Parameter selection and solution algorithm for TGV-based image restoration model

Yehu Lv¹

© Springer Nature Switzerland AG 2019

Abstract

In this paper, image restoration problem is formulated to solve a total generalized variation (TGV)-based minimization problem. The minimization problem includes an unknown regularization parameter. A Morozov's discrepancy principle-based method is used to choose a suitable regularization parameter. Computationally, by introducing two dual variables, the TGV-based image restoration problem is reformulated as a convex-concave saddle-point problem. Meanwhile, the Chambolle–Pock's first-order primal–dual algorithm is transformed into a different equivalent form which can be seen as a proximal-based primal–dual algorithm. Then, the different equivalent form is used to solve the saddle-point problem. At last, compared with several existing state-of-the-art methods, experimental results demonstrate the performance of our proposed method.

Keywords Image restoration · Total generalized variation (TGV) · Morozov's discrepancy principle · Primal–dual algorithm

1 Introduction

Image restoration is a fundamental and important problem in the field of image processing. The image restoration problem is often represented by a linear model as [1]

$$g = Hu + n1, \quad (1)$$

where H is a linear blurring operator, and $n1$ represents the Gaussian white noise with variance σ^2 . Our goal is to obtain the original image u from the blurred image g . The problem of finding u from (1) is a discrete linear inverse problem. The TV-based model [2] has been proposed to solve this inverse problem. The TV-based image restoration model can be written as

$$\min_u \int_{\Omega} |\nabla u| \, dx + \frac{\lambda}{2} \|Hu - g\|_2^2, \quad (2)$$

where $\Omega \subset \mathbb{R}^n$ is a bounded domain, $n \in \mathbb{N}$, $n \geq 1$ is a fixed space dimension, ∇u represents the gradient of u , and $\lambda > 0$ is a regularization parameter.

As described in [3] and [4], the TV-based image restoration model only takes the first derivative into account, and it often causes the staircase effect. In order to eliminate the staircase effect, high-order partial differential equation filters, such as fourth-order anisotropic diffusion strategy [5–9], and four-order partial differential equation regularization-based minimization schemes [10–13] have been applied to image restoration. Bredies et al. [3] proposed the concept of TGV; the TGV provides a way of balancing between the first and second derivative of a function. Experimental results in [3] show that the TGV-based image restoration model is effective in eliminating the staircase effect. Bredies et al. [3] proposed a TGV-based image restoration model which can be written as

$$\min_{u,w} \alpha_1 \int_{\Omega} |\nabla u - w| \, dx + \alpha_0 \int_{\Omega} |\varepsilon(w)| \, dx + \frac{\lambda}{2} \|Hu - g\|_2^2, \quad (3)$$

✉ Yehu Lv, mathyehul@163.com | ¹Institute of Mathematics, Hebei University of Technology, Tianjin 300401, China.

where w represents an approximation of the first-order gradient ∇u , α_1 and α_0 are the constants, $\varepsilon(w) = \frac{1}{2}(\nabla w + \nabla w^T)$ is a symmetrized derivative.

The regularization parameter plays an important role in the TGV-based image restoration problem. If the regularization parameter is too large, the regularized solution is under-smoothed, whereas if regularization parameter is too small, the regularized solution does not fit the given data properly. Recently, Bioucas-Dias et al. [14] proposed the majorization–minimization (MM) method to estimate the regularization parameter. Liao et al. [15] utilized the generalized cross-validation (GCV) method to estimate the regularization parameter. Babacan et al. [16] proposed the parameter estimation using the variational distribution approximations, which considered the Gamma distribution for the prior. Meanwhile, the discrepancy principle has been used to the problem of TV-based regularization parameter selection [17–20].

Our main contributions in this paper can be summarized into three aspects listed as follows. First, we consider the choice of regularization parameter in the TGV-based image restoration problem. Secondly, the Morozov's discrepancy principle is used to choose a suitable regularization parameter. Thirdly, a proximal-based primal–dual algorithm is used to solve the TGV-based image restoration problem.

The outline of this paper is organized as follows. In Sect. 2, the TGV-based image restoration model is introduced. In Sect. 3, when the regularization parameter is fixed, the numerical Algorithm 1 is given, and when the regularization parameter is chosen by the Morozov's discrepancy principle, the numerical Algorithm 2 is given. In Sect. 4, experimental results are given to show that Morozov's discrepancy principle-based parameter selection method is effective in improving the restoration effect. Finally, some discussions are given in Sect. 5.

2 The TGV-based image restoration model

According to [21], the TGV-based image restoration model can be written as

$$\min_{u,w} \alpha_1 \int_{\Omega} |\nabla u - w| \, dx + \alpha_0 \int_{\Omega} |\varepsilon(w)| \, dx + \frac{\lambda}{2} \|Hu - g\|_2^2, \quad (4)$$

its discrete form is

$$\min_{u,w} \alpha_1 \|\nabla u - w\|_1 + \alpha_0 \|\varepsilon(w)\|_1 + \frac{\lambda}{2} \|Hu - g\|_2^2. \quad (5)$$

