Image restoration based on the minimized surface regularization

Zhi-Feng Pang[†,‡], Li-Zhen Guo[†], Yuping Duan[♮] and Jian Lu[§]

[†] *College of Mathematics and Statistics, Henan University, Kaifeng, 475004, China.*

[‡] *Institute of Applied Mathematics, Henan University, Kaifeng, 475004, China.*

[♮] *Center for Applied Mathematics, Tianjing University, Tianjin, 300072, China.*

[§] *College of Mathematics and Statistics, Shenzhen University, Shenzhen, 518060, China.*

Abstract

We propose an image restoration model based on the minimized surface regularization. The proposed model closely relates to the classical smoothing ROF model [28]. We deduce two different conjugation forms via coupling the gradient operator with the smoothing parameter α or not, and then provide the existence of the minimizer in the continuous setting. In order to efficiently solving the proposed model, we employ the primal dual method by reformulating the proposed model as a min-max problem. Relying on the convex conjugation, the convergence of the algorithm is provided as well. Numerical implementations mainly emphasize the effectiveness of the proposed method by comparing it to other two well-known methods in terms of the CPU time and restoration quality.

Keywords: Minimized surface regularization; Smoothing ROF model; Image restoration; Convex conjugate; Primal dual method.

1. Introduction

Image restoration is one of the most fundamental and important problems in low-level image processing, which is the operation to recover (as good as possible) the clean image $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$, from a contaminated image $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ as

$$f = Ku + \eta,$$

where K is linear degraded operator (e.g. blur operator) and η is an additive noise. An ideal restoration model is expected to enhance image by reducing degradations and preserving edges as much as possible. However, it is often difficult to simultaneously remove the noise and enhance edges because both the noise and edges are high frequency signals.

During the past several decades, the models based on variational partial differential equation (PDE) have been attracted much attention such as TV-based models [28, 36, 9, 33, 19] and nonlocal-based models [5, 13] and also obtained some satisfactory results [2, 8, 29]. Different to aforementioned models, which are developed based on the

Email address: zhifengpang@163.com, lizhenguo99@126.com, doveduan@gmail.com, jianlu@szu.edu.cn (Zhi-Feng Pang[†,‡], Li-Zhen Guo[†], Yuping Duan[♮] and Jian Lu[§])

image domain, the authors in [31, 18, 37] proposed to consider the image as an embedded surface $\mathcal{M} \in \mathbb{R}^3$ denoted by

$$\Omega \to \mathcal{M} : \mathrm{x} \to \mathcal{U}(\mathrm{x}),$$

where $\mathrm{x} := (x_1, x_2)$ denotes the local coordinates of the surface and $\mathcal{U}(x_1, x_2) := (x_1, x_2, u(x_1, x_2))$. Note that Ω and $\mathcal{M}$ are viewed as the Riemannian manifold equipped with suitable metrics. By introducing metrics $\mathrm{d}^2 s = \alpha \mathrm{d}^2 x_1 + \alpha \mathrm{d}^2 x_2$ on Ω and $\mathrm{d}^2 \tilde{s} = \alpha \mathrm{d}^2 x_1 + \alpha \mathrm{d}^2 x_2 + \mathrm{d}^2 u$ on $\mathcal{M}$, we can obtain

$$\mathrm{d}^2 \tilde{\mathrm{s}} = (\mathrm{d}x_1, \mathrm{d}x_2) \begin{bmatrix} \alpha + u_{x_1}^2 & u_{x_1} u_{x_2} \\ u_{x_1} u_{x_2} & \alpha + u_{x_2}^2 \end{bmatrix} \begin{bmatrix} \mathrm{d}x_1 \\ \mathrm{d}x_2 \end{bmatrix}, \tag{1}$$

where $\alpha > 0$ is a shrinkage parameter for the local coordinates $(\mathrm{d}x_1, \mathrm{d}x_2)$. Here $\mathrm{d}^2(\cdot)$ denotes $(\mathrm{d}(\cdot))^2$ with the convention. In order to obtain a restored approximation u from f, we need to search for $\mathcal{M}$ with the minimal area. In this way, singularities are smoothed. Let g denote the determinant of the second-order square matrix in (1). We consider to minimize u as follows

$$\min_u \ \mathcal{J}_\alpha(\nabla u) := \int_\Omega \sqrt{g} \mathrm{dx} = \sqrt{\alpha} \int_\Omega \sqrt{\alpha + |\nabla u|^2} \mathrm{dx}. \tag{2}$$

With the assumption of the suitable boundary condition, the Euler-Lagrange equation of the minimized problem (2) can be deduced as

$$E_\alpha(u) = 0, \tag{3}$$

where

$$E_\alpha(u) := \mathrm{div}\left(\frac{\nabla u}{\sqrt{|\nabla u|^2 + \alpha}} \right).$$

It is clear that the mean curvature of $\mathcal{M}$ is zero when $\alpha = 1$. Surfaces of zero mean curvature are known as minimal surfaces. Thus, we can solve the equation (3) by embedding it into the following dynamical scheme

$$\frac{\mathrm{d}\mathcal{X}}{\mathrm{dt}}(t) = E_\alpha(u),$$

where $\mathcal{X}(t) = (x_1, x_2, u(t, x_1, x_2))$. However, this scheme only considers how to regularize the image while ignoring to preserve the image features. For example, if we use a (flat) ellipse as its boundary condition, then the minimization problem extends to a flat surface as showing in the first row of Figure 1. On the other hand, using the different boundary conditions can obtain the same results shown in 2nd-3rd rows of Figure 1. However, the image restoration problem usually pursues the restored image similar to the real image as much as possible. So it is essential for adding some prior information to the proposed model. Therefore, we propose a novel model by introducing a data fitting term as follows

$$\min_u \ \mathcal{J}(u) := \frac{\lambda}{2} \|Ku - f\|_2^2 + \int_\Omega \sqrt{\alpha + |\nabla u|^2} \mathrm{dx}, \tag{4}$$

where $\lambda > 0$ is a positive parameter and $\|\cdot\|_2$ denotes the L^2-norm. Obviously, the objective functional in the model (4) is strictly convex, so it has an unique solution.

Actually, the proposed model (4) reduces to the well-known ROF model proposed by Rudin, Osher and Fatemi (ROF model) [28] when $\alpha = 0$. On the other hand, when $\alpha > 0$ in the model (4), we introduce extra smoothness into the Total Variation (TV) model. Furthermore, we give two basic remarks on the proposed model (4):

① For (locally) small values of $|\nabla u|$, we can obtain

$$\sqrt{\alpha + |\nabla u|^2} \approx \frac{1}{2\sqrt{\alpha}} (2\alpha + |\nabla u|^2)$$

based on the Taylor expansion as $h(\tilde{t}) = \sqrt{1 + \tilde{t}^2} \approx 1 + \frac{\tilde{t}^2}{2}$ if $\tilde{t}$ is very small. This implies that the model (4) tends to a classical Gaussian filtering process to damp down the small perturbations of the gradient while integrating the coefficient $\frac{1}{2\sqrt{\alpha}}$ into the regularization parameter λ in (4).

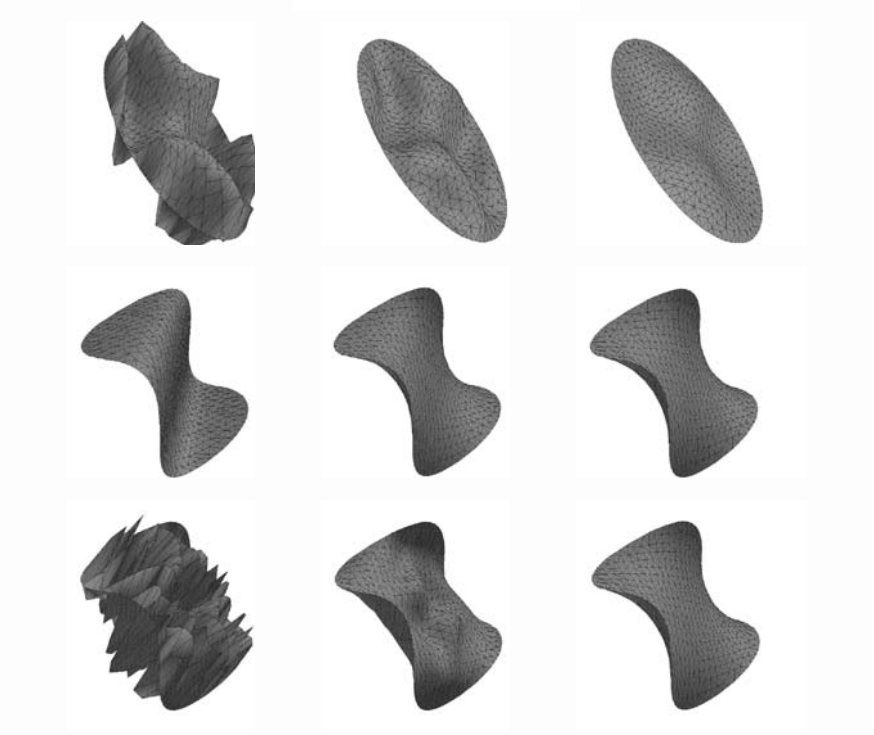


Figure 1. Using different boundary conditions to the equation (3) can obtain different evolution images. Here we set $\alpha = 1$. First Column: Original Image; Second Column: Middle Image; Second Column: Final Image.

