Research Article

Journal of Algorithms & Computational Technology

Journal of Algorithms & Computational Technology
2016, Vol. 10(4) 235–243

DOI: 10.1177/1748301816656298
act.sagepub.com
SAGE

Primal-dual method to smoothing TV-based model for image denoising

Zhanjiang Zhi[1], Baoli Shi[2] and Yi Sun[1]

Abstract

The total variation-based Rudin–Osher–Fatemi model is an effective and popular prior model in the image processing problem. Different to frequently using the splitting scheme to directly solve this model, we propose the primal dual method to solve the smoothing total variation-based Rudin–Osher–Fatemi model and give some convergence analysis of proposed method. Numerical implements show that our proposed model and method can efficiently improve the numerical results compared with the Rudin–Osher–Fatemi model.

Keywords

Smoothing total variation, primal dual method, image denoising, convex conjugation

Date received: 30 October 2015; accepted: 15 March 2016

Introduction

The degradation of an image is usually unavoidable during its acquisition, and perhaps the most fundamental image restoration problem is denoising. It forms a significant preliminary step in many machine vision tasks such as object detection and recognition. A major concern in designing image denoising models is to preserve important image features such as edges while removing noise. Variational PDE based models[1–3] have been widely used over the past decades for image denoising with edge preservation. A classical variational denoising method is the total variation (TV)-based minimizing process proposed by Rudin–Osher–Fatemi (ROF),[4] where they used the properties that TV can reduce oscillations and regularize the geometry of level sets without penalizing discontinuities. Formally, the ROF can be written as

$$\min_u \frac{\lambda}{2}\int_\Omega (u-f)^2 \mathrm{d}x + |u|_{TV} \tag{1}$$

where the regularization weight $\lambda > 0$ should be tuned to match the noise level contaminating f and $|u|_{TV} := \int_\Omega |\nabla u| \mathrm{d}x$ can be extended to the space $BV(\Omega)$ that contains functions with discontinuities. Here, Ω denotes the image domains.

However, from the view of numerical computation, solving (1) is a difficult task to find a good method since its corresponding Euler–Lagrangian equation includes the term $\frac{\nabla u}{|\nabla u|}$, which is degenerated when $|\nabla u| = 0$. Especially, this fact usually happens in the flat areas of image. A lot of methods have been proposed to overcome this drawback. The first class of techniques is to slightly perturb the TV-norm functional to get

$$\min_u \underbrace{\frac{\lambda}{2}\int_\Omega (u-f)^2 \mathrm{d}x + \int_\Omega \sqrt{|\nabla u|^2+\alpha}\mathrm{d}x}_{E(u,\alpha)} \tag{2}$$

where α is a small positive parameter, then to solve the associated Euler–Lagrange equation,

$$g(u) := \lambda(u-f) - \mathrm{div}\left(\frac{\nabla u}{\sqrt{|\nabla u|^2+\alpha}}\right) = 0 \tag{3}$$

with the homogeneous Neumann boundary condition. A number of methods have been proposed to solve the equation (3). Rudin et al.[4] used a time marching scheme to reach the steady state of the parabolic equation $u_t = g(u)$ with the initial condition f. It has been

[1]School of Information and Communication Engineering, Dalian University of Technology, Dalian, China
[2]College of Mathematics and Statistics, Henan University, Kaifeng, China

Corresponding author:
Baoli Shi, College of Mathematics and Statistics, Henan University, Kaifeng 475004, China.
Email: shibaoli1983@163.com

proved to be quite efficient for regularizing images without smoothing the boundaries of the objects. However, the convergence is usually slow due to the Courant–Friedrichs–Lewy (CFL) condition, especially in flat regions where $|\nabla u| \approx 0$. To completely get rid of CFL conditions, Vogel and Oman[5] proposed a lagged diffusion fixed point iteration scheme (FP) to solve the stationary Euler–Lagrange directly, where the $(n+1)$-th iteration is obtained by solving the sparse linear equation

$$\lambda\left(u^{n+1}-f\right)-\operatorname{div}\left(\frac{\nabla u^{n+1}}{\sqrt{|\nabla u^{n}|^{2}+\alpha}}\right)=0$$
$$\Rightarrow u^{n+1}=(\lambda I+L(u^{n}))^{-1}\lambda f$$

by defining $L(u^n) := -\frac{\nabla}{\sqrt{|\nabla u^n|^2+\alpha}}$. They showed that the fixed point iteration converged to the minimizer of the functional from the ROF model. However, inverse matrices still need to be computed in this method. There are some other methods based on using the dual problem of the original problem (1) such as semi-gradient and the primal dual method (PDM)[6–8] or the splitting scheme to overcome the nonsmoothing operator norm $|\nabla \cdot|_1$ such as the alternating direction method of multipliers (ADMM).[9–12]

