(\Omega) \mid TV_1(u) < +\infty \}.$$

Equipped with the norm $\|u\|_{BV(\Omega)} = \int_{\Omega} |Du| dx + \|u\|_{L^1(\Omega)}$ is a Banach space.

Without loss of generality, we set $TV_1(u)$ as $TV(u)$ in this paper. If $u \in W^{1,1}(\Omega)$, then $\int_{\Omega} |Du| = \int_{\Omega} |\nabla u| dx$. Therefore, the problem (1.1) can be formally written as

$$(2.1) \quad \min_u \int_{\Omega} |\nabla u| dx + \frac{\lambda}{2} \|u - g\|_{L^2(\Omega)}^2,$$

which Euler-Lagrange equation can be written as

$$(2.2) \quad u = g + \frac{1}{\lambda} \operatorname{div} \frac{\nabla u}{|\nabla u|}$$

with the Neumann boundary condition. Formally, since the restoration image u can be regarded as the smoothed image, we can set $\mathbf{n} := \frac{\nabla u}{|\nabla u|}$ to replace the input vector in the second step of the LOT model (1.4). From the numerical point of view, the total variation is not straightforward to be minimized, since it is not differentiable in zero. In order to overcome this non-differentiability, an idea from duality was proposed by Chambolle [11] to obtain the restoration image

$$(2.3) \quad u = g - \frac{1}{\lambda} \operatorname{div} \mathbf{p},$$

where the variable \mathbf{p} is closely related to Definition 2.1.

By comparing (2.2) with (2.3), we can transform the computation of the flow field \mathbf{n} in the first step of the original LOT model [22] into the computation of dual variable \mathbf{p} in the first step of our proposed two-step model. In fact, this conversion is easy to be implemented since there is only one variable to be considered. Here we solve the problem (2.1) by using the projection gradient method as did in [5].

Once the smoothed flow field \mathbf{n} is obtained, we want to find a surface as the restoration image to match it. In the second step, we thus minimize

$$(2.4) \quad \min_v \int_{\Omega} I(x) (|\nabla v| - (\nabla v)^T \cdot \mathbf{n}) dx + \gamma |v - g|_{L^1(\Omega)}$$

to get the restoration image v , where the regularization parameter $\gamma > 0$. In practical, the weighted function $I(x)$ can be chosen as

$$I(x) = \frac{1}{1 + \varrho |\nabla(G_{\sigma} * g)|^2} \quad \text{or} \quad I(x) = \exp(-k |\nabla(G_{\sigma} * g)|^2),$$

where $G_{\sigma}(x) = \frac{1}{2\pi\sigma^2} \exp(-\frac{x^2}{2\sigma^2})$ is a Gaussian kernel with a scale parameter σ . We choose the left form and set $\sigma = 0.4$ in this paper.

In the following, we will explain the above two-step model. The difference between (1.4) and (2.4) is the introduction of the edge indicator function $I(x)$ and the replacement of L^2 norm by L^1 norm. These modifications have two important impacts. At the fine structures like edges where $|\nabla G_{\sigma} * g|$ is large, $I(x)$ tends to zero. Hence, the first term in (2.4) tends to zero, only the fidelity term is remained. This allows the discontinuities across hyper surfaces in the reconstruction process. As a result, the edges can be perfectly found and retained. However, in the smooth regions where $|\nabla(G_{\sigma} * g)|$ is small, $I(x)$ goes to one, then the problem (2.4) is similar to the problem (1.4). Thus, the staircase effect in the smooth regions can be taken away effectively.

The L^1 norm as a measure of the fidelity term has been introduced and well studied in [1, 2, 3, 12, 24]. It has been proved to outperform the L^2 norm in some applications. The differences between the ROF model and the $TV - L^1$ model have been studied by Chan and Esedoğlu in [12]. They indicated that the L^1 norm as a fidelity term represents the contrast better than the L^2 norm.

2.2.. Algorithm and Its Convergence. In this subsection, we state some convergence results and propose an algorithm to solve our proposed model. According to the related work in [11], we have the following result for the ROF model (2.1), which closely relates to the first step of our proposed model.

Theorem 2.1. *When $\tau < \frac{1}{4}$, the sequence $\{u^n\}$ generated by*

$$\begin{cases} u^k = g - \lambda \operatorname{div} \mathbf{p}^k \\ \mathbf{p}^{k+1} = \mathcal{P}_{\mathcal{K}}(\mathbf{p}^k - \frac{\tau}{\lambda} \nabla u^k) \end{cases}$$

converges to the solution u^* of the ROF model (2.1), where \mathcal{K} is defined by

$$\mathcal{K} := \left\{ \mathbf{p} = (p^1, p^2)^T : \operatorname{div} \mathbf{p} \in L^2(\Omega), \|\mathbf{p}\|_\infty \leq 1 \text{ with } |\mathbf{p}| = \sqrt{(p^1)^2 + (p^2)^2} \right\}$$

and $\mathcal{P}_{\mathcal{K}}(\cdot)$ denotes the orthogonal projection operator on the convex set \mathcal{K} .

In fact this iteration method corresponds to the specific case of Bermudez-Moreno method [8, 5]. Since we can directly solve the projection problem on the closed convex set for each iteration, this iteration method is more efficient and convenient than the gradient descent method used in [11].