② For large values of $|\nabla u|$, the following relationship

$$\sqrt{\alpha + |\nabla u|^2} \approx |\nabla u|$$

exists and the model (4) reduces to the ROF model. Therefore, the model (4) is expected to preserve edges as the ROF model [28] while restoring degraded image.

These facts imply that choosing a smaller α compels the model (4) to keep these qualities.

In numerical implementations, since the model (4) relates to the smoothing ROF model, we then can employ some classical numerical methods such as the time marching scheme [28] and the fixed point iteration scheme [35] to solve it. These methods are usually restricted to the Courant-Friedrichs-Lewy (CFL) condition [32] and the data scale of the operator inversion. Different to the aforementioned works, we here propose two different primal-dual methods to solve the model (4) via choosing different conjugation schemes. Specifically, one of schemes dividually regards variables ∇u and α, another scheme regards them as a whole when using the conjugation theory in the regularization term. To the best of our knowledge, these methods have not been used to solve the model (4) although it has been verified on some non-smoothing models in the field of image processing and machine learning. Moreover, we firstly employ the Legendre-Fenchel transformation to reformulate the minimization problem (4) as a saddle-point problem and use the alternative updating scheme to solve the primal and dual variables. Since the primal dual method can avoid solving the nonlinear PDE directly as both primal and dual variables have closed form solution, it can avoid the numerical difficulties when working solely with the primal variable as the aforementioned methods [28, 35] to solve the model (3). Furthermore, we can classify the proposed method into the general framework of the primal dual method, which means the convergence of the algorithm can be obtained as well. We use numerical experiments to demonstrate the proposed primal-dual algorithms can achieve the solutions with satisfactory accuracy and in a reasonable time.

The contents of the paper are arranged as follows. In Section 2, we discuss some properties of the proposed model (4) such as the conjugation and existence of the solution. Based on the Section 3, we deduce two primal

dual methods by defining different functional spaces for the regularized term. We describe the numerical algorithms based on the discrete form of the proposed model (4) and the numerical methods in Section 4. Some numerical comparisons between our proposed algorithms and other related classic algorithms are arranged in Section 5. We give some concluding remarks in Section 6.

2. Some properties of the model (4)

In this section, we will consider some properties of the proposed model (4). In order to deduce the conjugate form for the regularization term in the model (4), we first introduce the simplified version of a theorem that allows to rephrase the optimization of an integrating functional as the pointwise optimization of its integrand [27].

Lemma 2.1. *Let $X = L^2(\Omega, \mathbb{R}^m)$ and $F : \Omega \times \mathbb{R}^m$ be a normal integrand , i. e., a function whose epigraphical mapping $E(x) = \mathrm{epi}F(x, \cdot)$ is closed-valued and measurable. Then, the integral functional $\Upsilon_F : X \to \mathbb{R} \cup \{+\infty\}$:*

$$\Upsilon_F(v) := \int_\Omega F(x, v(x))dx$$

has the property that minimization and integration can be exchanged, i. e.,

$$\inf_{v \in X} \Upsilon_F(v) = \int_\Omega \left[\inf_{z \in \mathbb{R}^m} F(x, z) \right] dx$$

as long as $\Upsilon_F \neq \infty$ on X. Likewise, unless this common value is $-\infty$, the minimizers of the functional satisfy

$$\bar{v} \in \arg\inf_{v \in X} \Upsilon_F(v) \Leftrightarrow \bar{v}(x) \in \arg\inf_{z \in X} F(x, z) \text{ for almost all } x \in \Omega.$$

Obviously, the above result is also right for solving the sup-problem. In addition, the condition on the integrands to be normal is rather straightforward to check

$$R(\boldsymbol{s}) := \int_\Omega \sqrt{|\boldsymbol{s}|^2 + \alpha}\mathrm{dx},$$

where $|\boldsymbol{s}| = \sqrt{s_1^2 + s_2^2}$. According to Example 14.29 in [27], $R(\boldsymbol{s})$ is the Carathéordory integrand, which means that are measurable in the first variable and continuous in the second, are normal integrands.

Theorem 2.1. *The convex conjugation of $R(\boldsymbol{s})$ is given by*

$$R^*(\boldsymbol{t}) = \int_\Omega r^*(\boldsymbol{t})dx \tag{5}$$

where $\boldsymbol{t} = (t_1, t_2)^T \in L^2(\Omega \times \Omega)$ and

$$r^*(\boldsymbol{t}) = \begin{cases} -\sqrt{\alpha}\sqrt{1 - |\boldsymbol{t}|^2} & \text{if } |\boldsymbol{t}| \leq 1, \\ +\infty & \text{if } |\boldsymbol{t}| > 1. \end{cases}$$

Proof. Based on the definition of convex conjugate and Lemma 2.1, we can get

$$R^*(\boldsymbol{t}) = \sup_{\boldsymbol{s} \in \mathbb{R}^2} [\langle \boldsymbol{s}, \boldsymbol{t} \rangle_X - R(\boldsymbol{s})] = \int_\Omega \sup_{\boldsymbol{s} \in \mathbb{R}^2} \left[\boldsymbol{s}^T \boldsymbol{t} - \sqrt{|\boldsymbol{s}|^2 + \alpha} \right] \mathrm{dx}.$$

The optimization of the integrated can be performed for each point $x \in \Omega$. So we can transform to consider

$$r^*(\boldsymbol{t}) = \sup_{\boldsymbol{s} \in \mathbb{R}^2} \left[\boldsymbol{s}^T \boldsymbol{t} - \sqrt{|\boldsymbol{s}|^2 + \alpha} \right].$$

By introducing the polar coordinates, this problem can be rewritten as

$$\begin{aligned} r^*(\boldsymbol{t}) &= \sup_{\rho\geq 0}\sup_{\theta\in[0,2\pi]}\left[\langle(\rho\cos\theta,\rho\sin\theta),\boldsymbol{t}\rangle - \sqrt{\rho^2+\alpha}\right] \\ &= \sup_{\rho\geq 0}\left[\rho|\boldsymbol{t}| - \sqrt{\rho^2+\alpha}\right]. \end{aligned}$$

Set $h(\rho) = \rho|\boldsymbol{t}| - \sqrt{\rho^2+\alpha}$, we can differentiate it as

$$h'(\rho) = |\boldsymbol{t}| - \frac{\rho}{\sqrt{\rho^2+\alpha}} \text{ and } h''(\rho) = -\frac{\alpha}{(\rho^2+\alpha)^{\frac{3}{2}}}.$$

Then it is easy to find that $h'(\rho) > 0$ when $|\boldsymbol{t}| > 1$, which implies that is increasing. So $r^*(\boldsymbol{t}) = +\infty$ when $\rho \to \infty$. For the case $|\boldsymbol{t}| = 1$, we can get $r^*(\boldsymbol{t}) = \lim_{\rho\to\infty} h(\rho) = 0$. When $|\boldsymbol{t}| < 1$, we can set $h'(\rho) = 0$ and then we have the critical value as $\bar{\rho} := \frac{|\boldsymbol{t}|\sqrt{\alpha}}{\sqrt{1-|\boldsymbol{t}|^2}}$. Using the strictly concave based on $h''(\rho) < 0$, we can deduce that $r^*(\boldsymbol{t}) = h(\bar{\rho}) = -\sqrt{\alpha}\sqrt{1-|\boldsymbol{t}|^2}$. In summary, the assertion is hold. □