This article focuses on the problem (2) by assuming $0 \le \alpha \le 1$. It is easy to find that $\int_\Omega \sqrt{|\nabla u|^2+\alpha}\mathrm{dx}$ as $\alpha \to 0$ tends to the original TV norm and $\int_\Omega \sqrt{|\nabla u|^2+\alpha}\mathrm{dx}$ as $\alpha \to 1$ tends to the surface area $\int_\Omega \sqrt{|\nabla u|^2+1}\mathrm{dx}$. Furthermore, using the Taylor approximation $\sqrt{1+x} \approx 1+x/2$, we have

$$\int_\Omega \sqrt{|\nabla u|^2+\alpha}\mathrm{dx} \approx \int_\Omega \left(\sqrt{\alpha}+|\nabla u|^2/2\right)\mathrm{dx} \qquad (4)$$

Then the Euler–Lagrange equation of problem (2) approximates to Laplace equation in flat region of image where $|\nabla u| \approx 0$. In this case, the problem (2) can reduce some circumstances such as staircasing effect. Actually the problem (2) was introduced in literature,[13,14] where they considered the image as embedded maps and minimal surfaces. Then they solved it in the sense of Laplace–Beltrami operator equation. Different from these aforementioned methods, we here propose the PDM to solve the problem (2). In this method, each iteration updates both a primal and a dual variable. It is thus able to avoid some of the difficulties that arise when working only on the primal or dual side.[4,8] Therefore, we first employ the Legrndre–Fenchel transform to transform the problem (2) to a saddle point problem. Then, we can solve it based on the PDM. Based on the triviality following the work in Chambolle and Pock,[8] the convergence of the proposed PDM can be kept. Our numerical results demonstrate that the algorithm instances achieve the solution of difficult convex optimization problems in reasonable time and without parameter tuning of algorithm.

The contents of the article are organized as follows. In "PDM" section, we give some preliminaries of the PDM and use it to solve our proposed model. We arrange some numerical implementations in "Numerical implementations" section. Some concluding remarks are presented in "Conclusions" section.

PDM

PDM was first presented by Arrow et al.[15] and named as the primal-dual hybrid gradient (PDHG) method in Arrow et al.[15] and Zhu et al.[16] This method enjoys to solve the primal variable and the dual variable alternatively and then widely applied to compute the saddle point of min–max problems.[17–19] A gradient descent method is employed to find the primal and the dual variables alternatively in Zhu et al.[20] and Esser et al.[21] A predictor-corrector scheme is used in the alternating direction iterations for finding the dual variable in Chen and Teboulle.[22] Recently, Chambolle and Pock[8] presented an unified form of PDM. They demonstrated that, with a properly specified step-size policy and an averaging scheme, this method can achieve the $O(1/\mathcal{N})$ rate of convergence, where $\mathcal{N}$ is total number of iterations. In this section, we will apply the algorithm in the problem (2) to seek the saddle point of our min–max problem. We therefore first give a brief introduction of the method here and then propose it to solve the problem (2).

PDM

Let X and Y be two finite-dimensional real vector spaces equipped with the Euclidian inner product $\langle\cdot,\cdot\rangle$ and the norm $\|\cdot\| = \langle\cdot,\cdot\rangle^{\frac{1}{2}}$. Define the operator norm as

$$\|\mathcal{A}\| = \max\{\|\mathcal{A}x\|, x \in X \text{ with } \|x\| \le 1\}$$

where $\mathcal{A}: X \to Y$.

Assume that functions $F: X \to \mathbb{R} \cup \{\infty\}$ and $G: Y \to \mathbb{R} \cup \{\infty\}$ are proper convex function, we consider the following problem:

$$\min_x F(x) + G(Ax) \qquad (5)$$

where A is a matrix operator. Based on the definition of convex conjugation as

$$H^*(s) = \sup_x \left\{\langle s,t\rangle - H(t)\right\}$$

for a function $H : X \to \mathbb{R} \cup \{\infty\}$, the dual problem of the problem (5) can be written as

$$\max_y -F^*(y) - G^*(-A^T y) \tag{6}$$

Similar to the work in Chambolle and Pock,[8] the primal problem (5) and its dual problem (6) can be connected by a generic saddle-point problem as

$$\min_x \max_y \mathcal{F}(x, y) := \langle Ax, y\rangle + F(x) - G^*(y), \quad \forall x \in X, y \in Y \tag{7}$$

Furthermore, the above fact holds if and only if $A^T y^* \in \partial F(x^*)$ or also $Ax^* \in \partial G(y^*)$. Thus, the problem (5) can be solved by the primal dual algorithm[8]

$$\begin{cases} x^{k+1} = (I + \tau\partial F)^{-1}(x^k - \tau A^T y^k), & \text{(8a)} \\ \bar{x}^k = 2x^{k+1} - x^k, & \text{(8b)} \\ y^{k+1} = (I + \sigma\partial G^*)^{-1}(y^k - \sigma A\bar{x}^k), & \text{(8c)} \end{cases}$$

where $\tau > 0$ and $\sigma > 0$ are step sizes of the primal and dual variables, respectively.

Obviously, the strategy (8a)–(8c) can be regarded as using the gradient descent method to solve the primal problem

$$\min_x \langle Ax, y\rangle + F(x)$$

and using the gradient ascent method to solve the dual problem

$$\max_y \langle Ax, y\rangle - G^*(y)$$

with the modified step (8b), where the step length are chosen as τ and σ corresponding to the variables x and y, respectively.