In the second step, we rewrite the problem (2.4) as

$$(2.5) \quad \min_v \int_{\Omega} (I(x)|\nabla v| - (\nabla v)^T \cdot (I(x)\mathbf{n})) \, dx + \gamma|v - g|_{L^1(\Omega)}.$$

Obviously, this model is difficult to be solved directly since there include two L^1 -norm terms. One of efficient methods first introduces an equality constraint $u_1 = v$ and then transforms to solve an unconstrained optimization problem based on the penalty method as

$$(2.6) \quad \min_{v, u_1} \int_{\Omega} (I(x)|\nabla v| - (\nabla u_1)^T \cdot (I(x)\mathbf{n})) \, dx + \frac{1}{2\delta} \|v - u_1\|_{L^2(\Omega)}^2 + \gamma|u_1 - g|_{L^1(\Omega)},$$

where $\delta > 0$ is the penalty parameter. This method was widely used in the image restoration problem such as the work in [6, 7, 10, 13, 9]. Obviously, the problem (2.6) is convex, so we can perform this minimization efficiently by alternately solving v and u_1 . Specifically, this course can be written as:

(i) u_1 being fixed, we search for v as a solution of

$$(2.7) \quad \min_v \int_{\Omega} I(x)|\nabla v| \, dx + \frac{1}{2\delta} \|v - u_1\|_{L^2(\Omega)}^2;$$

(ii) v being fixed, we search for u_1 as a solution of

$$(2.8) \quad \min_{u_1} \frac{1}{2\delta} \|v - u_1\|_{L^2(\Omega)}^2 + \gamma|u_1 - g|_{L^1(\Omega)} - \int_{\Omega} (\nabla u_1)^T \cdot (I(x)\mathbf{n}) \, dx.$$

Remark 2.1. *There are some remarks for solving above subproblems (2.7) and (2.8).*

- *The subproblem (2.7) is only similar to the ROF model (2.1) formally since there includes a weighted function $I(x)$. Usually the solution of (2.7) can be given by*

$$v = u_1 - \delta \operatorname{div} \mathbf{q},$$

where $\mathbf{q} = (q^1, q^2)^T$ satisfies that

$$-\nabla(\delta \operatorname{div} \mathbf{q} - u_1) + \frac{1}{I(x)} |\nabla(\delta \operatorname{div} \mathbf{q} - u_1)| \mathbf{q} = 0.$$

The above variable \mathbf{q} can be obtained by using semi-implicit gradient descent (or fixed point) algorithm as

$$(2.9) \quad \mathbf{q}^{n+1} = \frac{\mathbf{q}^n + t \left(\nabla \left(\operatorname{div} \mathbf{q}^n - \frac{u_1}{\delta} \right) \right)}{1 + \frac{t}{I(x)} \left| \nabla \left(\operatorname{div} \mathbf{q}^n - \frac{u_1}{\delta} \right) \right|}.$$

for choosing the original value $\mathbf{q}^0 = \mathbf{0}$. The generated sequence $\{\mathbf{q}^n\}$ is convergent when $t \leq \frac{1}{8}$, we refer to the literature [9] for the proof.

- For the subproblem (2.8), it can be equivalently rewritten as

$$\min_{u_1} \frac{1}{2\delta} \|u_1 - (v - \operatorname{div}(I(x)\mathbf{n}))\|_{L^2(\Omega)}^2 + \gamma|u_1 - g|_{L^1(\Omega)}.$$

It obviously corresponds to the classical shrinkage operator problem [17], so we can get the closed form solution as

$$(2.10) u_1 = \max(|v - g - \delta \operatorname{div}(I(x)\mathbf{n})| - \delta\gamma, 0) \frac{v - g - \delta \operatorname{div}(I(x)\mathbf{n})}{|v - g - \delta \operatorname{div}(I(x)\mathbf{n})|} + g.$$

In view of the above analysis and results, we summarize the following algorithm to solve our proposed two-step model.

Algorithm 2.1. *The algorithm for solving our proposed two-step model.*

- **Step 1:** *The projection gradient method for solving (2.1).*

- (1) Initialize: $u^0 = g$ and $\mathbf{p}^0 = \mathbf{0}$;
- (2) Compute \mathbf{p}^{k+1} by

$$\mathbf{p}^{k+1} = \frac{\mathbf{p}^k + \tau(\nabla(\operatorname{div}\mathbf{p}^k - \frac{g}{\lambda}))}{\max(1, |\mathbf{p}^k + \tau(\nabla(\operatorname{div}\mathbf{p}^k - \frac{g}{\lambda}))|)};$$

- (3) When the stopping criterion is satisfied, output the dual variable $\mathbf{p} := \mathbf{p}^{k+1}$ and let $u = g - \lambda \operatorname{div}\mathbf{p}$ as the restoration image by using the ROF model.