Formally, we can also couple the variables $\boldsymbol{s}$ with the constant α into a uniform form for the regularization term in the model (4) as

$$\bar{r}(\boldsymbol{s};\alpha) = |(\boldsymbol{s};\ \sqrt{\alpha})| = \sqrt{s_1^2+s_2^2+\alpha}.$$

So its conjugate form can be written as

$$\bar{r}^*(\bar{\boldsymbol{t}}) = \sup_{\boldsymbol{s}\in\mathbb{R}^2}\left[\boldsymbol{s}^T\check{\boldsymbol{t}} + \sqrt{\alpha}t_3 - \sqrt{|\boldsymbol{s}|^2+\alpha}\right] = \sup_{\boldsymbol{s}\in\mathbb{R}^2}\left[\boldsymbol{s}^T\check{\boldsymbol{t}} - \sqrt{|\boldsymbol{s}|^2+\alpha}\right] + \sqrt{\alpha}t_3$$

where $\bar{\boldsymbol{t}} = \left(\check{\boldsymbol{t}};t_3\right) := ((t_1,t_2);t_3)$.

Similar to the Theorem 2.1, we can obtain another convex conjugation form of $R(\boldsymbol{s})$.

Corollary 2.1. *The convex conjugate of $R(\boldsymbol{s})$ is given by*

$$R^*(\boldsymbol{t}) = \int_\Omega \bar{r}^*(\bar{\boldsymbol{t}})dx \tag{6}$$

where $\bar{r}^*(\bar{\boldsymbol{t}})$ is defined by

$$\bar{r}^*(\bar{\boldsymbol{t}}) = \begin{cases} \sqrt{\alpha}\left(t_3 - \sqrt{1-\left|\check{\boldsymbol{t}}\right|^2}\right) & \text{if } \left|\check{\boldsymbol{t}}\right| \leq 1, \\ +\infty & \text{if } \left|\check{\boldsymbol{t}}\right| > 1. \end{cases}$$

Now we consider the existence of the solution to the problem (4).

Theorem 2.2. *Assume that $u \in W^{1,1}(\Omega)$, $f \in L^2(\Omega)$ and the linear operator $K : L^1(\Omega) \to L^2(\Omega)$ is bounded, injective and satisfies the shift-invariance. For any $\alpha > 0$, the model (4) then has a unique solution.*

Proof. For the objective function $\mathcal{J}(u)$, its first Fréchet derivative in the direction $z \in C_c^1(\Omega)$ is

$$J(u,w) := \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{J}(u+tw)|_{t=0} = \lambda\int_\Omega K^T(Ku-f)\mathrm{d}x + \int_\Omega \frac{(\nabla u)^T\nabla w}{\sqrt{|\nabla u|^2+\alpha}}\mathrm{d}x.$$

Then we have the Hessian of $\mathcal{J}(u)$ at the direction z as

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}J(u+tw)|_{t=0} &= \lambda\int_\Omega K^TK\mathrm{d}x + \int_\Omega \frac{|\nabla w|^2\left(|\nabla u|^2+\alpha\right) - \left|(\nabla u)^T\nabla w\right|^2}{(|\nabla u|^2+\alpha)^{\frac{3}{2}}}\mathrm{d}x \\ &\geq \int_\Omega \lambda K^TK + \frac{|\nabla w|^2\alpha}{(|\nabla u|^2+\alpha)^{\frac{3}{2}}}\mathrm{d}x := \int_\Omega j(u,w)\mathrm{d}x. \end{aligned}$$

It is obvious that $j(u, w) > 0$, which implies that $\mathcal{J}(u)$ is strictly convex. Then the minimizer has to be unique if it indeed to be existed.

For the existence part, we first need the coercivity of the functional $\mathcal{J}(u)$, i.e., $\left|\mathcal{J}\left(u^k\right)\right| \to \infty$ whenever $\left\|u^k\right\| \to \infty$ for some sequence $\left\{u^k\right\} \in W^{1,1}(\Omega)$. For this propose, by assuming that $\left\|u^k\right\| \to \infty$, we can set $u^k = s^k v^k$ such that $s^k \geq 0$ and $\left\|v^k\right\|_2 = 1$ with $\left\{v^k\right\} \in W^{1,1}(\Omega)$. By employing the Sobolev inequality, there exists a constant $C_1 > 0$ and we can deduce that

$$\int_\Omega \sqrt{\left|\nabla u^k\right|^2 + \alpha}\mathrm{dx} \geq \int_\Omega \left|\nabla u^k\right| \mathrm{dx} = \int_\Omega \left|s^k \nabla v^k\right| \mathrm{dx} \geq C_1 s^k \left\|v^k\right\|_2 = C_1 s^k.$$

This implies that $\int_\Omega \sqrt{\left|\nabla u^k\right|^2 + \alpha} \to \infty$ when $k \to \infty$. By using the fact that $v^k \mapsto \|\cdot\|_2$ is coercive, we can deduce that $\mathcal{J}(u)$ is coercive. So we can deduce that $\left\{u^k\right\}$ is uniformly bounded. By compactness, there exists a subsequence, for convenience still denoted by $\left\{u^k\right\}$, which has an accumulation point u^* such that $u^* \in W^{1,1}(\Omega)$. By the lower semicontinuity of the smoothing convex function, we can deduce that

$$\mathcal{J}\left(u^*\right) = \liminf_{k\to\infty} \mathcal{J}\left(u^k\right)$$

Then the existence is proved. □

3. The primal-dual formulation

Let $\mathcal{X}$ and $\mathcal{Y}$ be two finite-dimensional real vector spaces equipped with an inner product $\langle\cdot,\cdot\rangle$ and the norm $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$. Now we consider the problem

$$\min_x f(x) + g(\mathcal{A}x), \tag{7}$$

where f and g are convex and lower semi-continuous function, $\mathcal{A}$ is a bounded linear operator with the operator norm $\|\mathcal{A}\|$ defined by

$$\|\mathcal{A}\| = \max\{\|\mathcal{A}\mathbf{z}\| \text{ with } \|\mathbf{z}\| \leq 1\}.$$

Formally the problem can be efficiently solved by some general optimization methods such as the Newton method or the gradient descent method when the objective functions are smooth. However, these methods are not suitable for the case as the problem including ℓ^1-norm or the operator $\mathcal{A}$ with complex structures or bigger scale. In order to cope with these, we can transform the problem (7) into the following min-max operator problem

$$\min_x \max_y f(x) + \langle \mathcal{A}x, y\rangle - g^*(y)$$

based on the definition of the conjugation function. To efficiently solve this saddle problem, the authors in [38] proposed to use a primal-dual hybrid method (named by PDHM) by applying gradient descending-ascending to primal-dual variable. They showed this method to be very efficient, especially when combining with an acceleration scheme consisting in decreasing the weight parameter in the updating of the original variable x and increasing the weight parameter in the updating of the dual variable y. Formally, Chambolle and Pock in [7] proposed to consider the following primal-dual method (PDM) as

$$\begin{cases} y^{k+1} = (\mathcal{I} + \varrho\partial g^*)^{-1}(y^k + \varrho\mathcal{A}\bar{x}^k) \\ x^{k+1} = (\mathcal{I} + \tau\partial f)^{-1}(x^k - \tau\mathcal{A}^* y^{k+1}) \\ \bar{x}^{k+1} = 2x^{k+1} - x^k \end{cases} \tag{8}$$

to solve this saddle problem, where $\mathcal{I}$ is the identity operator and ∂ denotes the sub-gradient defined by $\partial\hbar(\bar{y}) := \{v|\hbar(\bar{x}) - \hbar(y) \geq (v, \bar{x} - \bar{y})\}$ at the point $\bar{y}$ for a function $\hbar$. Then they also gave the following convergence result of the scheme (8) as follows.