Theorem 2.1 *Assume that the min–max problem (7) has a saddle point* (x^*, y^*) *and* $\tau\sigma\|A\|^2 < 1$, *then the sequence* $\{(x^n, y^n)\}$ *generated by the strategy (8a)–(8c) converges to* (x^*, y^*) *by choosing some suitable original values* $(x^0, y^0) \in X \times Y$ *and* $\bar{x} = x^0$.

For the proof, see Theorem 1 in Chambolle and Pock.[8] Obviously, the PDM is matrix inversion-free if the functions $F(x)$ and $G^*(y)$ are separable for the variables x and y, respectively. So this method can be used to solve large scale problem such as machine learning and medical image processing problem.[23,24]

The problem (2) based on the PDM

In this section, we consider solving the problem (2) based on the PDM. For simplicity, we assume that the image region Ω is squared. The problem (2) is computed numerically by projecting an image of $N = n \times n$ pixels. Set related Euclidean spaces $S = R^{n\times n}$ and $T = S \times S$. The usual scalar product can be denoted as $\langle v, w\rangle_S := \sum_{i=1}^n \sum_{j=1}^n v_{i,j} w_{i,j}$ for $v, w \in S$. If $u \in S$, we use $\nabla u = (\nabla_x^+ u, \nabla_y^+ u) \in T$ to denote the first-order forward difference operators with

$$\nabla_x^+ u_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j} & \text{for } 1 \le i < N \\ 0 & \text{for } i = N \end{cases} \quad \text{and}$$

$$\nabla_y^+ u_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j} & \text{for } 1 \le j < N \\ 0 & \text{for } j = N \end{cases}$$

and use $\mathrm{div}(p, q) = \nabla_x^- p + \nabla_y^- q$ to denote the first-order backward difference operators with

$$\nabla_x^- p_{i,j} = \begin{cases} -p_{i,j}, & \text{for } i = 1 \\ p_{i,j} - p_{i-1,j} & \text{for } 1 < i < N \\ p(i,j) & \text{for } i = N \end{cases} \quad \text{and}$$

$$\nabla_y^- p_{i,j} = \begin{cases} -p_{i,j} & \text{for } j = 1 \\ p_{i,j} - p_{i,j-1} & \text{for } 1 < j < N \\ p_{i,j} & \text{for } j = N \end{cases}$$

We also use the following notational brevities to represent ℓ^1 and ℓ^∞ norms as:

$$|\nabla u_{i,j}|_1 = \sqrt{\nabla_x u_{i,j}^2 + \nabla_y u_{i,j}^2} \quad \text{and}$$
$$|\mathbf{q}|_\infty = \max_{1\le i,j\le n} \{|\mathbf{q}_{i,j}|\}$$

for $|\mathbf{q}_{i,j}| = \sqrt{(q_{i,j}^x)^2 + (q_{i,j}^y)^2}$ with $\mathbf{q} \in T$. Based on these notations, we denote the discrete equivalent of (2) as

$$\min_u \left\{ \sum_{i=1}^n \sum_{j=1}^n \left(\frac{\lambda}{2}(u_{i,j} - f_{i,j})^2 + \sqrt{|\nabla u_{i,j}|^2 + \alpha} \right) \right\} \tag{8}$$

where $|\nabla u_{i,j}|^2 = \left(\nabla_x^+ u_{i,j}\right)^2 + (\nabla_y^+ u_{i,j})^2$. To solve the problem (5), we set $X = S$, $Y = T$, and

$$F(u) = \sum_{i=1}^n \sum_{j=1}^n \frac{\lambda}{2}(u_{i,j} - f_{i,j})^2 \quad \text{and}$$
$$G(Au) = \sum_{i=1}^n \sum_{j=1}^n \sqrt{|\nabla u_{i,j}|^2 + \alpha}$$

where $G(\cdot) = |\cdot|_1$ and $A = \nabla$. Using the property of the convex conjugate of a proper convex function,[25] as

$$g(t) = \sqrt{\alpha + t^2}$$
$$\Leftrightarrow g^{**}(t) = \sup_{|s|\le 1} \left\{ \langle t, s\rangle + \sqrt{\alpha(1 - s^2)} \right\}$$

then the problem (9) can be written to the min–max problem

$$\min_u \max_{|\mathbf{p}|_\infty \le 1} \left\{ \sum_{i=1}^{n}\sum_{j=1}^{n} \left(\frac{\lambda}{2}(u_{i,j} - f_{i,j})^2 - \langle u_{i,j}, \mathrm{div}\mathbf{p}_{i,j}\rangle + \sqrt{\alpha\left(1 - |\mathbf{p}_{i,j}|^2\right)} \right) \right\} \tag{9}$$

where $\mathbf{p}_{i,j} = (p^x_{i,j}, p^y_{i,j}) \in T$. Since the subjective function in (10) is proper convex, we can transfer the order of min and max. Thus, we can use the primal dual scheme (8) to solve it.

