

THE MODIFIED SECOND ACCELERATED PROXIMAL GRADIENT METHOD FOR TOTAL VARIATION IMAGE DEBLURRING

Daoling Lin and Wen Wang

Department of Mathematics, Jinan University, Guangzhou, 510632, P. R. China

Abstract

In this paper, we modify the second accelerated proximal gradient (APG) method proposed by Nesterov in [12] (and named by Tseng in [13]) for solving the total variation (TV) image deblurring problems. We also use the variables in [1] to decouple the problems so that we need not to solve the correlative subproblems in each iteration, borrow the idea of [19] to replace the Lipschitz constant by an appropriate positive-definite matrix, and use the continuation technique in [15] to provide good initial solutions for our algorithm.

Keywords: image deblurring, total variation, the second APG (the second accelerated proximal gradient), auxiliary variables, continuation technique.

*Corresponding author.

E-mail address: daolinglin@163.com (Daoling Lin).

Copyright © 2018 Scientific Advances Publishers

2010 Mathematics Subject Classification: 49M07.

Submitted by Omer Faruk Ertugrul.

Received February 28, 2018; Revised March 12, 2018

1. Introduction

Digital image restoration and reconstruction play an important part in various areas of applied sciences, such as medical and astronomical imaging, film restoration, and image and video coding. In the process of image acquisition and transmission, the images is often blurred and receive a variety of noises. This means that we need to restore the original clean images from the noisy blurred ones. Thus, the image deblurring is always a hotspot in the research of image processing.

The problem of estimating x from the observed blurred and noisy image b is called the image deblurring problems. In this paper, we mainly consider the TV deblurring problems of grayscale images defined on rectangle domains:

$$\min_{x \in B_{l,u}} \frac{1}{2} \|A(x) - b\|_F^2 + \lambda TV(x), \quad (1.1)$$

where $b \in \mathfrak{R}^{M \times N}$ be the observed blurred and noisy grayscale image, $x \in \mathfrak{R}^{M \times N}$ be the original image to recover, the linear mapping A is the blurring operator, $\|\cdot\|_F$ denotes the Frobenius norm, λ is the regularization parameter, $B_{l,u}$ is the constraint set of x :

$$B_{l,u} = \{x \in \mathfrak{R}^{M \times N} | l \leq x_{i,j} \leq u, \forall i, j\}, \quad (1.2)$$

and $TV(\cdot)$ is defined as

$$TV(x) \equiv \int_{\Omega} \|\nabla x\| d\Omega, \quad (1.3)$$

where Ω is the domain of the image x . In practice, we need to discretize $TV(x)$.

In the literature, two popular kinds of discrete TVs are isotropic TV and anisotropic TV [1], which are defined by

$$TV_{iso}(x) = \sum_{i=1}^M \sum_{j=1}^N \sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2}, \quad (1.4)$$

and

$$TV_{aniso}(x) = \sum_{i=1}^M \sum_{j=1}^N (|x_{i+1,j} - x_{i,j}| + |x_{i,j+1} - x_{i,j}|), \quad (1.5)$$

respectively, where we assume the reflexive boundary conditions for x :

$$\begin{aligned} x_{M+1,j} - x_{M,j} &= 0, \quad j = 1, \dots, N, \\ x_{i,N+1} - x_{i,N} &= 0, \quad i = 1, \dots, M. \end{aligned} \quad (1.6)$$

Due to the nondifferentiability and nonlinearity, the problem (1.1) is difficult to solve. Many researchers have proposed some specific algorithms such as kernel regression [2], soft thresholding approach [3, 4], wavelet approach [5], and dual approach [6, 7], etc. Besides, many scholars present some effective proximal algorithms for solving this kind of problem in recent years. The iterative shrinkage/thresholding (IST) algorithm proposed by Daubechies et al. [8] and references therein are effective methods to solve the linear inverse problems with a sparsity constraint. However, the IST method may be slow under some assumptions on linear and continuous operator A [9]. Some scholars keep improving the convergence rate of the IST algorithm. Two methods of particular interest to image deblurring problem are the fast iterative shrinkage/thresholding algorithm (FISTA) proposed by Beck and Teboulle [3] and the two-step iterative shrinkage/thresholding algorithm (TwISTA) proposed by Bioucas-Dias and Figueiredo [10]. And authors proved that the scheme of the sum of two convex functions exhibits a global convergence rate $O(1/k^2)$, where k is the iteration counter, see [4,