Theorem 3.1. *Let $\|\mathcal{A}\| = \mathcal{T}$ and assume that the problem (7) has a saddle point (x^*, y^*). If $\tau\varrho\mathcal{T}^2 < 1$ and choosing some suitable original values $\left\{\left(\bar{x}^0, y^0\right)\right\}$, then the sequence $\left\{\left(x^k, y^k\right)\right\}$ by the scheme (8) converges to this saddle. Furthermore, if the problem (7) has a global solution x^*, then the sequence $\left\{x^k\right\}$ converges to x^*.*

Remark 3.1. *The updating step in the iteration scheme (8) can be interpreted as an approximate extragradient and indeed it is possible to show the convergence of Theorem 3.1 with a convergence rate $O(1/k)$ as shown in [22]. In addition, above primal-dual method is intimately connected with other splitting methods such as the proximity point algorithm [16], Douglas-Rachford splitting method [17, 4], and alternating direction multiplier method [24, 14, 30], etc.*

4. Numerical method to the proposed model (4)

Throughout the rest of the paper, we assume the images are matrices with the size of $M \times N$. For the boundary conditions, there are many choices such as the zero Dirichlet boundary [3], the Neumann boundary (also called symmetric or reflective) [23], the reflexive boundary [21], the antireflective boundary [10], the periodic boundary or even without any boundary artifacts [26]. Here we only consider the periodic boundary condition. Let us define the Euclidean space $X = \mathbb{R}^{M\times N}$ and $Y = X \times X$. The usual scalar products can be denoted as $\langle \mathbf{v}, \mathbf{w}\rangle_X := \sum_{i=1}^{M}\sum_{j=1}^{N} \mathbf{v}_{i,j}\mathbf{w}_{i,j}$ with the norm $\|\mathbf{v}\|_X = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle_X}$ for $\mathbf{v}, \mathbf{w} \in X$ and $\langle \mathbf{p}, \mathbf{q}\rangle_Y := \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{\iota=1}^{2} \mathbf{p}^{\iota}_{i,j}\mathbf{q}^{\iota}_{i,j}$ for $\mathbf{p}, \mathbf{q} \in Y$. In the following, we define the discrete gradient $\nabla u = (D^+_x u, D^+_y u)$ with the forward difference operators

$$D^+_x u_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j} & \text{if } 1 \le i < M, 1 \le j \le N, \\ u_{1,j} - u_{i,j} & \text{if } i = M, 1 \le j \le N, \end{cases}$$
$$D^+_y u_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j} & \text{if } 1 \le i \le M, 1 \le j < N, \\ u_{i,1} - u_{i,j} & \text{if } 1 \le i \le M, j = N. \end{cases}$$

We also define the backward difference operators as

$$D^-_x p^1_{i,j} = \begin{cases} p^1_{i,j} - p^1_{i-1,j} & \text{if } 1 < i \le M, 1 \le j \le N, \\ p^1_{i,j} - p^1_{M,j} & \text{if } i = 1, 1 \le j \le N, \end{cases}$$
$$D^-_y p^2_{i,j} = \begin{cases} p^2_{i,j} - p^2_{i,j-1} & \text{if } 1 \le i \le M, 1 < j \le N, \\ p^2_{i,j} - p^2_{i,N} & \text{if } 1 \le i \le M, j = 1. \end{cases}$$

Based on the relation $\langle u, \mathrm{div}\mathbf{p}\rangle_X = -\langle \nabla u, \mathbf{p}\rangle_Y$ in [6], we can obtain the divergence operator as

$$\mathrm{div}\mathbf{p} = D^-_x p^1 + D^-_y p^2$$

for $\mathbf{p} \in Y$. Therefore, the discrete equivalent of (4) can be denoted by

$$\min_u \frac{\lambda}{2}\|Ku - f\|^2_X + \left\|\sqrt{\alpha + |\nabla u|^2}\right\|_X, \tag{9}$$

where $|\nabla u|^2 = (D^+_x u)^2 + (D^+_y u)^2$. In the following we consider two different schemes to solve this discrete model (9), which is a convex and smooth optimization problem.

4.1. Primal-dual method (I)

If setting $\mathcal{A} := \mathcal{A}_1 = \nabla$ and using the convex conjugate properties of Example 8.5 in [34] as

$$\theta(t) = \sqrt{\alpha + t^2} \Longleftrightarrow \theta(t) = \sup_{\|s\|_\infty \le 1}\left\{\langle t, s\rangle + \sqrt{\alpha\left(1 - s^2\right)}\right\},$$

we can rewrite the minimization problem (9) as a min-max problem

$$\min_{u} \max_{\|\mathbf{p}\|_\infty \le 1} \left\{ \frac{\lambda}{2}\|Ku - f\|_X^2 - \langle u, \mathrm{div}\mathbf{p}\rangle_X + \left\| \sqrt{\alpha\left(1 - |\mathbf{p}|^2\right)} \right\|_X \right\} \tag{10}$$

for $\mathbf{p} \in Y$. Since the subjective function (10) is proper convex, we can interchange the order of min and max and then solve it by using the primal-dual scheme [7]. Therefore, we separate the problem (10) into the following two subproblems.

- For the primal variable u in the problem (10): By ignoring the unrelated term to u, we can obtain

  $$\min_{u} \left\{ \frac{\lambda}{2}\|Ku - f\|_X^2 - \langle u, \mathrm{div}\mathbf{p}\rangle_X \right\},$$

  which optimality condition is

  $$\lambda K^T(Ku - f) - \mathrm{div}\mathbf{p} = 0,$$

  where T denotes the matrix transpose. In general, the blurring operator matrix K is ill-posed, we can use the gradient method to compute u

  $$u^k - u^{k+1} = \tau\left(\lambda K^T(Ku^{k+1} - f) - \mathrm{div}\mathbf{p}\right).$$

  Due to the assumption of the periodic boundary condition, we can use the Fast Fourier Transform (FFT) to solve u by

  $$u^{k+1} := \mathcal{U}^{k+1}(\mathbf{p}) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(u^k) + \lambda\tau\mathcal{F}(K^T f) + \tau\mathcal{F}(\mathrm{div}\mathbf{p})}{\mathcal{F}(I) + \lambda\tau\mathcal{F}(K^T K)}\right), \tag{11}$$

  where I is the identity matrix, $\mathcal{F}$ denotes the FFT operator and $\mathcal{F}^{-1}$ denotes its inverse transform.

- For the dual variable $\mathbf{p}$ in the problem (10): By ignoring the unrelated term to $\mathbf{p}$ and introducing an indicator function, we have

  $$\chi_{\mathcal{K}}(\mathbf{p}) = \begin{cases} 0, & \text{if } \mathbf{p} \in \mathcal{K}, \\ +\infty, & \text{if } \mathbf{p} \notin \mathcal{K}, \end{cases}$$

  where $\mathcal{K} := \{\mathbf{p} \mid \|\mathbf{p}\|_\infty \le 1\}$. Then we have

  $$\max_{\mathbf{p}} \left\{ \langle \nabla u, \mathbf{p}\rangle_Y + \left\| \sqrt{\alpha\left(1 - |\mathbf{p}|^2\right)} \right\|_X + \chi_{\mathcal{K}}(\mathbf{p}) \right\}.$$

  The optimality condition of the above maximization problem is

  $$\left(\nabla u + \partial\chi_{\mathcal{K}}(\mathbf{p})\right)\sqrt{1 - |\mathbf{p}|^2} - \sqrt{\alpha}\mathbf{p} = 0.$$

  Note that $\partial\chi_{\mathcal{K}}(\mathbf{p}) = 0$ because indicator function is a constant function. Based on the projection gradient method, we have

  $$\mathbf{p}^{k+1} := \mathcal{P}^{k+1}(u) = \frac{\mathbf{p}^k + \varrho\left(\nabla u\sqrt{1 - \left|\mathbf{p}^k\right|^2} - \sqrt{\alpha}\mathbf{p}^k\right)}{\max\left\{1, \left|\mathbf{p}^k + \varrho\left(\nabla u\sqrt{1 - \left|\mathbf{p}^k\right|^2} - \sqrt{\alpha}\mathbf{p}^k\right)\right|\right\}}. \tag{12}$$

By combining (11) with (12) and using Theorem 3.1, we have the following convergence results.