- For the original variation u: In the problem (10), ignoring the unrelated term to u, we can obtain

$$\min_u \left\{ \sum_{i=1}^{n}\sum_{j=1}^{n} \left(\frac{\lambda}{2}(u_{i,j} - f_{i,j})^2 - \langle u_{i,j}, \mathrm{div}\mathbf{p}_{i,j}\rangle \right) \right\}$$

  whose optimality condition is

$$\lambda\left(u_{i,j} - f_{i,j}\right) - \mathrm{div}\mathbf{p}_{i,j} = 0$$

- for $1 \le i,j \le n$. Then, we can use gradient ascent method to solve it as

$$\begin{aligned} u^k_{i,j} - u^{k+1}_{i,j} &= \tau\left(\lambda\left(u^{k+1}_{i,j} - f_{i,j}\right) - \mathrm{div}\mathbf{p}_{i,j}\right) \\ \Rightarrow u^{k+1}_{i,j} &= \frac{u^k_{i,j} + \lambda\tau f_{i,j} + \tau\mathrm{div}\mathbf{p}_{i,j}}{1 + \lambda\tau} \end{aligned} \tag{10}$$

- For the dual variation $\mathbf{p}$: In the problem (10), ignoring the unrelated term to $\mathbf{p}$ and introducing an indicator function

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

where $\mathcal{K} := \{|p_{i,j}|_\infty \le 1\}$, we then have

$$\max_{\mathbf{p}} \left\{ \sum_{i=1}^{n}\sum_{j=1}^{n} \left(-\langle u_{i,j}, \mathrm{div}\mathbf{p}_{i,j}\rangle + \sqrt{\alpha\left(1 - |\mathbf{p}_{i,j}|^2\right)} \right) + \chi_{\mathcal{K}}(\mathbf{p}_{i,j}) \right\}$$

which optimality condition is

$$\nabla u_{i,j}\sqrt{1 - |\mathbf{p}_{i,j}|^2} - \sqrt{\alpha}\mathbf{p}_{i,j} + \partial\chi_{\mathcal{K}}(\mathbf{p}_{i,j}) = 0$$

for $1 \le i,j \le n$. Note that $\partial\chi_{\mathcal{K}}(\mathbf{p}_{i,j}) = 0$ since indicator function is a constant function. Here ∂ denotes the subgradient defined by $\partial\hbar(\bar{y}) := \{v | \hbar(\bar{x}) - \hbar(y) \ge (v, \bar{x} - \bar{y})\}$ at the point $\bar{y}$ for a function $\hbar$. Then we use the projection gradient method as

$$\mathbf{p}^{k+1}_{i,j} = \frac{\mathbf{p}^k_{i,j} + \sigma\left(\nabla u_{i,j}\sqrt{1 - |\mathbf{p}^k_{i,j}|^2} - \sqrt{\alpha}\mathbf{p}^k_{i,j}\right)}{\max\left\{1, \left|\mathbf{p}^k_{i,j} + \sigma\left(\nabla u_{i,j}\sqrt{1 - |\mathbf{p}^k_{i,j}|^2} - \sqrt{\alpha}\mathbf{p}^k_{i,j}\right)\right|\right\}} \tag{11}$$

to solve it.

Based on two above subproblems, we summarize the PDM to solve the problem (9) as follows.

Algorithm 2.1. PDM to the problem (9)

- Set original values $u^0 = f$ and $\mathbf{p}^0 = 0$. Let $k := 0$.
- ***Repeat***

$$\begin{cases} u^{k+1}_{i,j} = \frac{u^k_{i,j} + \lambda\tau f_{i,j} + \tau\mathrm{div}\mathbf{p}^k_{i,j}}{1 + \lambda\tau} \\ \bar{u}^k_{i,j} = 2u^{k+1}_{i,j} - u^k_{i,j} \\ \mathbf{p}^{k+1}_{i,j} = \dfrac{\mathbf{p}^k_{i,j} + \sigma\left(\nabla\bar{u}^k_{i,j}\sqrt{1 - \left|\mathbf{p}^k_{i,j}\right|^2} - \sqrt{\alpha}\mathbf{p}^k_{i,j}\right)}{\max\left\{1, \left|\mathbf{p}^k_{i,j} + \sigma\left(\nabla\bar{u}^k_{i,j}\sqrt{1 - \left|\mathbf{p}^k_{i,j}\right|^2} - \sqrt{\alpha}\mathbf{p}^k_{i,j}\right)\right|\right\}} \end{cases} \tag{12}$$

Until *Convergence*.