By using the Legendre–Fenchel dual transform and the computation method in [22], Eq. (5) can be rewritten as

$$\min_{u,w} \max_{p \in P, q \in Q} \langle p, \nabla u - w \rangle + \langle q, \varepsilon(w) \rangle + \frac{\lambda}{2} \|Hu - g\|_2^2 - \delta_P(p) - \delta_Q(q), \quad (6)$$

where p and q are the introduced dual variables. The indicator function $\delta_P(p)$ and $\delta_Q(q)$ are defined as

$$\delta_P(p) = \begin{cases} 0, & \text{if } p \in P, \\ \infty, & \text{otherwise,} \end{cases}, \delta_Q(q) = \begin{cases} 0, & \text{if } q \in Q, \\ \infty, & \text{otherwise.} \end{cases} \quad (7)$$

Meanwhile, the associated feasible sets of two variables are defined as

$$P = \{p = (p_1, p_2)^T \mid \|p\|_{\infty} \leq \alpha_1\}, \\ Q = \left\{q = \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \mid \|q\|_{\infty} \leq \alpha_0\right\}, \quad (8)$$

where $\|p\|_{\infty} = \max_{ij} |p_{ij}|$, with $|p_{ij}| = \sqrt{(p_1)_{ij}^2 + (p_2)_{ij}^2}$, and $\|q\|_{\infty} = \max_{ij} |q_{ij}|$, with $|q_{ij}| = \sqrt{(q_{11})_{ij}^2 + (q_{12})_{ij}^2 + (q_{21})_{ij}^2 + (q_{22})_{ij}^2}$. Here, the symbol T represents the transpose operation. For $u, v \in \mathbb{R}^{m \times n}$, the inner product and induced norm are defined as

$$\langle u, v \rangle = \sum_{i=1}^m \sum_{j=1}^n u_{ij} v_{ij}, \quad \|u\|_2 = \left(\sum_{i=1}^m \sum_{j=1}^n u_{ij}^2 \right)^{1/2}. \quad (9)$$

The first-order forward and backward difference operators with suitable boundary conditions are defined as

$$(\partial_1^+ u)_{ij} = \begin{cases} u_{i+1,j} - u_{ij}, & \text{if } 1 \leq i < m, \\ 0, & \text{if } i = m, \end{cases} \\ (\partial_2^+ u)_{ij} = \begin{cases} u_{i,j+1} - u_{ij}, & \text{if } 1 \leq j < n, \\ 0, & \text{if } j = n, \end{cases} \quad (10)$$

$$(\partial_1^- u)_{ij} = \begin{cases} u_{ij} - u_{i-1,j}, & \text{if } 1 < i < m, \\ u_{ij}, & \text{if } i = 1, \\ -u_{i-1,j}, & \text{if } i = m, \end{cases} \\ (\partial_2^- u)_{ij} = \begin{cases} u_{ij} - u_{i,j-1}, & \text{if } 1 < j < n, \\ u_{ij}, & \text{if } j = 1, \\ -u_{i,j-1}, & \text{if } j = n. \end{cases} \quad (11)$$

The discrete gradient operator $\nabla : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n \times 2}$ is defined as $(\nabla u)_{ij} = ((\partial_1^+ u)_{ij}, (\partial_2^+ u)_{ij})$, and the symmetrized derivative $\varepsilon : \mathbb{R}^{m \times n \times 2} \rightarrow \mathbb{R}^{m \times n \times 4}$ can be separately expressed as

$$\varepsilon(w) = \frac{1}{2}(\nabla w + \nabla w^T) = \begin{pmatrix} \partial_1^+ w_1 & \frac{1}{2}(\partial_2^+ w_1 + \partial_1^+ w_2) \\ \frac{1}{2}(\partial_1^+ w_2 + \partial_2^+ w_1) & \partial_2^+ w_2 \end{pmatrix}. \quad (12)$$

If the discretized divergence operator div is defined as $\text{div} = -\nabla^*$, where ∇^* is the adjoint of ∇ , then for $u \in \mathbb{R}^{m \times n}$, $p \in \mathbb{R}^{m \times n \times 2}$, we have $\langle \nabla u, p \rangle = -\langle u, \text{div}(p) \rangle$. If the discretized divergence operator div^h is defined as $\text{div}^h = -\varepsilon^*$, where ε^* is the adjoint of ε , then for $w \in \mathbb{R}^{m \times n \times 2}$, $q \in \mathbb{R}^{m \times n \times 4}$, we have $\langle \varepsilon(w), q \rangle = -\langle w, \text{div}^h(q) \rangle$. Thus, $\text{div}(p)$ is defined as $\text{div}(p) = \partial_1^- p_1 + \partial_2^- p_2$, and $\text{div}^h(q)$ is defined as

$$\text{div}^h(q) = \begin{pmatrix} \partial_1^- q_{11} + \partial_2^- q_{12} \\ \partial_1^- q_{21} + \partial_2^- q_{22} \end{pmatrix}. \quad (13)$$

The norm $\|\cdot\|_1$ in the space $\mathbb{R}^{m \times n \times k}$ reflects that for $w \in \mathbb{R}^{m \times n \times k}$, we consider w_{ij} as a vector in $\mathbb{R}^k$ on which we use the Euclidean norm

$$\|w\|_1 = \sum_{i=1}^m \sum_{j=1}^n |w_{ij}| \quad (14)$$

$$\text{with } |w_{ij}| = \left(\sum_{l=1}^k w_{ij,l}^2 \right)^{1/2}.$$

3 Two numerical algorithms

In this section, the Chambolle–Pock’s first-order primal–dual algorithm is firstly transformed into a different equivalent form which can be seen as a proximal-based primal–dual algorithm. Then, two proximal-based primal–dual algorithms are proposed. On the one hand, when the regularization parameter λ is fixed, Algorithm 1 is given. On the other hand, when the regularization parameter λ is chosen by the Morozov’s discrepancy principle, Algorithm 2 is given.