11] for more details. But we note that the above algorithms, (e.g., ISTA, FISTA, and TwISTA and so on) need to solve some corresponding subproblems which might lead to the convergence rate slow. To avoid solving the image denoising subproblem, Zuo and Lin [1] proposed the generalized accelerated proximal gradient (GAPG) method (two auxiliary variables are added so that the TV-image formulas are suitable for GAPG), and obtain the convergence rate of $O(k^{-2})$.

In this paper, we extend the modified second APG method to handle both the isotropic and anisotropic TV-functions. Moreover, we use the two auxiliary variables in [1] to approximate the partial derivatives and borrow the idea of [19] to replace the Lipschitz constant by an appropriate positive-definite matrix. We also perform numerical experiments on two 256×256 images (Lena and Cameraman) to test our method.

The rest of this paper is organized as follows. In the next section, we introduce the necessary notation. In Section 3, we develop the mathematical framework for the second APG schemes of [12] (and named by Tseng in [13]) and present our algorithm based on the second APG method for the TV-based regularization models. We present numerical experiments in Section 4.

2. Notation and Preliminaries

2.1. Basic knowledge

In this paper, we use $\mathfrak{R}^{M \times N}$ to denote the $M \times N$ dimensional Euclidean space with inner product $\langle x, y \rangle = \text{Tr}(xy^T)$ and the norm of the matrix $x \in \mathfrak{R}^{M \times N}$ is the Frobenius norm $\|x\|_F = \sqrt{\text{Tr}(xx^T)}$.

Next, we recall some definitions concerning subdifferential calculus and fundamental properties, which play an important role in proximal gradient algorithms. For a proper closed convex function $g : \mathfrak{R}^{M \times N} \rightarrow (-\infty, +\infty]$ and any scalar $\mu > 0$, the proximal map of μg is defined as

$$prox_{\mu}(g)(x) := \arg \min_u \{g(u) + \frac{1}{2\mu} \|x - u\|^2\}, \tag{2.1}$$

where $\arg \min$ denotes the unique minimizer, and we refer the readers to ([14], Chapter 7) for properties of the proximal operator.

2.2. The difference operator

Inspired by the idea of the variable splitting method [15-17], we define two difference operators Φ_v and Φ_h to simplify the formulas (1.4) and (1.5).

Definition 2.1. (i) We first define the vertical and horizontal forward difference operators Φ_v and Φ_h as

$$p = \Phi_v(x) : \mathfrak{R}^{M \times N} \rightarrow \mathfrak{R}^{(M-1) \times N}, \tag{2.2}$$

and

$$q = \Phi_h(x) : \mathfrak{R}^{M \times N} \rightarrow \mathfrak{R}^{M \times (N-1)}, \tag{2.3}$$

respectively, where the matrices defined by

$$\begin{aligned} p_{i,j} &= x_{i,j} - x_{i+1,j}, \quad i = 1, 2, \dots, M - 1, \quad j = 1, 2, \dots, N, \\ q_{i,j} &= x_{i,j} - x_{i,j+1}, \quad i = 1, 2, \dots, M, \quad j = 1, 2, \dots, N - 1. \end{aligned} \tag{2.4}$$