Remark 2.1 *Using the gradient method to compute the variables u and* $\mathbf{p}$ *in (11) and (12) actually corresponds to the original primal dual scheme (8). For example, when we use the gradient ascent method to obtain the original variable* $u_{i,j}$ *in (11), it is also the optimization condition of the problem*

$$\min_u \sum_{i=1}^{n}\sum_{j=1}^{n} \underbrace{\frac{\lambda}{2}\left(u_{i,j} - f_{i,j}\right)^2}_{:=\mathcal{F}(u_{i,j})} + \frac{1}{2\tau}\left(u_{i,j} - \left(u^k_{i,j} + \tau\mathrm{div}\mathbf{p}^k_{i,j}\right)\right) = 0$$

Thus, we easily find to the formulation (8a) as

$$u^{k+1}_{i,j} = (I + \tau\nabla\mathcal{F})^{-1}\left(u_{i,j} - \left(u^k_{i,j} + \tau\mathrm{div}\mathbf{p}^k_{i,j}\right)\right) \tag{13}$$

For the dual variable $\mathbf{p}_{i,j}$ *in (12), we also obtain the similar formulation as (8c) by setting* $\mathcal{G}^*(p_{i,j}) := \sqrt{\alpha(1 - |\mathbf{p}_{i,j}|^2)} + \chi_{\mathcal{K}}(\mathbf{p}_{i,j})$.

Remark 2.2 *To the problem (10), it is easy to obtain that it has an unique saddle* $(u, \mathbf{p})$ *since the subjective function in (9) is strict convex and coercive. For the PDM, it is very important to compute the operator norm of* ∇ *to keep its convergence. Actually, it is not difficult to find that* $\|\nabla\|_2^2 \le 8$.

Based on the above remarks, we have the following assertion for Algorithm 2.1.

Theorem 2.2 *Assume that* $\tau\sigma < 1/8$ *in Algorithm 2.1, then the sequence* $\{(u^k, \mathbf{p}^k)\}$ *converges to the saddle point* $(u^*, \mathbf{p}^*)$. *Furthermore,* u^* *is the solution of the problem* (9).

Numerical implementations

In numerical implementations, we will terminate the iteration if the stop conditions is satisfied in term of the ℓ^2 relative error $\max\{\|u^{k+1} - u^k\|/\|u^k\|_2, \|\mathbf{p}^{k+1} - \mathbf{p}^k\|_2/\|\mathbf{p}^k\|_2\} < 10^{-5}$ or the iteration arriving to 500. To standardize discussions, we begin by normalizing pixel values of the inputting image $\bar{f}$ to [0,255]. We do this by using the linear-stretch formula as $f = 255 \times (\bar{f} - \min(\bar{f}))|/(\max(\bar{f}) - \min(\bar{f}))$, where max and min, respectively, represent maximum and minimum of $\bar{f}$. The simulations are preformed in Matlab R20014a on a PC with an Intel Core i5 M520 at 2.40 GHz and 4 GB of memory. Furthermore, to obtain noise images, we use the matlab function "randn" to add the white Gaussian noise to the original clean image.

For numerical comparisons, we consider to compare our proposed PDM with the numerical methods such as the time marching method (TM)[4] and the lagged diffusion fixed point method (FPM) in Vogel and Oman[5] to solve the smoothing TV-based ROF model.[15] Actually, our PDM degenerates to the PDM used in Chambolle and Pock[8] to solve the ROF model while $\alpha = 0$. Apart from some default settings, like the maximum number of iterations, the related parameters in the codes are chosen by trials and errors to give the best results of the respective methods. Here, we choose three different images with the size 256×256 as testing images shown in Figure 1. To measure more quantitatively, the similarity between the recovered image and the original (supposedly noise-free) image, we use both the signal-to-noise ratio (SNR) and the structural similarity index (SSIM), the latter being well known to better reflect perceived visual quality.

Examples 3.1 *Here we consider to use the PDM, FDM, and TM to solve the smoothing TV-based model (2) by setting* $\alpha = 0.01$ *and adding white Gaussian noise to Figure 1(a) with three different deviations as 10, 15, and 20 (Table 1). From convergence curves in Figure 2, it can be noticed that all methods can efficiently suppress noise. However, the energy curve generated by the TM need much more iteration number to attain steady state due to the effect by the CFL condition. Simultaneously, we also notice that the quality of restored images are reduced and the regularization parameter* λ *in the*

Figure 1. Original images in our implementations. (a) Cameraman (b) Lena (c) Synthetic image.

Table 1. The related results in Example 3.1.

(σ, λ)	(10, 0.08)			(15, 0.175)			(20, 0.032)		
Method	SNR	SSIM	Time	SNR	SSIM	Time	SNR	SSIM	Time
TM	19.1310	0.5056	64.7656	17.3887	0.4571	90.1563	15.8731	0.4037	78.9375
PDM	20.5998	0.5519	50.2969	17.8345	0.4559	50.0781	14.1136	0.3080	42.7813
FPM	20.5990	0.5523	32.0313	17.7935	0.4542	49.2656	14.0446	0.3081	78.8594

Here (σ, λ) denotes the level of noise and regularization parameter in the problem (2).