- **Step 2:** *The gradient descent algorithm for solving (2.6).*

- (I) Initialize: $\mathbf{q}^0 = \mathbf{0}$ and $u_1^0 = g$. Set $\mathbf{n} = -\mathbf{p}$;
- (II) Compute $(\mathbf{q}^{n+1}, v^{n+1}, u_1^{n+1})$ by

$$\begin{cases} \mathbf{q}^{n+1} = \frac{\mathbf{q}^n + t(\nabla(\operatorname{div}\mathbf{q}^n - \frac{u_1^n}{\delta}))}{1 + \frac{t}{I(x)} |\nabla(\operatorname{div}\mathbf{q}^n - \frac{u_1^n}{\delta})|}, \\ v^{n+1} = u_1^n - \delta \operatorname{div}\mathbf{q}^{n+1}, \\ u_1^{n+1} = \max(|v^{n+1} - g - \delta \operatorname{div}(I(x)\mathbf{n})| - \delta\gamma, 0) \frac{v^{n+1} - g - \delta \operatorname{div}(I(x)\mathbf{n})}{|v^{n+1} - g - \delta \operatorname{div}(I(x)\mathbf{n})|} + g; \end{cases}$$

- (III) When the stopping criterion is satisfied, output the variable v^{n+1} as the restoration image v .

In the following, we show the uniqueness of the solution for the problem (2.6) and also prove the convergence of the sequence $\{(v^n, u_1^n)\}$ generated by Step 2 in Algorithm 2.1. Denote the functional in (2.6) as

$$E(v, u_1) := |I(x)\nabla v|_{L^1(\Omega)} - \int_{\Omega} (\nabla u_1)^T \cdot (I(x)\mathbf{n}) dx + \frac{1}{2\delta} \|v - u_1\|_{L^2(\Omega)}^2 + \gamma|u_1 - g|_{L^1(\Omega)}.$$

Theorem 2.2. *Assume that the sequence $\{(v^n, u_1^n)\}$ is established in Step 2 of Algorithm 2.1 and $\int_{\Omega} ((\nabla(v^n - g))^T \cdot (I(x)\mathbf{n})) dx$ is bounded when $|v^{n+1} - g - \delta \operatorname{div}(I(x)\mathbf{n})| \leq \delta\gamma$. If set $u_1 \in BV(\Omega) \cap C^1(\Omega)$, then there exists in a unique solution of the problem (2.6). Furthermore, the sequence $\{(v^n, u_1^n)\}$ converges to this unique solution.*

Proof. For the problem (2.6), the existence of the solution can be directly obtained from the convexity and coercivity of the functional E . By using the similar approach as Theorem 2 in [7], we can obtain the uniqueness of its solution.

When solving the successive problems $\min_v E(v, \cdot)$ and $\min_u E(\cdot, u)$ in the problem (2.6), we can get

$$(2.11) \quad E(v^n, u_1^n) \geq E(v^{n+1}, u_1^n) \geq E(v^{n+1}, u_1^{n+1}).$$

Therefore, the sequence $\{E(v^n, u_1^n)\}$ is nonincreasing. Based on the fact that $|\mathbf{n}| = 1$, it is easy to find that $|I(x)\nabla v|_{L^1} \geq \int_{\Omega} (\nabla v)^T \cdot (I(x)\mathbf{n}) dx$ and

$$\begin{aligned} & E(v^{n+1}, u_1^{n+1}) \\ & \geq \int_{\Omega} (\nabla(v^{n+1} - u_1^{n+1}))^T \cdot (I(x)\mathbf{n}) dx + \frac{1}{2\delta} \|v^{n+1} - u_1^{n+1}\|_{L^2(\Omega)}^2 + \gamma |u_1^{n+1} - g|_{L^1(\Omega)}. \end{aligned}$$

Following from the third equation in Step 2 of Algorithm 2.1, we can find that

$$\int_{\Omega} (\nabla(v^{n+1} - u_1^{n+1}))^T \cdot (I(x)\mathbf{n}) dx = - \int_{\Omega} \delta(\operatorname{div}(I(x)\mathbf{n}))^2 dx$$

when $|v^{n+1} - g - \delta \operatorname{div}(I(x)\mathbf{n})| > \delta\gamma$. Then, using the assumption that $\int_{\Omega} ((\nabla(v^n - g))^T \cdot (I(x)\mathbf{n})) dx$ is bounded when $|v^{n+1} - g - \delta \operatorname{div}(I(x)\mathbf{n})| \leq \delta\gamma$, we can deduce that the sequence $E(v^n, u_1^n)$ is bounded. Thus the sequence $\{E(v^n, u_1^n)\}$ is convergent. Denote $m = \lim_{n \rightarrow \infty} E(v^n, u_1^n)$. In the following, we want to show that

$$m = \min_{v, u_1} E(v, u_1).$$

Since the functional E is coercive and the sequence $\{E(v^n, u_1^n)\}$ is convergent, we can deduce that $\{(v^n, u_1^n)\}$ is bounded. We can thus extract a subsequence $\{(v^{n_k}, u_1^{n_k})\}$ which converges to (\hat{v}, \hat{u}_1) as $n_k \rightarrow +\infty$. Moreover, for all n_k , v and u_1 , we have

$$(2.12) \quad E(v^{n_k+1}, u_1^{n_k}) \leq E(v, u_1^{n_k}) \quad \text{and} \quad E(v^{n_k}, u_1^{n_k}) \leq E(v^{n_k}, u_1).$$

Let \tilde{v} be a cluster point of the sequence $\{v^{n_k+1}\}$. Taking (2.11) into account, it is easy to find that

$$(2.13) \quad M = E(\tilde{v}, \hat{u}_1) = E(\hat{v}, \hat{u}_1).$$

Taking limits on both sides of $v^{n+1} = u_1^n - \delta \operatorname{div} q^{n+1}$, we obtain that \tilde{v} is the solution of the following problem

$$\min_u \left\{ |I(x)\nabla v|_{L^1(\Omega)} + \frac{1}{2\delta} \|v - \hat{u}_1\|_{L^2(\Omega)}^2 \right\}.$$