3.1 The transformed Chambolle–Pock’s first-order primal–dual algorithm

The primal–dual algorithm [23–28] has efficient iterative pattern, and it has been used in the field of image processing. In 2011, Chambolle et al. [29] proposed the first-order primal–dual algorithm to solve the saddle-point problem

$$\min_x \max_y \langle Kx, y \rangle - F^*(y) + G(x), \quad (15)$$

where x is primal variable, y is dual variable, F and G are proper, convex, lower-semicontinuous functions, F^* is the convex conjugate of F , and $\langle \cdot, \cdot \rangle$ represents the inner product, K is a linear operator. Here, the Chambolle–Pock’s first-order primal–dual algorithm is transformed into a different equivalent form which can be written as

$$\begin{aligned} x^{n+1} &= \arg \min_x G(x) + \langle Kx, y^n \rangle + \frac{1}{2s} \|x - x^n\|_2^2, \\ \hat{x}^{n+1} &= x^{n+1} + \theta (x^{n+1} - x^n), \\ y^{n+1} &= \arg \max_y \langle K\hat{x}^{n+1}, y \rangle - F^*(y) - \frac{1}{2t} \|y - y^n\|_2^2, \end{aligned} \quad (16)$$

where the parameters $s, t > 0$ are step sizes of the primal and dual variables, respectively, and θ is a fixed parameter. According to [29], it has also enjoyed convergence with rate $O(1/n)$ when $\theta = 1$ and the step sizes satisfy $s, t < 1/\|K\|_2^2$.

3.2 The first numerical algorithm

The transformed Chambolle–Pock’s first-order primal–dual algorithm can be seen as a proximal-based primal–dual algorithm. In order to apply the proximal-based primal–dual algorithm to (6), Eq. (6) is needed to be formulated as the form of (15). Denote

$$K = \begin{pmatrix} \nabla & -I \\ 0 & \varepsilon \end{pmatrix}, x = \begin{pmatrix} u \\ w \end{pmatrix}, y = \begin{pmatrix} p \\ q \end{pmatrix}, \quad (17)$$

$F^*(y) = \delta_p(p) + \delta_q(q)$, $G(x) = \frac{\lambda}{2} \|Hu - g\|_2^2$, where I is the identity matrix. It is easy to check that (6) completely fits into the framework of (15).

When the regularization parameter λ is fixed, the resulting proximal-based primal–dual algorithm is summarized into Algorithm 1.

firstly introduced. Then, when the regularization parameter is chosen by the Morozov's discrepancy principle, the numerical Algorithm 2 is given.

Algorithm 1 Proximal-Based Primal-Dual Algorithm (λ is fixed)

Step 1 Initialization: choose $s > 0, t > 0, u^0, w^0, p^0, q^0$, let $\theta = 1$;

Step 2 Compute:

$$\left\{ \begin{array}{l} u^{n+1} = \arg \min_u \langle p^n, \nabla u - w \rangle + \frac{\lambda}{2} \|Hu - g\|_2^2 + \frac{1}{2s} \|u - u^n\|_2^2 \\ w^{n+1} = \arg \min_w \langle p^n, \nabla u - w \rangle + \langle q^n, \varepsilon(w) \rangle + \frac{1}{2s} \|w - w^n\|_2^2 \\ \hat{u}^{n+1} = u^{n+1} + \theta (u^{n+1} - u^n) \\ \hat{w}^{n+1} = w^{n+1} + \theta (w^{n+1} - w^n) \\ p^{n+1} = \arg \max_{p \in P} \langle p, \nabla \hat{u}^{n+1} - \hat{w}^{n+1} \rangle - \delta_P(p) - \frac{1}{2t} \|p - p^n\|_2^2 \\ q^{n+1} = \arg \max_{q \in Q} \langle q, \varepsilon(\hat{w}^{n+1}) \rangle - \delta_Q(q) - \frac{1}{2t} \|q - q^n\|_2^2 \\ u^n = u^{n+1}, w^n = w^{n+1}, p^n = p^{n+1}, q^n = q^{n+1}, \end{array} \right. \quad (18)$$

Step 3 When the stopping criterion is fulfilled, the iterative process is stopped; otherwise, let $n = n + 1$, then go to Step 2.

In Algorithm 1, the regularization parameter λ is only included in the subproblem for u ; the subproblem for u can be written as

$$\begin{aligned} u^{n+1} &= \arg \min_u \langle p^n, \nabla u - w \rangle \\ &\quad + \frac{\lambda}{2} \|Hu - g\|_2^2 + \frac{1}{2s} \|u - u^n\|_2^2 \\ &= \arg \min_u \langle p^n, \nabla u \rangle \\ &\quad + \frac{\lambda}{2} \|Hu - g\|_2^2 + \frac{1}{2s} \|u - u^n\|_2^2. \end{aligned} \quad (19)$$

Its optimality condition is

$$-\operatorname{div}(p^n) + \lambda H^T (Hu - g) + \frac{1}{s} (u - u^n) = 0. \quad (20)$$

Thus, we have

$$u^{n+1} = (I + \lambda s H^T H)^{-1} (u^n + s(\operatorname{div}(p^n) + \lambda H^T g)). \quad (21)$$

3.3 The Morozov's discrepancy principle

Here, the basic theory of choosing the regularization parameter by the Morozov's discrepancy principle is

According to the Morozov's discrepancy principle, we can adjust u is always in a feasible region D . The feasible region D can be written as

$$D = \{u : \|Hu - g\|_2^2 \leq c^2\}, \quad (22)$$

where $c^2 = \rho n_1 n_2 \sigma^2$, and $\rho \in (0, 1]$ is default parameter. According to [30], one sets $\rho = 1$. $n_1 \times n_2$ represents the size of image, σ^2 is the variance of the noise. If the variance σ^2 of the noise is unknown, it can be estimated using the median rule [31, 32].