Corollary 4.1. *Assume that $\tau\varrho \le 1/8$ and choosing $\bar{u}^0 = f$, then the sequence $\left\{\left(u^{k+1}, \boldsymbol{p}^{k+1}\right)\right\}$ generated by*

$$\begin{cases} \boldsymbol{p}^{k+1} := \mathcal{P}^{k+1}\left(\bar{u}^k\right) \text{ ←-- using the formula (12)}; \\ u^{k+1} := \mathcal{U}^{k+1}\left(\boldsymbol{p}^{k+1}\right) \text{ ←-- using the formula (11)}; \\ \overline{u}^k \ = 2u^{k+1} - u^k; \end{cases} \tag{13}$$

converges to the saddle point $(u^*, \boldsymbol{p}^*)$ of the problem (10). Furthermore, u^* is the solution of the problem (9).

Remark 4.1. *For Corollary 4.1, if we set $F(u) = \frac{\lambda}{2}\|Ku - f\|_X^2$ and $G(\nabla u) = \left\| \sqrt{|\nabla u|^2 + \alpha} \right\|_X$, the iteration scheme (13) is the exact Algorithm 1 used in [7]. Furthermore, the convergence can be kept since $\|\nabla\|_Y \le 1/8$ (See Theorem 3.1 in [6]).*

4.2. Primal-dual method (II)

As mentioned earlier, the image can be regarded as an embedded shrinkage surface $\left(\sqrt{\alpha}x_1,\ \sqrt{\alpha}x_2, u(x_1, x_2)\right)$ in $\mathcal{M}$. One thus considers the surface $\phi(x_1, x_2) - \sqrt{\alpha}z = 0$ and motivated by the work in [39], we can set the operator $\mathcal{A} := \mathcal{A}_2 = \left(\nabla, -\sqrt{\alpha}\right)$. As it makes no difference for replacing $\mathcal{A}_2 = \left(\nabla,\ \sqrt{\alpha}\right)$ by $\mathcal{A}_2 = \left(\nabla, -\sqrt{\alpha}\right)$ in the proposed model, so we transform the model (9) into following min-max saddle model as

$$\min_u \max_{\mathbf{q}} \ \left\{\frac{\lambda}{2}\|Ku - f\|_X^2 + \langle (u, 1), \mathcal{A}_2^*\mathbf{q}\rangle + \chi_C(\mathbf{q})\right\}, \tag{14}$$

where $\mathbf{q} = \left(q^1, q^2; q^3\right)$ and $C =: \{\mathbf{q} \mid \|\mathbf{q}\|_\infty \le 1\}$. Based on the divergence theorem and the definition of the operator $\mathcal{A}_2$, we can deduce that $\mathcal{A}_2^*\mathbf{q} = \left(-\mathrm{div}\left(q^1, q^2\right),\ \sqrt{\alpha}q^3\right)$. Set $\bar{\mathbf{q}} = (q^1, q^2)$, then the solution of the subproblem u^k in (14) can be similarly obtained by

$$u^{n+1} := \mathcal{U}^{n+1}(\mathbf{q}) = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(u^n) + \lambda\tau\mathcal{F}(K^T f) + \tau\mathcal{F}(\mathrm{div}\mathbf{q})}{\mathcal{F}(I) + \lambda\tau\mathcal{F}(K^T K)}\right). \tag{15}$$

To solve the subproblem $\mathbf{q}$, we first focus on the optimization condition (14) to be written by

$$(\nabla u,\ \sqrt{\alpha}) + \partial\chi_C(\mathbf{q}) = 0,$$

so we still use the projection gradient method to obtain the numerical solution as

$$\mathbf{q}^{n+1} := Q^{n+1}(u) = \frac{\mathbf{q}^n + \varrho\left(\nabla u,\ \sqrt{\alpha}\right)}{\max\left\{1,\ \left|\mathbf{q}^n + \varrho\left(\nabla u,\ \sqrt{\alpha}\right)\right|\right\}}. \tag{16}$$

In order to keep the convergence of above scheme, we also need to know the spectral radius of the operator $\mathcal{A}_2$. Formally, we can deduce that

$$\begin{aligned} \|\mathcal{A}_2^*\mathbf{q}\|_X^2 =& \sum_{1\le i\le M, 1\le j\le N} \left[\left(q^1_{(i,j)} - q^1_{(i-1,j)} + q^2_{(i,j)} - q^2_{(i,j-1)}\right)^2 + \alpha\left(q^3_{(i,j)}\right)^2\right] \\ \le& \sum_{1\le i\le M, 1\le j\le N} \left[4\left(\left(q^1_{(i,j)}\right)^2 + \left(q^1_{(i-1,j)}\right)^2 + \left(q^2_{(i,j)}\right)^2 + \left(q^2_{(i,j-1)}\right)^2\right) + \alpha\left(q^3_{(i,j)}\right)^2\right] \\ \le& 8\|\bar{\mathbf{q}}\|_Y^2 + \alpha\left\|q^3\right\|_X^2 \le \max(8, \alpha)\|\mathbf{q}\|_{Y\times X}^2. \end{aligned}$$

Then based on above fact, we obtain the convergence result as

Corollary 4.2. *Assume that $\tau\varrho \le 1/\max(8, \alpha)$ and choosing $\bar{u}^0 = f$, then the sequence $\left\{\left(u^{n+1}, \boldsymbol{q}^{n+1}\right)\right\}$ generated by*

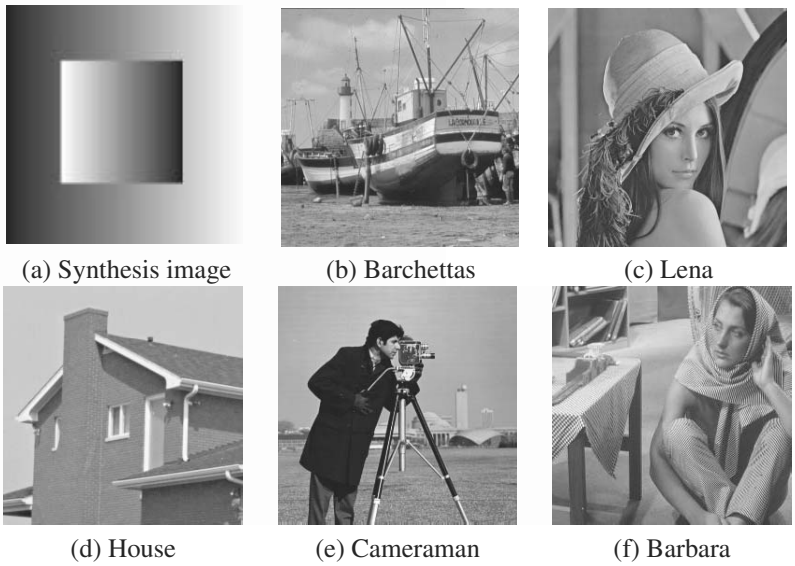
$$\begin{cases} \boldsymbol{q}^{n+1} := Q^{n+1}\left(\bar{u}^n\right) \text{ ←-- using the formula (16)}; \\ u^{n+1} := \mathcal{U}^{n+1}\left(\boldsymbol{q}^{n+1}\right) \text{ ←-- using the formula (15)}; \\ \overline{u}^n \ = 2u^{n+1} - u^n; \end{cases} \tag{17}$$

converges to the saddle point $(u^*, \boldsymbol{p}^*)$ of the problem (10). Furthermore, u^* is the solution of the problem (9).

5. Numerical implementations

In numerical implementations, we only consider to use the proposed model (4) for the basic image restoration problems, i.e., image denoising and deblurring. In fact, the model (4) has many other applications, for example the CT or MRI reconstruction problems and image inpainting problems with different operators K, etc.. In order to demonstrate the advantage of the proposed Primal-Dual Method (PDM) by respectively setting as the PDM(I) for the iteration scheme (13) and setting as the PDM(II) for the iteration scheme (17). We also compare it with another two classic numerical methods, which are the Time Marching Method (TMM) [28] and the Fixed Point Method (FPM) [9], respectively. All the algorithms will stop when $\max\left\{\frac{\|u^{k+1}-u^k\|_2}{\|u^k\|_2}, \frac{\|E(u^{k+1})-E(u^k)\|}{|E(u^k)|}\right\} \leq 10^{-5}$ or the iteration arrives to 500. The simulations are preformed in Matlab (R20016a) on a PC with an Intel Core i7-5500U at 2.40 GHz and 8 GB of memory.