(ii) The operators Φ_v^* and Φ_h^* which are the adjoint of Φ_v and Φ_h are given by

$$z = \Phi_v^*(p) : \mathfrak{R}^{(M-1) \times N} \rightarrow \mathfrak{R}^{M \times N}, \tag{2.5}$$

and

$$z = \Phi_h^*(q) : \mathfrak{R}^{M \times (N-1)} \rightarrow \mathfrak{R}^{M \times N}, \tag{2.6}$$

respectively, where the matrices defined by

$$\begin{aligned} z_{i,j} &= p_{i-1,j} - p_{i,j}, \quad i = 1, 2, \dots, M, \quad j = 1, 2, \dots, N, \\ z_{i,j} &= q_{i,j-1} - q_{i,j}, \quad i = 1, 2, \dots, M, \quad j = 1, 2, \dots, N, \end{aligned} \quad (2.7)$$

where we assume that $p_{0,j} = p_{M,j} = q_{i,0} = q_{i,N}$ for any $i = 1, 2, \dots, M$, $j = 1, 2, \dots, N$.

(iii) P_C is the orthogonal projection operator on the set C . Thus, for instance, if $C = B_{l,u}$, then $P_{B_{l,u}}$ is explicitly given by

$$P_{B_{l,u}}(x)_{i,j} = \begin{cases} l & x_{ij} < l, \\ x_{ij} & l \leq x_{ij} \leq u, \\ u & x_{ij} > u. \end{cases} \quad (2.8)$$

Noting that if we arrange x, p, q column by column into vectors, then there are matrices Φ_v and Φ_h such that $p = \Phi_v x$ and $q = \Phi_h x$, respectively. And the adjoint operators Φ_v^* and Φ_h^* are associated with matrices Φ_v^T and Φ_h^T , respectively, where the superscript T denotes the transpose. Thus, we can rewrite the matrix norm of the anisotropic and anisotropic TVs as

$$TV_{iso} = \|\Phi_v(x) \quad \Phi_h(x)\|_{l_2}, \quad (2.9)$$

and

$$TV_{aniso} = \|\Phi_v(x)\|_{l_1} + \|\Phi_h(x)\|_{l_1}, \quad (2.10)$$

respectively, where l_2 -norm refer to the sqart of the sum of the square of each variable and the matrix l_1 -norm here refer to the sum of the absolute values.

3. The Modified Second APG Method for Solving Discrete TV-Image Deblurring Problems

We are interested in using a first-order method such as the accelerated proximal gradient (APG) algorithm and the steepest descent to solve the large-scale optimization problems. There are three APG methods introduced by Nesterov [12] for solving the sum of a smooth convex function and a nonsmooth convex function, here we focus on the second APG algorithm. For convenience, we first consider the following model which naturally extends the problem formulation (1.1):

$$\min_{x \in \mathfrak{R}^{M \times N}} F(x) := f(x) + g(x), \quad (3.1)$$

where $g : \mathfrak{R}^{M \times N} \rightarrow (-\infty, +\infty]$ is a proper closed convex function and $f : \mathfrak{R}^{M \times N} \rightarrow \mathfrak{R}$ is continuously differentiable function that has a Lipschitz continuous gradient $L > 0$:

$$\|\nabla f(x) - \nabla f(y)\|_F \leq L\|x - y\|_F, \quad \forall x, y \in \mathfrak{R}^{M \times N}. \quad (3.2)$$

We assume in addition that F is level-bounded, which implies that the set of the global minimizers of (3.1) is nonempty.

Many first-order methods such as the fast iterative shrinkage-thresholding algorithm (FISTA) [11] (which is based on the first APG method introduced by [13]) are used for solving the above problem (3.1) since the size of this kind of problem is typically large scale. The theory of the accelerated deformations of the accelerated APG method and their convergence rate are shown by many papers (see [4, 11] and the references therein). Here we will extend modified second APG method (APG_s) introduced in [18] such it is suitable for the general model (3.1), and we will establish the improved complexity results. Now we collate the following algorithm:

Algorithm 1: The modified second APG (APG_g) method:

Take $x^0, z^0 \in \mathfrak{R}^{M \times N}$, and $\{\theta_k\} \subseteq (0, 1]$,

for $k = 0, 1, 2, \dots$ **and do**

$$y^k = (1 - \theta_k)x^k + \theta_k z^k,$$

$$z^{k+1} = \arg \min_{z \in \mathfrak{R}^{M \times N}} \left\{ \frac{L}{2} \|z - [z^k - \frac{1}{\theta_k L} \nabla f(y^k)]\|_F^2 + \frac{1}{\theta_k} g(z) \right\}, \quad (3.3)$$

$$x^{k+1} = (1 - \theta_k)x^k + \theta_k z^{k+1},$$

end for

$$\theta_k = \frac{\sqrt{\theta_{k-1}^4 + 4\theta_{k-1}^2} - \theta_{k-1}^2}{2} \quad (3.4)$$

A possible drawback of the basic scheme of the above algorithm is that the Lipschitz constant L is generally larger. On the other hand, the Lipschitz constant in the regularization term depends on the maximum eigenvalue of $A^T A$, it not always easy compute especially for large-scale problems. Noting that when the function f satisfies the Lipschitz gradient condition (3.2), the following inequality holds:

$$f(x) \leq f(y) + \langle x - y, \nabla f(y) \rangle + \frac{L}{2} \|x - y\|_F^2, \quad \forall x, y. \quad (3.5)$$

In order to obtain smaller Lipschitz constant to accelerate convergence, Zuo and Lin [1] introduce a positive-definite matrix L_f to change the inequality (3.2), where inner product and the L_f -norm of the matrix L_f are $\langle x, y \rangle_{L_f} = x^T L_f y$ and $\|x\|_{L_f} = \sqrt{\langle x, x \rangle_{L_f}}$, respectively. Then (3.5) can be generalized as

$$f(x) \leq f(y) + \langle x - y, \nabla f(y) \rangle + \frac{1}{2} \|x - y\|_{L_f}^2, \quad \forall x, y. \quad (3.6)$$

The matrix L_f exists for a broad class of f . For instance, we can choose L_f as LI when f satisfies the Lipschitz gradient condition (3.2), where I is the identify matrix. Here, we extend the modified second APG to a more general and efficient class, so called the generalized modified second accelerated gradient ($GMAPG_s$) method. Hence, we rewrite the Algorithm 1 for accelerating convergence to solve (3.1):

Algorithm 2: The generalized modified second accelerated gradient:

Define L_f be a positive-definite matrix, let $x^0, z^0 \in \mathfrak{R}^{M \times N}$, and $\{\theta_k\} \subseteq (0, 1]$,

for $k = 0, 1, 2, \dots$ **and do**

$$y^k = (1 - \theta_k)x^k + \theta_k z^k,$$

$$z^{k+1} = \arg \min_{z \in \mathfrak{R}^{M \times N}} \left\{ \frac{1}{2} \|z - [z^k - \frac{1}{\theta_k} L_f^{-1} \nabla f(y^k)]\|_{L_f}^2 + \frac{1}{\theta_k} g(z) \right\}, \quad (3.7)$$

$$x^{k+1} = (1 - \theta_k)x^k + \theta_k z^{k+1},$$

update θ_k **by** (3.4)

end for

4. Algorithm for TV-Image Deblurring

Noting that the difficulty of solving the TV-image deblurring problems is due to the nonsmoothness of the TV functions. It is worth mentioning that Chambolle [6] advised to use the dual approach to overcome this difficulty, and proposed a gradient based algorithm for solving the resulting dual problems, which in the unconstrained case was shown to be a convex quadratic program. In view of the success of the method in effecting to solve these problems with TV regularized function. Here we transform the TV-image problems into a quadratic programming by using two auxiliary variables [10] and dual approach [6], such it suitable for Algorithm 2 in the following:

Replace $TV(x)$ in (1.1) by (1.4) and (1.5), respectively, we have

$$\min_{x \in B_{l,u}} \frac{1}{2} \|Ax - b\|_F^2 + \lambda TV_{iso}(x), \quad (4.1)$$

and

$$\min_{x \in B_{l,u}} \frac{1}{2} \|Ax - b\|_F^2 + \lambda TV_{aniso}(x), \quad (4.2)$$

where A is associated with the linear mapping A when x and b are treated as vectors. Here, we first using the matrix pairs to rewrite the problems (4.1) and (4.2) making them become two constrained problems, then we follow the approach proposed in [15] and construct a dual of these problems.