On the other hand, it follows from (2.13) that

$$|I(x)\nabla \tilde{v}|_{L^1(\Omega)} + \frac{1}{2\delta} \|\tilde{v} - \hat{u}_1\|_{L^2(\Omega)}^2 = |I(x)\nabla \hat{v}|_{L^1(\Omega)} + \frac{1}{2\delta} \|\hat{v} - \hat{u}_1\|_{L^2(\Omega)}^2.$$

Based on the uniqueness of the solution for the problem (2.7), we deduce that $\tilde{v} = \hat{v}$ and then $v^{n_k+1} \rightarrow \hat{v}$ as $n_k + 1 \rightarrow +\infty$. By passing to the limits in (2.12), we get

$$E(\hat{v}, \hat{u}_1) \leq E(v, \hat{u}_1) \quad \text{and} \quad E(\hat{v}, \hat{u}_1) \leq E(\hat{v}, u_1).$$

The above formulas can be rewritten as

$$E(\hat{v}, \hat{u}_1) = \min_v E(v, \hat{u}_1) \quad \text{and} \quad E(\hat{v}, \hat{u}_1) = \min_{u_1} E(\hat{v}, u_1).$$

From the definition of $E(v, u_1)$, we can easily obtain the equivalent form corresponding to above two problems

$$\begin{cases} 0 \in \frac{1}{\delta} (\hat{v} - \hat{u}_1) + \partial(|I(x)\nabla \hat{v}|), \\ 0 \in \frac{1}{\delta} (\hat{u}_1 - \hat{v}) + \operatorname{div}(I(x)\mathbf{n}) + \gamma \partial(|\hat{u}_1 - g|). \end{cases}$$

Since the subdifferential of E at (\hat{v}, \hat{u}_1) is given by

$$\partial E(\hat{v}, \hat{u}_1) = \left(\begin{array}{c} \frac{1}{\delta} (\hat{v} - \hat{u}_1) + \partial(|I(x)\nabla \hat{v}|) \\ \frac{1}{\delta} (\hat{u}_1 - \hat{v}) + \operatorname{div}(I(x)\mathbf{n}) + \gamma \partial(|\hat{u}_1 - g|) \end{array} \right),$$

it follows that

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \partial E(\widehat{v}, \widehat{u}_1),$$

which is equivalent to $E(\widehat{v}, \widehat{u}_1) = \min_{v, u_1} E(v, u_1) = M$. Hence, the sequence $\{E(v^n, u_1^n)\}$ converges to M . By the uniqueness of the solution for the minimization problem (2.6), we can deduce that the sequence $\{(v^n, u_1^n)\}$ converges to $(\widehat{v}, \widehat{u}_1)$ as $n \rightarrow +\infty$. \square

3.. Numerical Experiments

In this section, we will give some experimental results to compare our proposed model with the ROF model (1.1), the higher-order model (1.2), the LOT model (1.3)-(1.4) and other two-step models proposed in [18, 26]. For the ROF model, the restoration image can be directly obtained by *Step 1* of Algorithm 2.1. Actually, this course is the projection gradient method so that it is more efficient and faster than the original semi-implicit gradient method in [5]. Similarly, we also use this method to solve the higher-order model (1.2) as follows

$$\begin{cases} \xi^{k+1} = \frac{\xi^k + \nu(\nabla^2(\operatorname{div}^2 \xi^k - \frac{g}{\xi}))}{\max(1, |\xi^k + \nu(\nabla^2(\operatorname{div}^2 \xi^k - \frac{g}{\xi}))|)} \\ u^{k+1} = g - \zeta \operatorname{div}^2 \xi^{k+1} \end{cases}$$

for choosing some suitable values $u^0 = g$ and $\xi^0 = \mathbf{0}$. Formally, we can get the convergence of the above algorithm when $\nu \leq \frac{1}{32}$. In order to estimate the improvement of our proposed method compared with other models, we give the definitions of the signal to noise ratio (*SNR*) and the mean square error (*MSE*). For a clean image w and its noisy observation w_0 , the noise can be denoted as $\eta = w_0 - w$. Then *SNR* and *MSE* are defined by

$$SNR = 10 \log_{10} \left(\|w_0 - \bar{w}_0\|_{L^2(\Omega)}^2 / \|\eta - \bar{\eta}\|_{L^2(\Omega)}^2 \right) \quad \text{and} \quad MSE = \|w - w_0\|_{L^2(\Omega)}^2 / |\Omega|,$$

where $|\Omega| = \int_{\Omega} dx$, $\bar{w}_0 = \int_{\Omega} w_0 dx / |\Omega|$ and $\bar{\eta} = \int_{\Omega} \eta dx / |\Omega|$. It is not difficult to find that the higher *SNR* and the lower *MSE*, the better quality of the restoration.

All numerical results are generated on Matlab 7.12 and the related testing images are shown in Figure 1. The first image is a synthetic image, which only includes simple geometric structures such as smoothing regions and boundary regions. The last two images include more complex structures than the synthetic image. For choices of the regularization parameters in models, we first set the noisy images as the format of 'mat' in order to fix the inputting images and then tune them carefully by a large number of experiments. When solving the ROF model (1.1) and the higher-order model (1.2), we set $t = \frac{1}{4}$ and $\nu = \frac{1}{32}$. Since the first step of the model in [26] can be directly obtained from the ROF model (1.1), we only set the regularization parameter α_2 in the second step for this model.