We use the "Lena" image of different sizes, i.e., 128×128, 256×256, 512×512 and another five images of size 256×256 in the numerical experiments, which are shown in Figure 2. To standardize the discussions, we first normalize the pixel values of the test image $\bar{f}$ to be [0,255] by using the linear-stretch formula as $f = \frac{255\times(\bar{f}-\min(\bar{f}))}{\max(\bar{f})-\min(\bar{f})}$, where max and min represent the maximum and minimum of $\bar{f}$, respectively.



(a) Synthesis image (b) Barchettas (c) Lena

(d) House (e) Cameraman (f) Barbara

Figure 2. Test images in numerical implementations.

For the sake of simplicity, we use σ to denote standard deviation of the white Gaussian noise and $G(hsize, \tilde{\sigma})$ denotes the symmetric Gaussian low-pass filter of size $hsize$ with the standard deviation $\tilde{\sigma}$. In order to compare the visual perception and quality metric point of view, the performance of each method is evaluated in terms of the signal to noise ratio (SNR) and the structural similarity index ($SSIM$): the higher SNR and $SSIM$ implies better restoration results. In addition, we explicitly give the update scheme of the TMM [28] and the FPM [34, 35] as follows

$$\frac{u^{k+1}-u^k}{dt} = \left(\lambda K^T\left(Ku^k - f\right) - E_\alpha\left(u^k\right)\right) \qquad (\rightarrow TMM) \tag{18}$$

$$\left(\lambda K^T K - \operatorname{div}\left(\frac{\nabla}{\sqrt{|\nabla u^k|^2+\alpha}}\right)\right)u^{k+1} = \lambda K^T f \qquad (\rightarrow FPM) \tag{19}$$

by choosing a suitable original value u^0. We set $dt = 0.05$ for the TMM which can give a better numerical results for all of implementations. To the PDM, we set $\tau = \varrho = \frac{1}{\sqrt{8}}$ to keep its convergence. Note the matrix operator in the left of the FPM is symmetric and positive definite. Therefore, we can employ the conjugate gradient method to solve it as [34, 35].

5.1. The choice of regularized parameter λ

In the proposed model, the restoration effectiveness depends on the regularization parameter λ that controls how much filtering is introduced by the regularization. Often the key issue in connection with these methods is to find a regularization parameter that gives a good balance, filtering out enough noise without losing too much information in the computed solution. Over estimating the regularized parameter λ may lead to the domination for the regularization term $\mathcal{R}(u)$. So it will over-penalize image details and obtain over smoothed blurry result with a low energy functional. On the contrary, while underestimating it may leave the noise in the image unfiltered. Then we obtain a noisy image and a high energy. So it is very important to find a suitable λ based on these relation. Some classic methods to choose the suitable λ can be employed such as the L-curve method [15], the generalized cross validation (GCV) [12], the discrepancy principle [20], or the variational Bayes approach [25].

In our experiments, the main aim is focused on the advantages of the proposed adaptive norms model and especially the proposed numerical method. If above methods for selecting regularization parameter are used, it will bring great uncertainties for the comparison between the proposed and other traditional methods. Therefore, for a fair comparison between different models, the regularization parameter was manually determined by attempting a series of values and selecting the one with the highest SNR (in simulated experiments) or the best visual effect (in real experiments). Specifically, we set the related parameters into a bigger range as $[a, b]$ and then find a suitable sunset as $[c, d] \subset [a, b]$. In the next we find a more suitable parameter in $[c, d]$. When the difference between the successive SNR is below 0.001, we set this parameter as the best value to the regularization parameter λ.

5.2. Advantages of the proposed model

In order to distinctly illustrate the performance of the two approaches, we consider to restore the noisy image

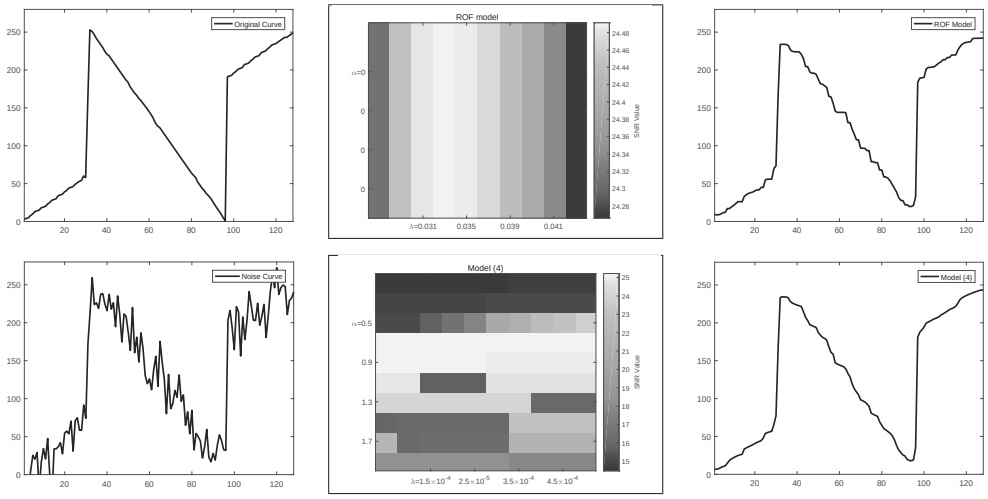


Figure 3. The SNR images with different parameters and related curves based on the image in Figure 2(a).

degraded by the white Gaussian noise with $\sigma = 20$. The algorithm will be stopped when the stopping condition is met. We first show the abilities to keep edges and smoothing regions of the model (4) by employing the image (*a*) in Figure 2 as the testing image. This choice is based on the fact that this image has clear edges and abundantly smoothing regions. In order to choose suitable values of α and λ in the proposed model (4) comparing with the ROF model, i.e., setting $\alpha = 0$ in the model (4). We first estimate original intervals of parameters and careful change their values in these intervals with the suitable step distance. Specifically, we choose $\lambda \in [0.025, 0.44]$ with the stepsize 0.02 for the ROF model and $\lambda \in \left[10^{-5}, 5 \times 10^{-4}\right]$ with the stepsize 5×10^{-5} and $\alpha \in [0.1, 2]$ with the stepwise 0.1 for the proposed model (4). The plot indicates that the SNR has a global maximum for different values of parameters. Especially, for the ROF model, we set $\lambda = 0.032$ to obtain the best SNR of 24.4951 dB. Similarly, we choose $\lambda = 5 \times 10^{-5}$ and $\alpha = 0.8$ in the proposed model (4), which gives the best SNR of 25.1823 dB. In order to show the detail comparisons for these two models, we show a piece of slice with the 80th row of the related images in Figure 3. It is obviously that

the proposed model (4) have a robust restored abilities comparing with the ROF model, especially for suppressing the staircase effect in the smoothing regions.

5.3. Impact of parameter α

As we can see the results from subsection 4.1 that α is one important parameter in the proposed model. Thus, we test our proposed PDM(I) with different contaminated levels by the noise or the Gaussian blurring to validate their effect in (10). We use the image "Barchetta" shown in Figure 2 (b). Comparing with the artificial image used in the subsection 4.1, this image includes more detail structures, so it is more suitable to test α by the effect of contamination. We still follow the strategy for choosing suitable parameter λ and α used in subsection 4.1. Here the degraded images are generated by the white Gaussian noise with $\sigma = 10, 20, 30$ or the Gaussian blurring with $G(21, 0.6)$ and these noises. To the ROF model, we again set the parameter $\alpha = 0$ and solve it by the primal dual method as used in the subsection 4.1. It is easy to deduce that the proposed model (4) can obtain a better restored image than the ROF model based on the numerical results in Table 1. Furthermore, we should observe that the choice of α rarely depends on the noise change if there are not any blurs. However, we need to decrease the value of α while increasing the noise level in the blurring image. Actually, it is because of that more details are destroyed by the blurring and noise. On the other hand, we need to decease the value of regularization parameter λ while increasing the noise level. This is based on the fact we need to increase the weigh of data regularization term to keep the image details.