When the regularization parameter λ is fixed, according to (21), one can see that

$$u^{n+1} = (I + \lambda s H^T H)^{-1} (u^n + s(\operatorname{div}(p^n) + \lambda H^T g)). \quad (23)$$

Similarly, if the regularization parameter λ is chosen by the Morozov's discrepancy principle, we can let

$$\begin{aligned} u^{n+1} &= (I + \lambda_{n+1} s H^T H)^{-1} (u^n + s(\operatorname{div}(p^n) \\ &\quad + \lambda_{n+1} H^T g)), \end{aligned} \quad (24)$$

which can also obtained by (32). If we let

$$\begin{aligned} e_{n+1} &= Hu^{n+1} - g = H(I + \lambda_{n+1} s H^T H)^{-1} (\tilde{u}^n \\ &\quad + s \lambda_{n+1} H^T g) - g \end{aligned} \quad (25)$$

with $\tilde{u}^n = u^n + s \operatorname{div}(p^n)$. Using identity $H(I + \lambda_{n+1} s H^T H)^{-1} = (\lambda_{n+1} s H H^T + I)^{-1} H$, we obtain

$$e_{n+1} = (\lambda_{n+1} s H H^T + I)^{-1} (H \tilde{u}^n - g). \quad (26)$$

Define the function $\kappa(\lambda, \tilde{u})$ as

$$\kappa(\lambda, \tilde{u}) = \left\| (\lambda s H H^T + I)^{-1} (H \tilde{u} - g) \right\|_2^2. \quad (27)$$

It is obvious that $\kappa(\lambda_{n+1}, \tilde{u}^n) = \|e_{n+1}\|_2^2$. By directly computing the first and second-order derivatives of $\kappa(\lambda, \tilde{u})$ with respect to λ , it is easy to certify that $\kappa(\lambda, \tilde{u})$ is a strictly positive and strictly monotonically decreasing strictly convex function of λ . Moreover, combining with the strictly convexity of the function $\kappa(\lambda, \tilde{u})$, the equation $\kappa(\lambda, \tilde{u}) = n_1 n_2 \sigma^2$ has a unique solution when

$$\|r_0\|_2^2 \leq n_1 n_2 \sigma^2 < \kappa(0, \tilde{u}) = \|H \tilde{u} - g\|_2^2, \quad (28)$$

where r_0 denotes the orthogonal projection of $H \tilde{u} - g$ onto the null space of $H H^T$. Thus, when $\|H \tilde{u}^n - g\|_2^2 > n_1 n_2 \sigma^2$, i.e., $\tilde{u}^n \notin D$, the equation $\kappa(\lambda, \tilde{u}^n) = n_1 n_2 \sigma^2$ has a unique solution, which means that a unique solution $\lambda_{n+1} > 0$ can be found to satisfy $\|H u^{n+1} - g\|_2^2 = n_1 n_2 \sigma^2$. When $\|H \tilde{u}^n - g\|_2^2 \leq n_1 n_2 \sigma^2$, i.e., $\tilde{u}^n \in D$, we can simply let $\lambda_{n+1} = 0$. This leads to $u^{n+1} = \tilde{u}^n$, see (25) and (26). Therefore, there always exists a unique $\lambda_{n+1} \geq 0$ such that $u^{n+1} \in D$. Here, the Newton method is used to solve the nonlinear equation $\kappa(\lambda_{n+1}, \tilde{u}^n) = n_1 n_2 \sigma^2$.

3.4 The second numerical algorithm

When the regularization parameter λ is chosen by the Morozov's discrepancy principle, the resulting proximal-based primal-dual algorithm is summarized into Algorithm 2.

Algorithm 2 Proximal-Based Primal-Dual Algorithm (λ is chosen by the Morozov's discrepancy principle)

Step 1 Initialization: choose $s > 0, t > 0, u^0, w^0, p^0, q^0$, let $\theta = 1$;

Step 2 Compute:

$$\left\{ \begin{array}{l} \tilde{u}^n = u^n + s \operatorname{div}(p^n); \\ \text{if } \tilde{u}^n \in D \\ \quad \lambda_{n+1} = 0; \\ \text{else} \\ \quad \text{solve } \kappa(\lambda_{n+1}, \tilde{u}^n) = c^2; \\ u^{n+1} = \arg \min_u \langle p^n, \nabla u - w \rangle + \frac{\lambda_{n+1}}{2} \|Hu - g\|_2^2 + \frac{1}{2s} \|u - u^n\|_2^2 \\ w^{n+1} = \arg \min_w \langle p^n, \nabla u - w \rangle + \langle q^n, \varepsilon(w) \rangle + \frac{1}{2s} \|w - w^n\|_2^2 \\ \hat{u}^{n+1} = u^{n+1} + \theta (u^{n+1} - u^n) \\ \hat{w}^{n+1} = w^{n+1} + \theta (w^{n+1} - w^n) \\ p^{n+1} = \arg \max_{p \in P} \langle p, \nabla \hat{u}^{n+1} - \hat{w}^{n+1} \rangle - \delta_P(p) - \frac{1}{2t} \|p - p^n\|_2^2 \\ q^{n+1} = \arg \max_{q \in Q} \langle q, \varepsilon(\hat{w}^{n+1}) \rangle - \delta_Q(q) - \frac{1}{2t} \|q - q^n\|_2^2 \\ u^n = u^{n+1}, w^n = w^{n+1}, p^n = p^{n+1}, q^n = q^{n+1}, \end{array} \right. \quad (29)$$