Example 3.1. *In this example, we consider the synthetic image shown in Figure 1(a), where the noisy image with $SNR = 11.2740$ and $MSE = 256.5994$ is contaminated by the white Gaussian noise with the standard deviation $\sigma = 16$. The related data and parameters are arranged in Table 1. In order to understand the restoration results of these models, we not only plot a slice curve of restoration images shown in Figure 3 which corresponding station is shown in Figure 1(a) but also give the differences between the related restoration images and the original noisy image shown in Figure 4.*

From the restoration images shown in Figure 2, it is clear to see that much of the noise is suppressed. As expected, the ROF model gives rise to the undesirable



FIGURE 1. The related original images.

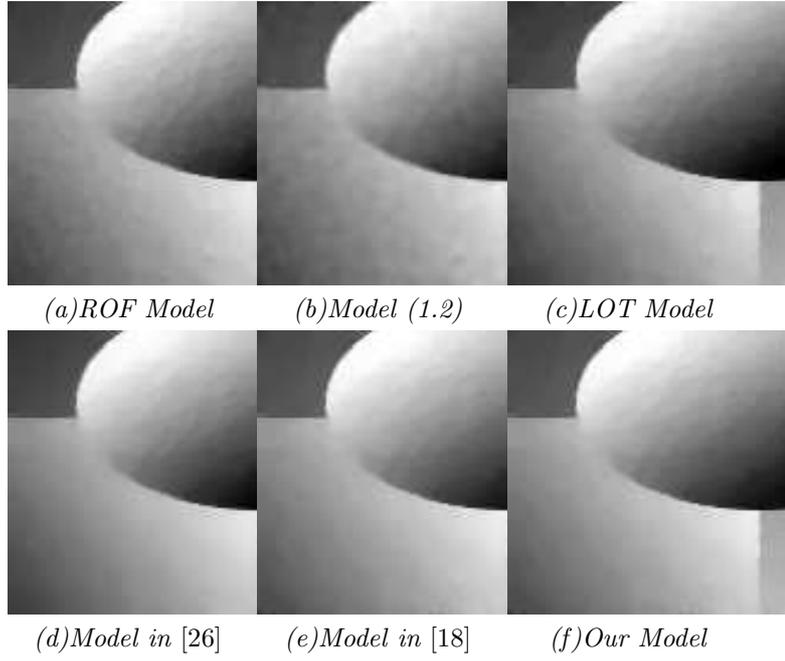


FIGURE 2. The restoration images in Example 3.1

TABLE 1. The related data in Example 3.1.

<i>Model</i>	<i>Iterations</i>	<i>Related parameters</i>	<i>SNR</i>	<i>MSE</i>
<i>ROF</i>	<i>100/0</i>	$\lambda = 17.5$	<i>22.3558</i>	<i>17.9605</i>
<i>(1.2)</i>	<i>100/0</i>	$\zeta = 8$	<i>21.6274</i>	<i>21.7448</i>
<i>LOT</i>	<i>150/150</i>	$\alpha = 20, \beta = 200$	<i>23.0540</i>	<i>15.5923</i>
<i>[26]</i>	<i>100/100</i>	$\alpha_2 = 12$	<i>22.8539</i>	<i>15.7980</i>
<i>[18]</i>	<i>150/250</i>	$\alpha = 6.667, \gamma = 0.05, \delta = 0.1$	<i>23.2376</i>	<i>14.8113</i>
<i>Our</i>	<i>100/250</i>	$\gamma = 0.001, \delta = 0.1$	<i>23.2261</i>	<i>14.8450</i>

staircase effect and the model (2.4) suffers from edge blurring. However, the two-step models (including the LOT model (1.3)-(1.4), the models proposed in [18, 26] and our proposed model) can efficiently suppress these drawbacks. Actually we can

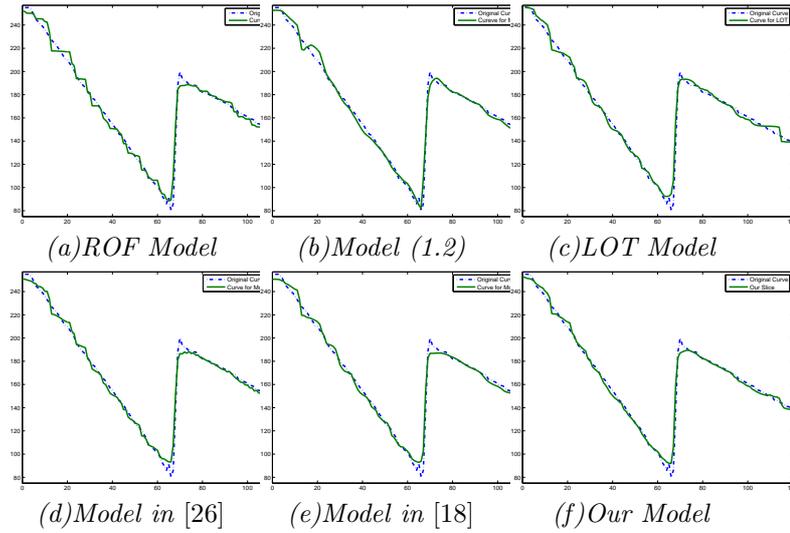


FIGURE 3. The comparison of related slice curves between the restoration images and the original image in Example 3.1.