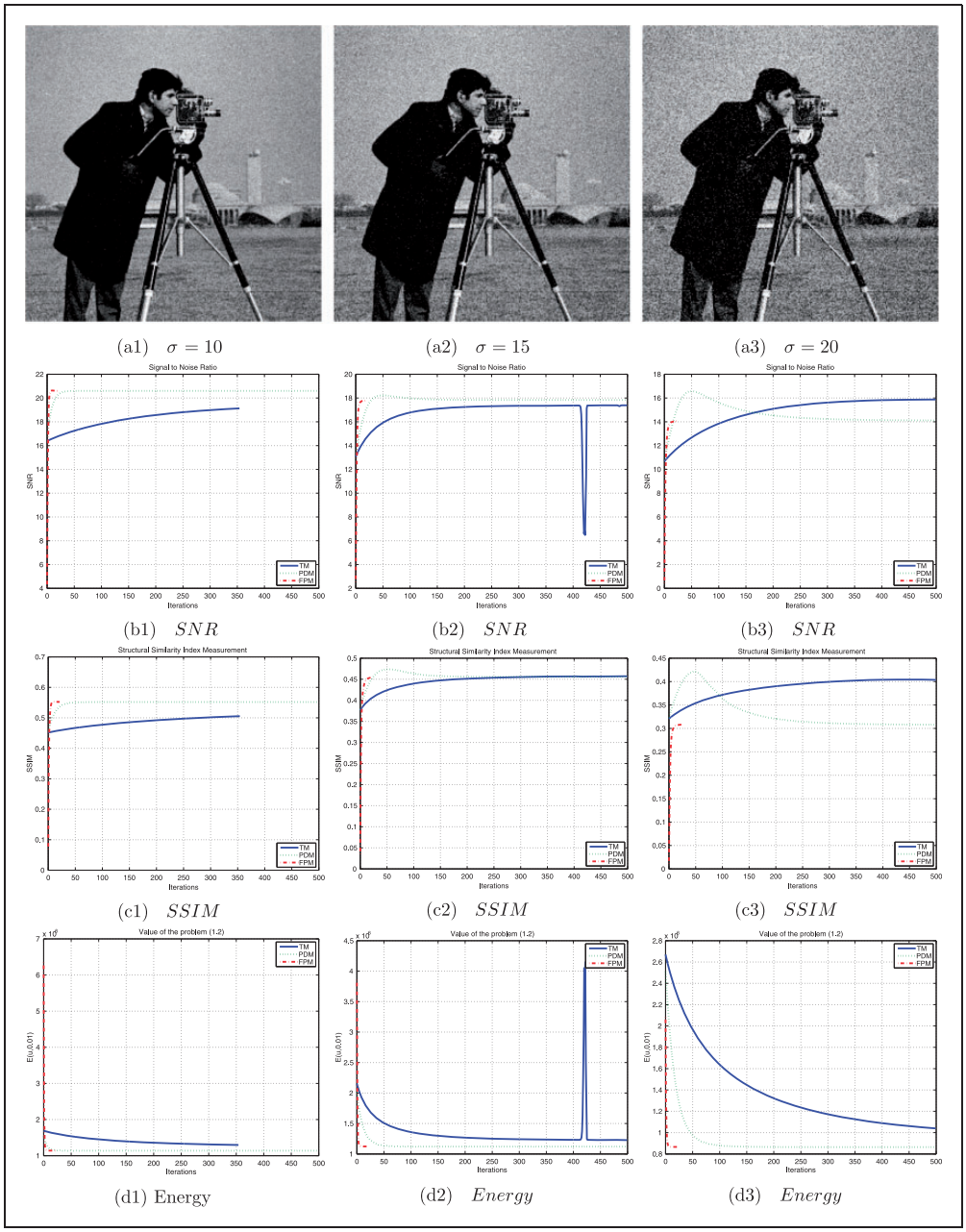

Figure 2. Numerical curves of *SNR*, *SSIM*, and the energy of the problem (2). Top-row denotes the original images by adding to different white Gaussian noise.

Table 2. The related results in Example 3.2.

Noise Parameter	σ=10			σ=20			σ=30		
	SNR	SSIM	λ	SNR	SSIM	λ	SNR	SSIM	λ
$\alpha = 0$	19.1983	0.7457	0.18	15.4873	0.6320	0.07	13.3730	0.5478	0.039
$\alpha = 0.001$	19.2249	0.7492	0.18	15.4639	0.6413	0.07	13.3582	0.5594	0.039
$\alpha = 0.01$	19.2005	0.7448	0.18	15.4999	0.6320	0.07	13.3768	0.5495	0.038
$\alpha = 0.1$	19.2091	0.7469	0.17	15.5119	0.6306	0.07	13.3844	0.5522	0.036
$\alpha = 1$	19.2206	0.7474	0.15	15.5275	0.6338	0.06	13.4102	0.5573	0.030
$\alpha = 1.5$	19.2249	0.7492	0.14	15.4639	0.6413	0.05	13.3582	0.5594	0.026