Step 3 When the stopping criterion is fulfilled, the iterative process is stopped; otherwise, let $n = n + 1$, then go to Step 2.

In Algorithm 2, the subproblem for u can be written as

$$\begin{aligned} u^{n+1} &= \arg \min_u \langle p^n, \nabla u - w \rangle \\ &\quad + \frac{\lambda_{n+1}}{2} \|Hu - g\|_2^2 + \frac{1}{2s} \|u - u^n\|_2^2 \\ &= \arg \min_u \langle p^n, \nabla u \rangle \\ &\quad + \frac{\lambda_{n+1}}{2} \|Hu - g\|_2^2 + \frac{1}{2s} \|u - u^n\|_2^2. \end{aligned} \quad (30)$$

Its optimality condition is

$$-\operatorname{div}(p^n) + \lambda_{n+1} H^T (Hu - g) + \frac{1}{s} (u - u^n) = 0. \quad (31)$$

Thus, we have

$$u^{n+1} = (I + \lambda_{n+1} s H^T H)^{-1} (u^n + s(\operatorname{div}(p^n) + \lambda_{n+1} H^T g)). \quad (32)$$

The subproblem for w can be written as

$$\begin{aligned} w^{n+1} &= \arg \min_w \langle p^n, \nabla u - w \rangle + \langle q^n, \varepsilon(w) \rangle \\ &\quad + \frac{1}{2s} \|w - w^n\|_2^2 \\ &= \arg \min_w \langle p^n, -w \rangle + \langle q^n, \varepsilon(w) \rangle \\ &\quad + \frac{1}{2s} \|w - w^n\|_2^2. \end{aligned} \quad (33)$$

Its optimality condition is

$$-p^n - \operatorname{div}^h(q^n) + \frac{1}{s} (w - w^n) = 0. \quad (34)$$

Thus, we have

$$w = w^n + s(p^n + \operatorname{div}^h(q^n)). \quad (35)$$

The subproblem for the dual variable p can be written as

$$\begin{aligned} p^{n+1} &= \arg \max_{p \in P} \langle p, \nabla \hat{u}^{n+1} - \hat{w}^{n+1} \rangle \\ &\quad - \delta_p(p) - \frac{1}{2t} \|p - p^n\|_2^2 \\ &= \arg \min_{p \in P} \frac{1}{2t} \|p - p^n\|_2^2 - \langle p, \nabla \hat{u}^{n+1} - \hat{w}^{n+1} \rangle \\ &= \arg \min_{p \in P} \left\| p - (p^n + t(\nabla \hat{u}^{n+1} - \hat{w}^{n+1})) \right\|_2^2. \end{aligned} \quad (36)$$

Combining with the definition of the projection operator $\mathcal{P}_P(z)$, i.e., $\mathcal{P}_P(z) = \arg \min_{p \in P} \|p - z\|_2^2$ for any $z \in \mathbb{R}^{m \times n \times 2}$, we have $p^{n+1} = \mathcal{P}_P(p^n + t(\nabla \hat{u}^{n+1} - \hat{w}^{n+1}))$. The Euclidean projector $\mathcal{P}_P$ represents the projection onto the convex set P , which can be computed by

$$\mathcal{P}_P(\bar{p}^n) = \frac{\bar{p}^n}{\max(1, |\bar{p}^n|/\alpha_1)}, \quad (37)$$

where $\bar{p}^n = p^n + t(\nabla \hat{u}^{n+1} - \hat{w}^{n+1})$.

Similarly, for the dual variable q , we have

$$q^{n+1} = \mathcal{P}_Q(q^n + t \varepsilon(\hat{w}^{n+1})). \quad (38)$$

If we denote $\bar{q}^n = q^n + t \varepsilon(\hat{w}^{n+1})$, then

$$q^{n+1} = \frac{\bar{q}^n}{\max(1, |\bar{q}^n|/\alpha_0)}. \quad (39)$$

3.5 Convergence analysis

As is discussed above, Algorithm 1 and Algorithm 2 can be seen as the transformed Chambolle–Pock’s first-order primal–dual algorithms. According to [29] and [33], they enjoy convergence with rate $O(1/n)$ when $\theta = 1$ and the step sizes satisfy $s t \|K\|_2^2 < 1$. Thus, an estimate on $\|K\|_2^2$ is needed. According to the definition of the divergence operator div , it is easy to show that