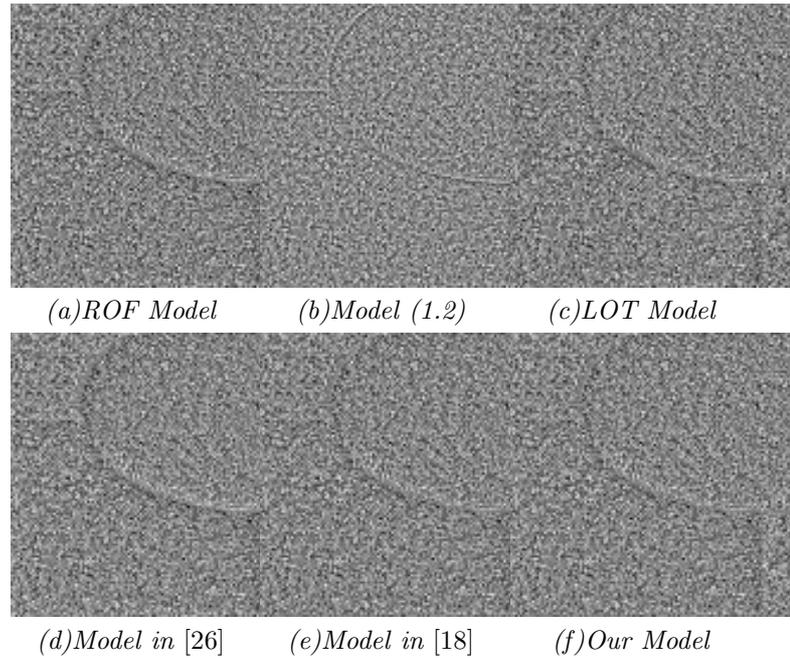


FIGURE 4. The differences between the restoration images and the noisy image in Example 3.1.

also observe these advantages of two step models from the slice curves shown in Figure 3. Furthermore, the images restored by two step models are more natural than other models. Simultaneously, the differences shown in Figure 3 imply that the noisy image can be more efficiently restored by using the two-step models. Additionally, the values of SNR and MSE shown in Table 1 also illustrate that the

two-step models outperform the ROF Model (1.1) and the higher-order model (1.2). Simultaneously, it's worth noting that our proposed model almost obtain the similar restoration result as did by the model in [18].

Example 3.2. We take the well-known Lena image as the testing image shown in Figure 1(b). The noisy image is added to the white Gaussian noise with the standard

TABLE 2. The related data in Example 3.2.

Model	Iterations	Related parameters	SNR	MSE
ROF	100/0	$\lambda = 5.8$	18.8108	34.9009
(1.2)	100/0	$\zeta = 2.75$	19.1163	32.9968
LOT	500/150	$\alpha = 0.2, \beta = 0.2$	18.7634	35.7674
[26]	100/8	$\alpha_2 = 0.45$	18.8182	34.7885
[18]	150/210	$\alpha = 2, \gamma = 0.0005, \delta = 0.03$	19.0093	33.6215
Our	100/210	$\gamma = 0.0002, \delta = 0.03$	19.0080	33.6288

deviation $\sigma = 10$. Before processing, the SNR and MSE of the noisy image are 14.5094 and 100.4604. The related data and parameters are shown in Table 2. In order to understand the related models, especially in regions with smooth signals or discontinuities, we consider a part of shoulder regions of the restoration images in Figure 5 and plot them in Figure 6. It is easy to find that the ROF model transforms smooth regions into piecewise constant regions and the model (1.2) leads to edges blurring. However, all of the two-step models can alleviate these effects in related regions. Furthermore, our proposed model and the model in [18] are more efficient than other two-step models [22, 26].



FIGURE 5. The related restoration image in Example 3.2.

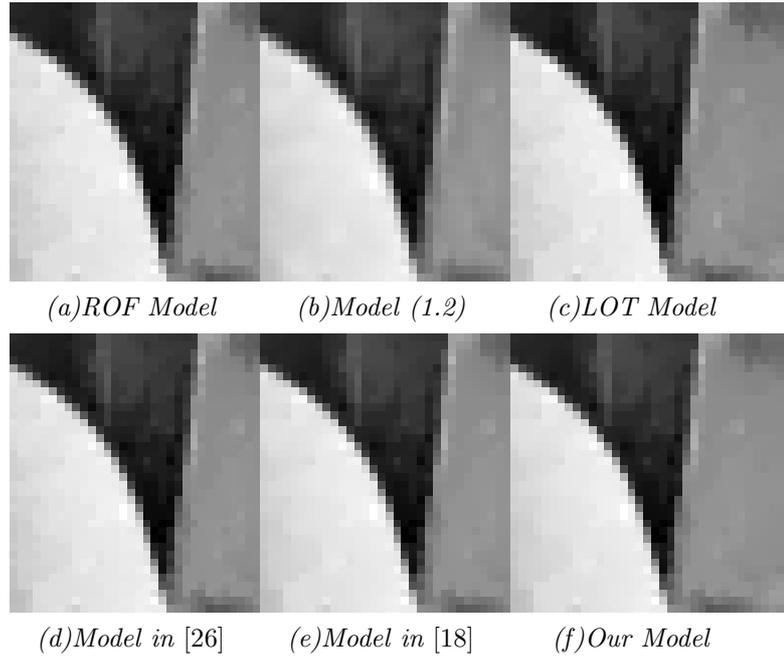


FIGURE 6. A part of the restoration image in Example 3.2.