σ	10		20		30	
Model	ROF	Model (4)	ROF	Model (4)	ROF	Model (4)
λ	0.19	0.12	0.07	0.05	0.04	0.03
α	0	0.05	0	0.06	0	0.06
SNR	18.6239	18.6320	14.4890	14.5216	12.4151	12.4878
SSIM	0.7537	0.7541	0.6112	0.6127	0.5197	0.5281
(σ, G)	$(10, G(21, 0.6))$		$(20, G(21, 0.6))$		$(30, G(21, 0.6))$	
Model	ROF	Model (4)	ROF	Model (4)	ROF	Model (4)
λ	0.33	0.32	0.08	0.09	0.05	0.05
α	0	0.006	0	0.001	0	0.0001
SNR	15.7181	15.7211	12.6904	12.7489	11.2536	11.2799
SSIM	0.6722	0.6723	0.5335	0.5371	0.4621	0.4630

Table 1. The related data by restoring contaminated "Barchettas" images with different values of α. Here, λ denotes the regularization parameter.

5.4. Comparison with the TMM and FPM

We analyze the performance of our proposed PDM by comparing it with the TMM and FPM for restoration of "Lena" images with different sizes. Here, the blur and noisy images are corrupted by additive white Gaussian noise with $\sigma = 10$ and the Gaussian blur with $G(21, 0.6)$. Table 2 illustrates the values of SNR, $SSIM$ and CPU time when the numerical algorithms stop. We can observe that SNR and CPU time increases while $SSIM$ decreases when the image size increases. In fact, large parameter λ is required to penalize the data fitting term when we increase the size of the test image without increasing the level of noises. From the comparisons of SNR, $SSIM$, $Time$ and Ite, our proposed PDM is shown better than both the TMM and FPM, especially the CPU time. These also follow from the observation in the Figure 4. Table 2 clearly shows that the TMM usually takes more iterations to obtain the steady solution while although the FPM need less iteration, it requires to solve a linear equation using the conjugated gradient method with inner iterations. Therefore, our proposed PDM is the most efficient one among the three algorithms, especially, when dealing with the large scale image.

Image	Lena image contaminated by noise with $\sigma = 10$.											
(S, λ, α)	$(128 \times 128, 0.16, 0.009)$				$(256 \times 256, 0.17, 0.03)$				$(512 \times 512, 0.12, 0.013)$			
Method	TMM	FPM	PDM(13)	PDM(17)	TMM	FPM	PDM(13)	PDM(17)	TMM	FPM	PDM(13)	PDM(17)
SNR	17.8674	17.9446	18.1622	18.0441	18.7203	19.3058	19.3350	19.3397	20.0060	20.0094	20.2870	20.1919
SSIM	0.8284	0.8408	0.8354	0.8433	0.7133	0.7406	0.7410	0.7404	0.6187	0.6041	0.6344	0.6159
Time(s)	13.5156	6.8750	0.9844	4.2500	27.8906	21.9531	7.7188	15.4844	384.1875	79.8438	21.4063	93.0781
Ite	303	20	27	149	180	19	65	136	500	18	33	152
Image	Lena image contaminated by noise with $\sigma = 10$ and Gaussian blur with $G(21, 0.6)$.											
(S, λ, α)	$(128 \times 128, 0.29, 0.02)$				$(256 \times 256, 0.26, 0.011)$				$(512 \times 512, 0.23, 0.02)$			
Method	TMM	FPM	PDM(13)	PDM(17)	TMM	FPM	PDM(13)	PDM(17)	TMM	FPM	PDM(13)	PDM(17)
SNR	13.7845	15.3528	15.3807	14.3120	15.4783	16.8419	16.8822	15.9692	17.5351	18.7662	18.7888	17.8123
SSIM	0.7146	0.7932	0.7905	0.7351	0.6069	0.6828	0.6821	0.6304	0.5350	0.5755	0.5773	0.5468
Time(s)	2.3906	31.7344	3.0781	1.5625	14.6406	94.5156	13.0625	6.5000	104.1875	293.9375	48.1406	39.4375
Ite	38	29	77	34	66	29	81	44	123	23	61	53

Table 2. The related data by restoring contaminated "Lena" image of different sizes.

On the other hand, we fix the iteration number to be 500 and compare the energy decay of the objective functional, the variations of the SNR, SSIM and the computational time. In Figure 4, we use the "Lena" with size of 256×256 as an example and consider the same two cases, i.e., the test image is corrupted by only white Gaussian noise $\sigma = 10$ (the first row in Figure 4, and both white Gaussian noise $\sigma = 10$ and Gaussian blur $G(21, 0.6)$ (the second row in Figure 4, respectively. As shown in Figure 4, we observe that the FPM, PDM (13) and (17) perform better than the TMM, the numerical energies of which converge to the same value and are smaller than the TMM. Indeed, the FPM, PDM (13) and (17) produce quite satisfactory and accurate restoration results. Our proposed PDM (13) and (17) outperforms the FPM in computational costs due to its low expenses in each iteration, even though the FPM converges faster than the PDM (13) and (17). In addition, we notice that the curves for the CPU time of the TMM, PDM (13) and (17) linearly depend on the iteration. However, this curve is nonlinear to the FPM since using the conjugate gradient method [35] to solve the linear equation system (19) effects the tendency of this curve. Now we need to give some discussions for the PDM-type (13) and (17). Formally, we can find that the PDM (13) in most cases outperforms the PDM (17). The reason is of that the dual variable in the PDM (13) has more compact form than in the PDM (17). So we can expect to obtain more rust numerical results.

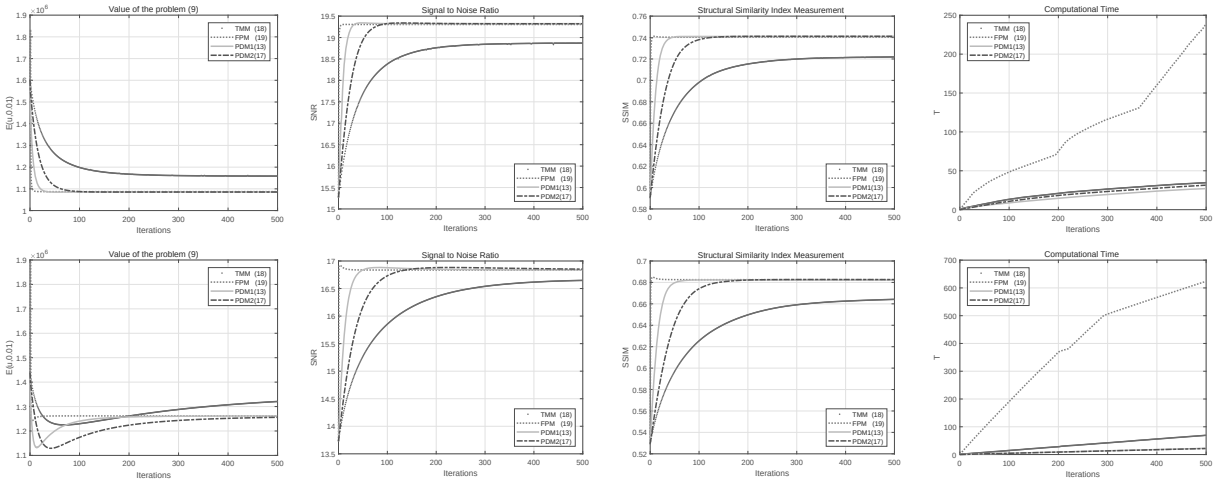


Figure 4. Related curves by restoring the Lena image with the size 256×256 and fixed iteration as 500. First Column: Enegy decay of the objective functional in the model (4); Second Column: Values of SNR; Third Column: Values of SSIM; Forth Column: CPU Time.

5.5. Extend to other three images

Here we test our proposed PDM on other degraded images, where each test image has its own specialty, i.e., "House" has much more sharp edges, "Cameraman" has more affine regions, and "Barbara" includes many texture structures. To generate the degraded images, we add the additive white Gaussian noise and apply the Gaussian convolution to the images. Similar results are obtained on these three test images, as shown in Table 3. Similarly, we are able to conclude that our proposed PDM-types are more efficient than the other two methods.