$$\begin{aligned} \|\operatorname{div}(p)\|_2^2 &= \sum_{ij} ((p_1)_{ij} - (p_1)_{i-1,j} + (p_2)_{ij} - (p_2)_{i,j-1})^2 \\ &\leq 2 \sum_{ij} [((p_1)_{ij} - (p_1)_{i-1,j})^2 \\ &\quad + ((p_2)_{ij} - (p_2)_{i,j-1})^2] \\ &\leq 4 \sum_{ij} [((p_1)_{ij})^2 + ((p_1)_{i-1,j})^2 \\ &\quad + ((p_2)_{ij})^2 + ((p_2)_{i,j-1})^2] \\ &\leq 8 \|p\|_2^2, \end{aligned} \quad (40)$$

thus $\|\nabla\|_2^2 = \|\operatorname{div}\|_2^2 \leq 8$. Similarly, one can see that $\|\varepsilon\|_2^2 = \|\operatorname{div}^h\|_2^2 \leq 8$ which leads, after some computations, to the estimate $\|K\|_2^2 \leq \frac{17+\sqrt{33}}{2} < 12$.

4 Experimental results

The blurred signal-to-noise ratio (BSNR) is used to measure the quality of the observed images, and the improved signal-to-noise ratio (ISNR) is used to measure the quality of the restored images. They are defined as

$$\begin{aligned} \text{BSNR} &= 10 \log_{10} (\|g\|_2^2 / \|n\|_2^2), \\ \text{ISNR} &= 10 \log_{10} (\|g - u\|_2^2 / \|\hat{u} - u\|_2^2). \end{aligned} \quad (41)$$

Meanwhile, the signal-to-noise ratio (SNR), the peak signal-to-noise ratio (PSNR) and the mean square error (MSE) are also used to measure the quality of image restoration. They are defined as follows

Fig. 1 The first row are the original Lena, Barbara and Cameraman images (from **a** to **c**). The second row are the blurred Lena, Barbara and Cameraman images under the uniform blur of size 9×9 (from **d** to **f**). The third row are the blurred Lena, Barbara and Cameraman images under the Gaussian blur of size 9×9 with variance 9 (from **g** to **i**)

$$\begin{aligned} \text{SNR} &= 20 \log_{10} \left(\frac{\|u\|_2}{\|\hat{u} - u\|_2} \right), \quad \text{MSE} = \frac{1}{mn} \|u - \hat{u}\|_2^2, \\ \text{PSNR} &= 10 \log_{10} \left(255^2 / \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (u_{ij} - \hat{u}_{ij})^2 \right), \end{aligned} \quad (42)$$

where u denotes the original clean image, $\hat{u}$ is the restored image. Generally speaking, the larger the SNR, PSNR and ISNR, and the lower the MSE, the better the performance.

The Lena image, the Barbara image and the Cameraman image are used for testing, all with size 256×256 . The MATLAB commands `fspecial('average', 9)` and `fspecial('Gaussian', [9 9], 3)` are used to generate the uniform blur of size 9×9 and the Gaussian blur of size 9×9 with variance 9, respectively. In order to show that our method is effective in image restoration, our method is compared with the methods of [14] and [34]. Meanwhile, for the TGV-based image restoration problem, our method is also compared with the method of fixed regularization

parameter. For simplicity, the method of [14] is called as MM, the method of [34] is called as TV-LR, and the method of fixed regularization parameter is called as TGV-fixed. All experiments were run on MATLAB R2014a, and the experiments were performed under Windows 10 with a 2.5GHz Intel Core i5 processor and 4GB of RAM. In this paper, $\alpha_1 = 1, \alpha_0 = 2, s = 1/\sqrt{12}, t = 1/\sqrt{12}$ are set for Algorithm 2. $\alpha_1 = 1, \alpha_0 = 2, s = 1/\sqrt{12}, t = 1/\sqrt{12}, \lambda = 15$ are set for Algorithm 1. Algorithm 1 is the solving algorithm of the TGV-fixed method. On the one hand, these default parameters can guarantee the convergence; on the other hand, these default parameters can guarantee to obtain better restoration effect. Whether for Algorithm 2 or the MM, TV-LR and TGV-fixed methods, the stopping criterion all is

$$\frac{\|u^{n+1} - u^n\|_2}{\|u^n\|_2} \leq 10^{-4}, \quad (43)$$

or the iteration number larger than 150.

Fig. 2 The restored Lena images using four different methods under the uniform blur of size 9×9 . **a** MM **b** TV-LR **c** TGV-fixed **d** our method

Figure 1 consists of the original Lena, Barbara and Cameraman images, and the blurred Lena, Barbara and Cameraman images. In Fig. 1, the first row are the original three images, the second row are the blurred three images under the uniform blur of size 9×9 , and the third row are the blurred three images under the Gaussian blur of size 9×9 with variance 9.

The restored Lena images obtained by four different methods under the uniform blur of size 9×9 are provided in Fig. 2. Table 1 is the restoration effect comparison of Lena image under the uniform blur of size 9×9 and the Gaussian blur of size 9×9 with variance 9. By combining with Fig. 2 and Table 1, it is clear to see that our proposed method is effective in improving the restoration effect. And the SNR and PSNR values of our proposed method are 3–4 dB higher than the MM, TV-LR and TGV-fixed methods.

Similarly, the restored Barbara and Cameraman images obtained by four different methods under the uniform blur of

Fig. 3 The restored Barbara images using four different methods under the uniform blur of size 9×9 . **a** MM **b** TV-LR **c** TGV-fixed **d** our method

size 9×9 are provided in Figs. 3 and 4, respectively. Tables 2 and 3 are the restoration effect comparison of the Barbara image and the Cameraman image, respectively. The same observation can be made as described above. As might be expected, whether for the Barbara image or the Cameraman image, our method outperforms the methods of MM, TV-LR and TGV-fixed in terms of the SNR, PSNR and MSE values.