Example 3.3. Since the Barbara image includes many high and low intensity textures, we choose it as the testing image in order to compare the restored ability of aforementioned models. Here the noisy image is contaminated by the white Gaussian noise with the standard deviation $\sigma = 10$. Usually, the contour plot of the image help us to evaluate the restoration while revealing level curves, thus we only

TABLE 3. The related data in Example 3.3.

Model	Iterations	Related parameters	SNR	MSE
ROF	100/0	$\lambda = 2.75$	15.5120	73.8138
(1.2)	100/0	$\zeta = 1.1$	15.5959	73.2808
LOT	150/150	$\alpha = 4, \beta = 0.5$	15.4910	75.2175
[26]	100/100	$\alpha_2 = 1.35$	15.5122	73.8908
[18]	150/210	$\alpha = 2, \gamma = 0.0005, \delta = 0.015$	15.5391	73.4760
Our	100/210	$\gamma = 0.0001, \delta = 0.015$	15.5376	73.4956

plot contours of related images shown in Figure 7(a-f). From contours, we can deduce that two-step models except the LOT model can give more smooth surfaces and uniformly separated smooth level lines. Especially, the contours generated by the model (1.2) and the LOT model are looser than those restored by other models. Furthermore, the data in Table 3 again imply that the LOT model gives the worst restoration image. However, it's worth noting that the model (1.2) seemly gives the best restoration image based on the values of SNR and MSE. Actually, it is due to the fact that the textures such as the scarf includes much more smooth structures, which is advantageous to use the higher-order model.

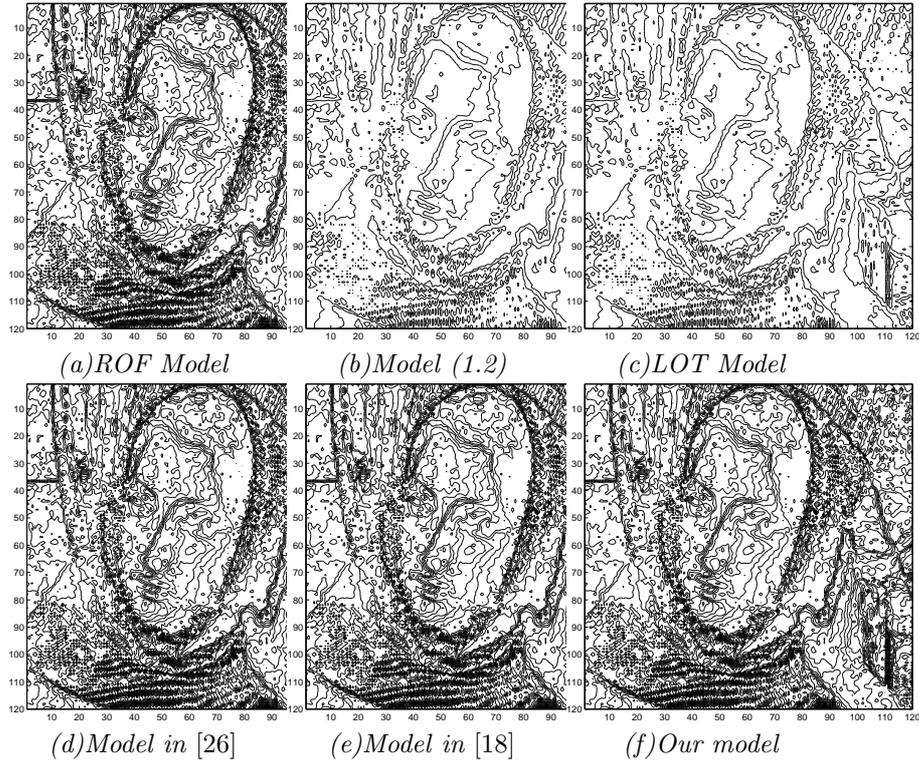


FIGURE 7. The contour plots correspond to the related restoration images in Example 3.3.

4.. Conclusion

In this paper, a two-step model and the corresponding algorithm are proposed for noise removal. In the first step, we use a projection gradient method to solve the ROF model in order to get the dual variable, which can be regarded as the smoothed unit normal vector \mathbf{n} . In the second step, we try to find a surface as the restoration image to fit the smoothed normal vector \mathbf{n} . Furthermore, in order to preserve the edges well and keep contrast better, we introduce the edge indicator function $I(x)$ and employ the L^1 norm for the fidelity term in the second step. Numerical experiments indicate that our proposed method can preserve the edges well and alleviate the staircase effect successfully. Its efficiency is similar to the modified model in [18], but our proposed model is easier to be implemented and choose suitable parameters in the first step. Simultaneously, it's worth noting that the restoration images extremely depend on the edge indicator function $I(x)$. In the future research, we hope to study how to choose a more robust edge indicator function $I(x)$ and extend our model to restore the vector-value images.

Acknowledgments

The first author would like to thank the Mathematical Imaging and Vision Group in Division of Mathematical Sciences at Nanyang Technological University in Singapore for the hospitality during the visit. The authors also appreciate Prof. Xue-Cheng Tai, Li-Lian Wang and Yu-Fei Yang for many discussions. This research is supported by Singapore MOE Grant T207B2202, Singapore NRF2007IDM-IDM002-010, NSF of China(U1304610) and University Research Fund of Henan university (N0.2012YBZR029).