Image	House				Cameraman				Barbara			
(σ,λ,α)	$(20, 0.041, 0.025)$				$(22, 0.05, 0.007)$				$(20, 0.06)$			
Method	TMM	FPM	PDM(I)	PDM(II)	TMM	FPM	PDM(I)	PDM(II)	TMM	FPM	PDM(I)	PDM(II)
SNR	116.4014	16.0192	16.9087	16.4325	15.2826	16.0762	16.1427	16.1467	10.0807	10.4062	10.5097	10.5236
SSIM	0.3896	0.3256	0.3820	0.3366	0.3825	0.3947	0.3963	0.3971	0.5605	0.5883	0.5922	0.5929
Time(s)	79.4688	29.8594	9.3594	47.4531	78.4219	30.6406	20.1250	40.5625	56.5156	34.3906	12.3594	41.5938
Ite	500	20	77	401	497	23	165	353	382	22	163	370

Image	House				Cameraman				Barbara			
$(\sigma,G,\lambda,\alpha)$	$(20, G(21, 0.6), 0.058, 0.001)$				$(22, G(21, 0.6), 0.06, 0.003)$				$(25, G(21, 0.6), 0.06, 0.002)$			
Method	TMM	FPM	PDM(I)	PDM(II)	TMM	FPM	PDM(I)	PDM(II)	TMM	FPM	PDM(I)	PDM(II)
SNR	12.2995	15.7617	15.8030	15.7303	13.4439	13.5288	13.9426	13.9381	8.5448	8.5463	8.7650	8.7631
SSIM	0.3140	0.3246	0.3444	0.3520	0.3368	0.3285	0.3503	0.3512	0.4616	0.4894	0.4903	0.4891
Time(s)	28.7031	65.2344	16.7188	33.0938	61.1406	154.3594	14.7969	28.2188	56.1719	47.4688	14.3281	16.2188
Ite	141	10	90	210	315	25	80	210	331	9	74	197

Table 3. The related data by restoring contaminated image "House", "Cameraman" and "Barbara".

6. Conclusions

We presented an image restoration model based on the minimized surface regularization, which closely relates to the smoothing ROF model [28]. By using the property of conjugate function, we reformulated the proposed model as a min-max problem and used two different kinds of primal-dual methods [7] to solve it. The convexity of the model can guarantee the proposed algorithm to converge to a unique global minimizer. Numerical experiments demonstrated that the proposed methods hold the potential for efficient and stable computation by comparing to the classic time marching method (TMM) [28] and the lagged diffusivity fixed point method (FPM) [34, 35], especially for the large-scale images. In the future, we would like to extend the proposed methods to other image processing problems such as image inpainting, reconstruction, registration and also for vector value images, etc..

Acknowledgements

This work is partially supported by the National Basic Research Program of China (973 Program) (No. 2015CB856 003) and National Natural Science Foundation of China (Nos. U1304601, 11401170, 11526208, 61373087). Dr. Jian Lu is supported by Guangdong Natural Science Foundation of China (Nos. 2015A030313557, 2015A030313550).

References

[1] R. Acar and C. Vogel. Analysis of bounded variation penalty methods for ill-posed problems. Inverse problems, 10:1217-1229, 1994.
[2] G. Aubert and P. Kornprobst. Mathematical Problem in Image Processing: partial differential equations and the calculus of variations. Springer, 2008.
[3] M. Bertero and P. Boccacci. Introduction to inverse problems in imaging. CRC Press, 1998.
[4] K. Bredies and H. Sun. Preconditioned Douglas-Rachford algorithms for TV- and TGV -regularized variational imaging problems. Journal of Mathematical Imaging and Vision, 52(3):317-344, 2015.
[5] A. Buades, B. Coll, and J. Morel. A review of image denoising methods, with a new one. Multiscale Modeling and Simulation, 4(2):490-530, 2006.
[6] A. Chambolle. An algorithm for total variation minimization and applications. Journal of Mathematical Imaging and Vision, 20(1-2):89-97, 2004.
[7] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision, 40(1):120-145, 2011.
[8] T. Chan and J. Shen. Image Processing and Analysis-Variational, PDE, Wavelet, and Stochastic Methods. SIAM, Philadephia, 2005.
[9] T. Chan, G. 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.
[10] M. Christiansen and M. Hanke. Deblurring methods using antireflective boundary conditions. SIAM Journal on Scientific Computing, 30(2):855-872, 2008.
[11] I. Ekeland and T. Turbull. Infinite Dimensional Optimization and Convexity. The University of Chicago Press, 1983.
[12] N. Galatsanos and A. Katsaggelos. Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. IEEE Transactions on Image Processing, 1(3):322-336, 1992.
[13] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. Multiscale Modeling and Simulation, 7(3):1005-1028, 2008.
[14] T. Goldstein and S. Osher. The split Bregman method for L1-regularized problems. SIAM Journal on Imaging Sciences, 2(2):323-343, 2009.
[15] P. Hansen and D. Leary. The use of the L-curve in the regularization of discrete ill-posed problems. SIAM Journal on Scientific Computing, 14(6):1487-1503, 1993.
[16] B. He and X. Yuan. Convergence analysis of primal-dual algorithms for a saddle-point problem: from contraction perspective. SIAM Journal on Imaging Sciences, 5(1):119-149, 2012.
[17] B. He and X. Yuan. On the O(1/t) convergence rate of the Douglas-Rachford alternating direction method. SIAM Journal on Numerical Analysis, 50(2):700-709, 2012.
[18] R. Kimmel, R. Malladi, and N. Sochen. Images as embedded maps and minimal surfaces: movies, color, texture, and volumetric medical images. International Journal of Computer Vision, 39(2): 111-129, 2000.
[19] J. Lu, L. Shen, C. Xu, and Y. Xu. Multiplicative noise removal in imaging:An exp-model and its fixed-point proximity algorithm. Applied and Computational Harmonic Analysis, 41(2):518-539, 2016.
[20] V. Morozov. *Methods for solving incorrectly posed problems*. Springer-Verlag, New York, 1984.
[21] J. Nagy, M. Ng, and L. Perrone. Kronecker product approximations for image Restoration with reflexive boundary conditions. SIAM Journal on Matrix Analysis and Applications, 25(3):829-841, 2003.
[22] A. Nemirovski. Prox-method with rate of convergence O(1/t) for variational inequalities with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems. SIAM Journal on Optimization, 15(1):229-251, 2004.
[23] M. Ng, R. Chan, and W. Tang. A fast algorithm for deblurring models with Neumann boundary conditions. SIAM Journal on Scientific Computing, 21(3):851-866, 1999.
[24] M. Ng, F. Wang, and X. Yuan. Inexact alternating direction methods for image recovery. SIAM Journal on Scientific Computing, 33(4):1643-1668, 2011.

[25] J. Oliveira, J. Bioucas-Dias, and M. Figueiredo. Adaptive total variation image deblurring: A majorizationCminimization approach. Signal Processing, 89(9):1683-1693, 2009.
[26] S. Reeves. Fast image restoration without boundary artifacts. IEEE Transactions on Image Processing, 14(10):1448-1153, 2005.
[27] R. Rockafellar and R. Wets. Variational Analysis. Springer, 1998.
[28] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. Physica D, 60(1-4):259-268, 1992.
[29] O. Scherzer. Handbook of Mathematical Methods in Imaging. Springer New York, 2015.
[30] R. Shefi and M. Teboulle. Rate of convergence analysis of decomposition methods based on the proximal method of multipliers for convex minimization. SIAM Journal on Optimization, 24(1):269-297, 2014.
[31] N. Sochen, R. Kimmel, and R. Malladi. A general framework for low level vision. IEEE Transactions on Image Processing, 7(3):310-318, 1998.
[32] J. Thomas. Numerical Partial Differential Equations: Finite Difference Methods. New York: Springer-Verlag, 1995.
[33] C. Wu and X. Tai. Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models. SIAM Journal on Imaging Sciences, 3(3):300-339, 2010.
[34] C. Vogel. Computational Methods for Inverse Problems. SIAM, Philadelphia, 2002.
[35] C. Vogel and M. Oman. Iterative methods for total variation denoising. SIAM Journal on Scientific Computing, 17(1):227-238, 1996.
[36] D. Yi. An Algorithm for image removals and decompositions without inverse matrices. Journal of Computational and Applied Mathematics, 225(2):428-439, 2009.
[37] D. Zosso and A. Bustin. A primal-dual projected gradient algorithm for efficient Beltrami regularization. UCLA CAM Report 14-52.
[38] M. Zhu, S. Wright, and T. Chan. Duality-based algorithms for total-variation-regularized image restoration. Computational Optimization and Applications, 47(3):377-400, 2010.
[39] W. Zhu, X. Tai and T. Chan. Augmented Lagrangian method for a mean curvature based image denoising model. Inverse Problems and Imaging, 7(4):1409-1432, 2013.