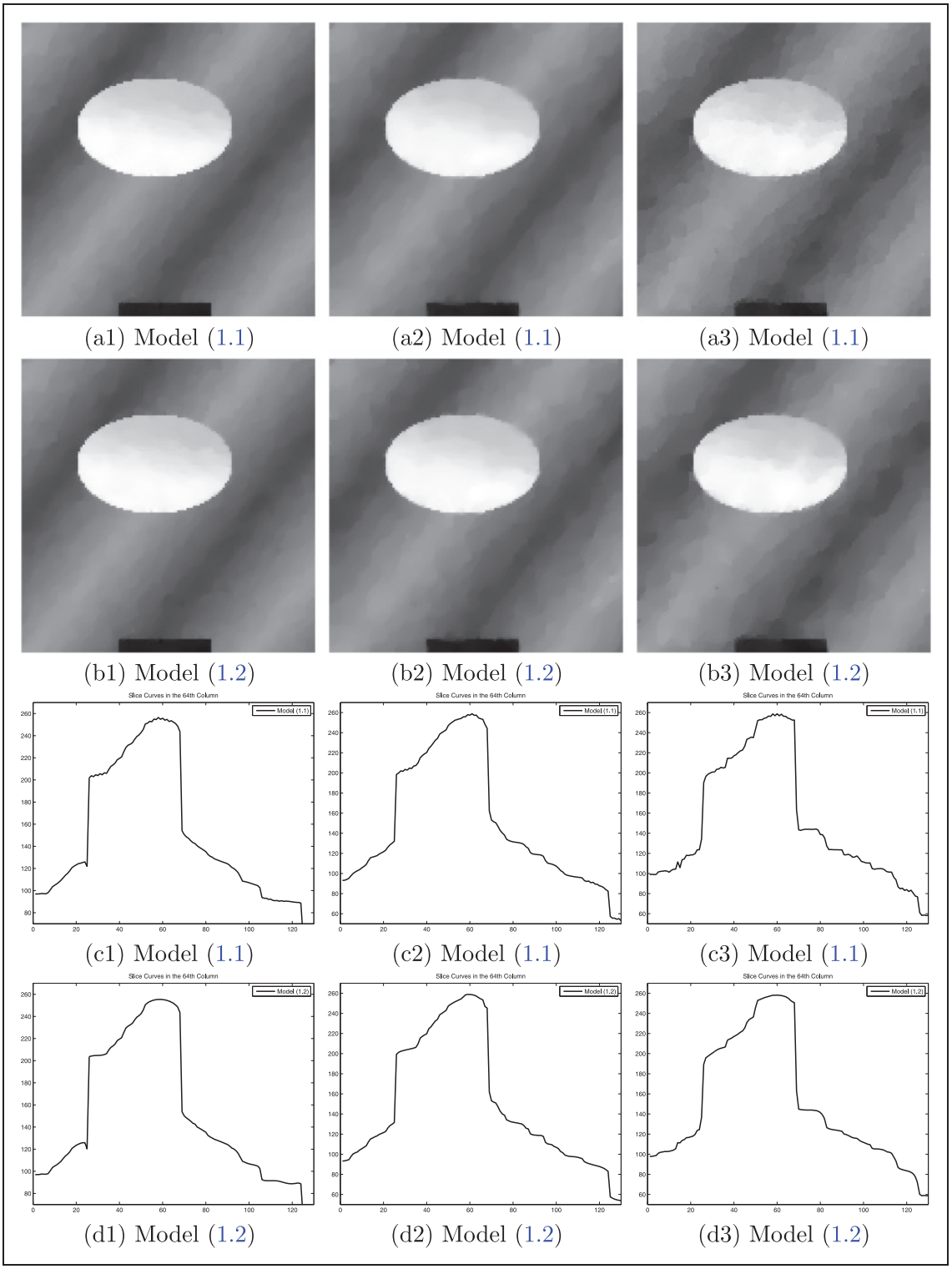

Figure 3. Numerical results by using the models (1) and (2) to solve the noisy images. To clear comparison, we only show the upper-left part of restoration images with size 128×128 and plot the slice of the 64th column. First to third columns are added to the white Gaussian noise with three different deviations, such as 10, 20, and 30.

model (2) needs to be decreased while increasing the noise in the image. That is because that the capability of reconstructing becomes harder and the regularization term $\int_\Omega \sqrt{|\nabla u|^2+\alpha}\mathrm{dx}$ *to be dominance. From the averages of CPU time, our proposed method is obviously superior to other two methods, especially for the large scale image such in image denoising problem in the medical image. Simultaneously, we notice that our method will give the best SNR and SSIM when choosing a suitable stopping condition by comparing the convergence curves in the second and third rows of Figure 2.*

Examples 3.2 *We consider the effect while choosing the different parameter* α*. Here the Lena image as the testing image is added to different noise levels with different deviations shown in Table 2. As we can see from the data, it is easy to find that the smoothing TV-based model can improve the restoration quality then the results generated by using suitable* α*. Note that it is of using the PDM to solve the original ROF model when setting* $\alpha=0$*. Furthermore, the regularization parameter* λ *is need to be reduced when we increase the smoothing parameter* α*. This is because that we need a higher* λ *to penalty* $\int_\Omega \sqrt{|\nabla u|^2+\alpha}\mathrm{dx}$ *while decreasing* α*.*