In order to show that the TGV-based image restoration model and our method can effectively eliminate the staircase effect and preserve the edge of image, the samples of the red bounding box in Fig. 3 are taken for experiments. The locally enlarged restored Barbara images using four different methods under the uniform blur of size 9×9 are provided in Fig. 5. From Fig. 5a, b, it is clear to see that the MM and TV-LR methods lead to very blocky images in the smooth regions. From Fig. 5c, it is clear to see that the TGV-fixed method has the advantage for eliminating the

Table 1 The restoration effect comparison of Lena image under the uniform blur of size 9×9 (BSNR=40) and the Gaussian blur of size 9×9 with variance 9 (BSNR=40)

Method	The uniform blur			The Gaussian blur		
	SNR	PSNR	MSE	SNR	PSNR	MSE
MM [14]	21.05	28.47	92.54	20.29	27.70	110.37
TV-LR [34]	22.96	30.37	59.71	21.62	29.03	81.27
TGV-fixed	21.32	28.73	87.02	20.51	27.92	104.93
Our method	24.68	32.10	40.12	23.34	30.76	54.60

Fig. 4 The restored Cameraman images using four different methods under the uniform blur of size 9×9 . **a** MM **b** TV-LR **c** TGV-fixed **d** our method

staircase effect, but it causes the unexpected edge blurring. Figure 5d is the locally enlarged restored Barbara image obtained by our method. It is clear to see that our method can effectively restore the patterned fabric. In a word, our method can effectively eliminate the staircase effect, and it is able to overcome the blocky images while preserving edge details.

Finally, the figure of ISNR versus iteration number obtained by Algorithm 2 is given in Fig. 6. From Fig. 6, it

Fig. 5 Locally enlarged restored Barbara images using four different methods under the uniform blur of size 9×9 . **a** MM **b** TV-LR **c** TGV-fixed **d** our method

is clear to see that as the iteration number increases, the ISNR values of three different images increase first and stabilize to a certain value at last. The figure of the regularization parameter λ versus iteration number obtained by Algorithm 2 is given in Fig. 7. Figure 7 demonstrates that our method can automatically choose the regularization parameter in each iteration. From Figs. 6 and 7, we can see that Algorithm 2 usually converges within 100 iterations,

Table 2 The restoration effect comparison of Barbara image under the uniform blur of size 9×9 (BSNR=40) and the Gaussian blur of size 9×9 with variance 9 (BSNR=40)

Method	The uniform blur			The Gaussian blur		
	SNR	PSNR	MSE	SNR	PSNR	MSE
MM [14]	20.72	26.16	157.44	18.89	24.33	239.88
TV-LR [34]	22.67	28.12	100.36	21.66	27.10	126.84
TGV-fixed	20.83	26.27	153.61	19.05	24.49	231.15
Our method	24.69	30.13	63.14	23.51	28.96	82.70

Table 3 The restoration effect comparison of Cameraman image under the uniform blur of size 9×9 (BSNR=40) and the Gaussian blur of size 9×9 with variance 9 (BSNR=40)

Method	The uniform blur			The Gaussian blur		
	SNR	PSNR	MSE	SNR	PSNR	MSE
MM [14]	23.58	29.16	78.83	21.88	27.46	116.75
TV-LR [34]	23.67	29.25	77.27	22.05	27.63	112.17
TGV-fixed	21.55	27.14	125.76	20.12	25.70	175.08
Our method	25.52	31.10	50.46	24.05	29.63	70.77

Fig. 6 ISNR versus iterations for the Lena, Barbara and Cameraman images. **a** and **b** are the experimental results under the uniform blur of size 9×9 , **c** and **d** are the experimental results under the Gaussian blur of size 9×9 with variance 9

and the smaller the value of BSNR, i.e., the higher the level of the noise, the smaller the regularization parameter λ , as one would expect. We also observe that λ_{n+1} stabilizes within 100 iterations.

5 Conclusion

In this paper, the choice of regularization parameter in the TGV-based image restoration model is considered, and the regularization parameter is chosen by the Morozov's

discrepancy principle. The TGV-based image restoration problem is reformulated as a saddle-point problem. When the regularization parameter is chosen by our proposed method, the proximal-based primal-dual algorithm is applied to solve the saddle-point problem. For the subproblems of the primal variables and dual variables, each subproblem has a closed-form solution. Therefore, the regularization parameter associated with the solution of the restored image is chosen by the Morozov's discrepancy principle in the each iteration. Experimental results show that the Morozov's discrepancy principle-based

Fig. 7 Regularization parameter versus iterations for the Lena, Barbara and Cameraman images. **a** and **b** are the experimental results under the uniform blur of size 9×9 , **c** and **d** are the experimental results under the Gaussian blur of size 9×9 with variance 9

regularization parameter selection method can effectively improve the effect of TGV-based image restoration in terms of SNR, PSNR and MSE quality, compared with several existing state-of-the-art methods.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest, whether financial or non-financial.

Human participants and animals This research did not involve human participants and animals.