By introducing two auxiliary variables ϕ_v and ϕ_h [10] such that $\phi_v = \Phi_v x$ and $\phi_h = \Phi_h x$, and using (2.9) and (2.10), we structure the constraint problems of the above problems, then we further write their unconstrained formulation in the following, respectively:

$$\min_{x, \phi_v, \phi_h} \frac{\eta}{2} \|Ax - b\|_F^2 + \frac{1}{2} \|\phi_v - \Phi_v x\|_F^2 + \frac{1}{2} \|\phi_h - \Phi_h x\|_F^2 + \lambda \eta (\|\phi_v - \phi_h\|_{l_2} + \chi_{B_{l,u}}(x)), \quad (4.3)$$

and

$$\min_{x, \phi_v, \phi_h} \frac{\eta}{2} \|Ax - b\|_F^2 + \frac{1}{2} \|\phi_v - \Phi_v x\|_F^2 + \frac{1}{2} \|\phi_h - \Phi_h x\|_F^2 + \lambda \eta (\|\phi_v\|_1 + \|\phi_h\|_1 + \chi_{B_{l,u}}(x)), \tag{4.4}$$

where η is a relaxation parameter and $\chi_{B_{l,u}}$ is the characteristic function of $B_{l,u}$, i.e.,

$$\chi_{B_{l,u}}(x) = \begin{cases} 0, & x \in B_{l,u}, \\ +\infty, & x \notin B_{l,u}. \end{cases}$$

To utilize Algorithm 2 for the problems (4.3) and (4.4), we readily see that the decomposition in (3.1) should be chosen as

$$\begin{aligned} f(\hat{x}) &= \frac{\eta}{2} \|Ax - b\|_F^2 + \frac{1}{2} \|\phi_v - \Phi_v x\|_F^2 + \frac{1}{2} \|\phi_h - \Phi_h x\|_F^2, \\ g_{l_2}(\hat{x}) &= \lambda \eta (\|\phi_v\|_{l_2} + \|\phi_h\|_{l_2} + \chi_{B_{l,u}}(x)), \\ g_{l_1}(\hat{x}) &= \lambda \eta (\|\phi_v\|_{l_1} + \|\phi_h\|_{l_1} + \chi_{B_{l,u}}(x)), \end{aligned} \tag{4.5}$$

where $\hat{x} = (x^T, \phi_v^T, \phi_h^T)^T$.

Below, we consider the details of Algorithm 2 for solving the TV-image deblurring:

- (1) We choose the L_f similarly as [19]:

$$L_f = \text{diag}(\lambda_{\max} I, \sigma I, \sigma I), \tag{4.6}$$

where λ_{\max} is the large eigenvalue of $\eta A^T A + \sigma \Phi_v^T \Phi_v + \sigma \Phi_h^T \Phi_h$ and $\sigma \geq 2$. Moreover, because the size of the image x is large, then we use the more convenient way to estimate λ_{\max} proposed by Zuo and Lin [1] as in the following:

$$\lambda_{\max} \leq (\sqrt{\eta} \|A\|_2 + \sqrt{\sigma} \|\Phi_v\|_2 + \sqrt{\sigma} \|\Phi_h\|_2)^2, \tag{4.7}$$

where $\|A\|_2$ can be estimated as

$$\|A\|_2 \leq \min(\|A\|_F, \sqrt{\|A\|_1 \|A\|_\infty}), \quad (4.8)$$

due to the fact that $\|X\|_2 \leq \|X\|_F$ and the equality $\|X\|_2 \leq \sqrt{\|X\|_1 \|X\|_\infty}$ ([17]).