Examples 3.3 *Since the original ROF model (1) usually generates the staircase effect in the smoothing region while restoring degraded image. Simultaneously, from the introduction, we can find the smoothing TV-based ROF model (2) approximately own the diffusion ability with isotropy, that is to say* $\int_\Omega \sqrt{|\nabla u|^2+\alpha}\mathrm{dx} \approx \int_\Omega (1+|\nabla u|^2/2)\mathrm{dx}$*, when* $|\nabla u| \approx 0$*. So we expect that the model (2) can efficiently reduce staircase effect and keep image edges when choosing suitable the smoothing parameter* α*. This fact has been observed from Example 3.2 by choosing* $\alpha=1$*. Here we mainly focus on reducing the staircase effect by choosing suitable* α*, and numerical results are arranged in Figure 3. To clear comparisons, we also plot the slices of the 60th in the restored image. Obviously, the smoothing TV-based model (2) can reduce the staircase effect compared with the ROF model (2) and keep image edges, this phenomenon is especially obvious in peak regions and the corner regions in Figure 3(c) and (d).*

Conclusions

In this article, we have introduced the PDM based on the work in Chambolle and Pock[8] to solve the smoothing TV-based ROF model (2). Numerical implementations show that our method is superior to the classic methods such as the time marching method[4] and the FPM[5] on the stability and high efficiency, especially to restore the large-scale image. Furthermore, the proposed model can effectively reduce the staircase effect by choosing suitable smoothing parameter α.

Acknowledgments

The author is grateful to the anonymous referee for his/her careful checking of details and helpful comments that improved this article.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: NSF of China (No. 11401170) and Foundation of Henan Educational Committee of China (No.14A110018).

References

1. Aubert G and Kornprobst P. *Mathematical problem in image processing: partial differential equations and the calculus of variations*. New York, NY: Springer, 2008.
2. Chan TF and Shen J. *Image processing and analysis-variational, PDE, wavelet, and stochastic methods*. Philadelphia: SIAM, 2005.
3. Scherzer O. *Handbook of mathematical methods in imaging*. New York: Springer, 2015.
4. Rudin L, Osher S and Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D* 1992; 60: 259–268.
5. Vogel C and Oman M. Iterative methods for total variation denoising. *SIAM J Sci Comput* 1996; 17: 227–238.
6. Chan T, Golub G and Mulet P. A nonlinear primal-dual method for total variation-based image restoration. *SIAM J Sci Comput* 1999; 20: 1964–1977.
7. Chambolle A. An algorithm for total variation minimization and applications. *J Math Imag Vis* 2004; 20: 89–97.
8. Chambolle A and Pock T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J Math Imag Vis* 2011; 40: 120–145.
9. Boyd S, Parikh N, Chu E, et al. Distributed optimization and statistical learning via the alternating direction method of multiplier. *Found Trends Mach Learn* 2011; 3: 1–122.
10. Wu C and Tai X. Augmented Lagrangian method, dual methods, and split Bregman iteration for ROF, vectorial TV, and high order models. *SIAM J Imag Sci* 2010; 3: 300–339.
11. Setzer S. Split Bregman algorithm, Douglas-Rachford splitting and frame shrinkage. In: *Proceedings of the second international conference on scale space and variational methods in computer vision*, 2009, Voss, Norway, vol. 5567, pp.464–476.
12. Goldstein T and Osher S. The split Bregman method for L1 regularized problems. *SIAM J Imag Sci* 2009; 2: 323–343.

13. Malladi R and Sethian JA. Image processing: Flows under min/max curvature and mean curvature. *Graph Models Image Process* 1996; 58: 127–141.
14. Fallah A and Ford G. On mean curvature diffusion in nonlinear image filtering. *Pattern Recog Lett* 1998; 19: 433–437.
15. Arrow K, Hurwicz L and Uzawa H. *Studies in linear and non-linear programming*. Stanford, CA: Stanford University Press, 1958.
16. Zhu M, Wright S and Chan T. Duality-based algorithms for total-variation-regularized image restoration. *Comput Opt Appl* 2010; 47: 377–400.
17. Combettes P and Pesquet J. A proximal decomposition method for solving convex variational inverse problems. *Inverse Prob* 2008; 24: 065014.
18. Dupe F, Fadili J and Starck J. A proximal iteration for deconvolving Poisson noisy images using sparse representations. *IEEE Trans Image Process* 2009; 18: 310–321.
19. Eckstein J and Bertsekas D. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math Prog A* 1992; 55: 293–318.
20. Zhu M, Wright S and Chan T. Duality-based algorithms for total-variation-regularized image restoration. *Comput Opt Appl* 2010; 47: 377–400.
21. Esser E, Zhang X and Chan T. A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. *SIAM J Imag Sci* 2010; 3: 1015–1046.
22. Chen G and Teboulle M. A proximal-based decomposition method for convex minimization problems. *Math Prog A* 1994; 64: 81–101.
23. Sidky E, Jøgensen J and Pan X. Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm. *Phys Med Biol* 2012; 57: 3065–3091.
24. Zhang Y and Xiao L. Stochastic primal-dual coordinate method for regularized empirical risk minimization. In: *The 32nd international conference on machine learning*, Lille, France, 2015.
25. Ekeland I and Turbull T. *Infinite dimensional optimization and convexity*. Chicago, IL: The University of Chicago Press, 1983.