References

1. Hanif M (2014) An EM-based hybrid Fourier-wavelet image deconvolution algorithm. In: IEEE international conference on image processing
2. Rudin L, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. *Phys D* 60(1–4):259–268
3. Bredies K, Kunisch K, Pock TT (2010) Total generalized variation. *SIAM J Imag Sci* 3(3):492–526
4. Chan R, Chan T, Yip A (2010) Numerical methods and applications in total variation image restoration. In: Handbook of mathematical methods in imaging, pp 1059–1094
5. Greer JB, Bertozzi AL (2004) Traveling wave solutions of fourth order PDEs for image processing. *SIAM J Math Anal* 36(1):38–68
6. Hajiaboli MR (2009) An anisotropic fourth-order partial differential equation for noise removal. In: Tai XC, Mørken K, Lysaker M,

- Lie KA (eds) Scale space and variational methods in computer vision. SSVM 2009. Lecture Notes in Computer Science, vol 5567. Springer, Berlin
7. Hajiaboli MR (2010) A self-governing fourth-order nonlinear diffusion filter for image noise removal. *IPSJ Trans Comput Vis Appl* 2:94–103
 8. Hajiaboli MR (2011) An anisotropic fourth-order diffusion filter for image noise removal. *Int J Comput Vis* 92(2):177–191
 9. You YL, Kaveh M (2000) Fourth-order partial differential equations for noise removal. *IEEE Trans Image Process* 9(10):1723–1730
 10. Chan T, Marquina A, Mulet P (2000) High-order total variation-based image restoration. *SIAM J Sci Comput* 22:503–516
 11. Chen HZ, Song JP, Tai XC (2009) A dual algorithm for minimization of the LLT model. *Adv Comput Math* 31:115–130
 12. Chen B, Cai JL, Chen WS (2012) A multiplicative noise removal approach based on partial differential equation model. *Math Probl Eng* 2012:1035–1052
 13. Jin ZM, Yang XP (2010) Analysis of a new variational model for multiplicative noise removal. *J Math Anal Appl* 362(2):415–426
 14. Bioucas-Dias JM, Figueiredo MAT, Oliveira JP (2006) Total variation based image deconvolution: a majorization-minimization approach. In: *IEEE international conference on acoustics speech and signal processing*
 15. Liao HY, Li F, Ng MK (2009) Selection of regularization parameter in total variation image restoration. *J Opt Soc Am A* 26(11):2311–2320
 16. Derin B, Rafael M, Katsaggelos AK (2008) Parameter estimation in TV image restoration using variational distribution approximation. *IEEE Trans Image Process* 17(3):326–339
 17. Aujol JF, Gilboa G (2006) Constrained and SNR-based solutions for TV-Hilbert space image denoising. *J Math Imag Vis* 26(1–2):217–237
 18. Blomgren P, Chan TF (2010) Modular solvers for image restoration problems using the discrepancy principle. *Numer Linear Algebr* 9(5):347–358
 19. Ng MK, Weiss P, Yuan XM (2010) Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods. *SIAM J Sci Comput* 32(5):2710–2736
 20. Weiss P, Blanc-Fraud L, Aubert G (2008) Efficient schemes for total variation minimization under constraints in image processing. *SIAM J One Comput* 31(3):2047–2080
 21. Florian K, Kristian B, Thomas P, Rudolf S (2011) Second order total generalized variation (TGV) for MRI. *Magn Reson Med* 65(2):480–491
 22. Bredies K (2014) Recovering piecewise smooth multichannel images by minimization of convex functionals with total generalized variation penalty. In: Bruhn A, Pock T, Tai XC (eds) *Efficient algorithms for global optimization methods in computer vision*. Lecture Notes in Computer Science, vol 8293. Springer, Berlin
 23. Ono S (2017) Primal-dual plug-and-play image restoration. *IEEE Signal Process Lett* 24(8):1108–1112
 24. Valkonen T (2013) A primal-dual hybrid gradient method for nonlinear operators with applications to MRI. *Inverse Probl* 30(5):900–914
 25. Gong C, Teboulle M (1994) A proximal-based decomposition method for convex minimization problems. *Math Program* 64(1–3):81–101
 26. Combettes PL, Pesquet JC (2008) A proximal decomposition method for solving convex variational inverse problems. *Inverse Probl* 24(6):065014
 27. FranOis-Xavier D, Fadili JM, Jean-Luc S (2009) A proximal iteration for deconvolving poisson noisy images using sparse representations. *IEEE Trans Image Process* 18(2):310–321
 28. Rockafellar RT (1976) Augmented lagrangians and applications of the proximal point algorithm in convex programming. *Math Oper Res* 1(2):97–116
 29. Chambolle A, Pock T (2011) A first-order primal-dual algorithm for convex problems with applications to imaging. *J Math Imag Vis* 40(1):120–145
 30. Galatsanos NP, Katsaggelos AK (1992) Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. *IEEE Trans Image Process* 1(3):322–336
 31. Mallat S (1999) *A wavelet tour of signal processing*, 2nd edn. Academic, San Diego, San Diego
 32. Rosenkranz T, Puder H (2012) Integrating recursive minimum tracking and codebook-based noise estimation for improved reduction of non-stationary noise. *Signal Process* 92(3):767–779
 33. Chambolle A (2004) An algorithm for total variation minimization and applications. *J Math Imag Vis* 20(1–2):89–97
 34. Andreas L (2015) Automated parameter selection for total variation minimization in image restoration. *J Math Imag Vis* 57(2):1–30

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.