References

- [1] S. Alliney. A property of the minimum vectors of a regularizing functional defined by means of the absolute norm. *IEEE Transactions on Signal Processing*, 45(4):913-917, 1997.
- [2] S. Alliney. Digital filter as absolute norm regularizers. *IEEE Transactions on Signal Processing*, 40(6):1548-1562, 1992.
- [3] S. Alliney. Recursive median filter of increasing order: a variational approach, *IEEE Transactions on Signal Processing*, 44(6):1346-1354, 1996.
- [4] G. Aubert and P. Kornprobst. *Mathematical Problem in Image Processing: partial differential equations and the calculus of variations*. Applied Mathematical Science, 147, 2nd edn, Springer, 2008.
- [5] J. F. Aujol. Some first-order algorithms for total variation based image restoration. *Journal of Mathematical Imaging and Vision*, 34(3):307-327, 2009.
- [6] J. F. Aujol, A. Chambolle, Dual norms and image decomposition models. *International Journal of Computer Vision*, 63(1):85-104, 2005.
- [7] J. F. Aujol, G. Gilboa, T. F. Chan, and S. Osher. Structure-texture image decomposition-modeling. algorithms, and parameter selection, *International Journal of Computer Vision*, 67(1):111-136, 2006.
- [8] A. Bermudez and C. Moreno. Duality methods for solving variational inequalities. *Computers & Mathematics with Applications*, 7(1):43-58, 1981.
- [9] X. Bresson, S. Esedoğlu, P. Vanderheynt, J. P. Thiran, and S. Osher. Fast global minimization of the active contour/ snake model. *Journal of Mathematical Imaging and Vision*, 28(2):151-167, 2007.
- [10] A. Chambolle. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1-2):89-97, 2004.
- [11] A. Chambolle. Total variation minimization and a class of binary MRF models. In *EMM-CVPR 05*, volume 3757 of *Lecture Notes in Computer Sciences*, 136-152, 2005
- [12] T. F. Chan and S. Esedoğlu. Aspects of total variation regularized L^1 function approximation. *SIAM Journal on Applied Mathematics*, 65(5):1817-1837, 2005.
- [13] T. F. Chan, G. H. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM Journal on Scientific Computing*, 20(6):1964-1977, 1999.
- [14] T. F. Chan, A. Marquina, and P. Mulet. Higher-order total variation-based images restoration. *SIAM Journal on Scientific Computing*, 22(2):503-516, 2000.
- [15] T. F. Chan and J. Shen. *Image Processing and Analysis-Variational, PDE, Wavelet, and Stochastic Methods*. SIAM, Philadelphia, 2005.
- [16] H. Z. Chen, J. P. Song, and X. C. Tai. A dual algorithm for minimization of the LLT model. *Advances in Computational Mathematics*, 31(1-3):115-130, 2009.
- [17] D. Donoho. De-Noising by Soft-Thresholding. *IEEE Transactions on Information Theory*, 41(33):613-627, 1995.
- [18] F. F. Dong, Z. Liu, D. X. Kong and K. F. Liu. An improved LOT model for image restoration. *Journal of Mathematical Imaging and Vision*, 34(1):89-97, 2009.
- [19] T. Goldstein and S. Osher. The split Bregman methods for L^1 regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323-343, 2009.
- [20] K. Joo and S. Kim. PDE-based image restoration, I: Anti-staircasing and anti-diffusion. <http://www.ms.uky.edu/math/MAREport/PDF/2003-07.pdf>, 2003.
- [21] M. Lysaker, A. Lundervold and 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*, 12(12):1579-1590, 2003.
- [22] M. Lysaker, S. Osher, and X. C. Tai. Noise removal using smoothed normals and surface fitting. *IEEE Transactions on Image Processing*, 13(10):1345-1357, 2004.

- [23] F. Li, C. M. Shen, J. S. Fan, and C. L. Shen. Image restoration combining a total variational filter and a fourth-order filter. *Journal of Visual Communication and Image Representation*, 18(4):322-330, 2007.
- [24] M. Nikolova. Minimiziers of cost-functions involving nonsmooth data-fidelity terms: Application to the processing of outliers. *SIAM Journal on Numerical Analysis*, 40(3):965-994, 2002.
- [25] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *SIAM Multiscale Modeling and Simulation*, 4(2):460-489, 2005.
- [26] Z. F. Pang and Y. F. Yang. A two-step model for image denoising using a duality strategy and surface fitting. *Journal of Computational and Applied Mathematics*, 235(1):82-90, 2010.
- [27] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259-268, 1992.
- [28] T. Rahman, X. C. Tai, and S. Osher. A TV-stokes denoising algorithm. In *Scale Space and Variational Methods in Computer Vision*, Lecture Notes in Computer Science 4485, Springer Berlin / Heidelberg, 473-483, 2007.
- [29] Y. L. You and M. Kaveh. Four-order partial differential equation for noise removal. *IEEE Transactions on Image Processing*, 9(10):1723-1730, 2000.

College of Mathematics and Information Science, Henan University, Kaifeng, 475004, Henan, China.

E-mail: zhifengpang@163.com and blshi@hotmail.com