(2) Consider the corresponding subproblem of anisotropic:

$$\hat{z}^{k+1} = \arg \min_{\hat{z}} \left\{ \frac{1}{2} \|\hat{z} - [\hat{z}^k - \frac{1}{\theta_k} L_f^{-1} \nabla f(\hat{y}^k)]\|_{L_f}^2 + \frac{1}{\theta_k} g_{l_1}(z) \right\}, \quad (4.9)$$

where

$$\hat{z} = (z_x^T, z_{\phi_\nu}^T, z_{\phi_h}^T)^T, \hat{z}^k = ((z_x^k)^T, (z_{\phi_\nu}^k)^T, (z_{\phi_h}^k)^T)^T, \hat{y}^k = ((y_x^k)^T,$$

$(y_{\phi_\nu}^k)^T, (y_{\phi_h}^k)^T)^T$). This subproblem could be divided into the following

three independent smaller problems:

$$z_x^{k+1} = \arg \min_z \left\{ \frac{\lambda_{\max}}{2} \|z - \tilde{z}^k\|_F^2 + \frac{1}{\theta_k} \lambda \eta \chi_{B_{l,u}}(z) \right\}, \quad (4.10)$$

$$z_{\phi_\nu}^{k+1} = \arg \min_{\phi_\nu} \left\{ \lambda \eta \|\phi_\nu\|_{l_1} + \frac{\sigma}{2} \|\phi_\nu - \tilde{\phi}_\nu^k\|_F^2 \right\}, \quad (4.11)$$

$$z_{\phi_h}^{k+1} = \arg \min_{\phi_h} \left\{ \lambda \eta \|\phi_h\|_{l_1} + \frac{\sigma}{2} \|\phi_h - \tilde{\phi}_h^k\|_F^2 \right\}, \quad (4.12)$$

where

$$\tilde{z}^k = z_x^k - \frac{1}{\theta_k \lambda_{\max}} [\eta A^T (A y_x^k - b) + \Phi_\nu^T (\Phi_\nu y_x^k - y_{\phi_\nu}^k) + \Phi_h^T (\Phi_h y_x^k - y_{\phi_h}^k)], \quad (4.13)$$

$$\tilde{\phi}_\nu^k = z_{\phi_\nu}^k - \frac{1}{\theta_k \sigma} (y_{\phi_\nu}^k - \Phi_\nu y_x^k), \quad (4.14)$$

$$\tilde{\phi}_h^k = z_{\phi_h}^k - \frac{1}{\theta_k \sigma} (y_{\phi_h}^k - \Phi_h y_x^k). \quad (4.15)$$

It is easy to verify that

$$z_x^{k+1} = P_{B_{l,u}}(\tilde{z}^k). \quad (4.16)$$

Moreover, by ([15], Lemma 3.3), we have

$$(z_{\phi_v}^{k+1})_{i,j} = T_{\lambda\eta/\sigma}((\tilde{\phi}_v^k)_{i,j}), (z_{\phi_h}^{k+1})_{i,j} = T_{\lambda\eta/\sigma}((\tilde{\phi}_h^k)_{i,j}), \quad \forall i, j, \quad (4.17)$$

where the shrinkage $T_{\lambda\eta/\sigma}(\cdot)$ is defined as

$$T_{\lambda\eta/\sigma}(t) = \frac{t}{|t|} \max\{|t| - \lambda\eta/\sigma, 0\}. \quad (4.18)$$

(3) In view of the success of continuation technique in GAPG for updating the ‘‘Lipschitz constant’’ η .

Here, we incorporate continuation technique in [15] to accelerate Algorithm 2. We first assume the initial value $\eta_0 = \|b\|_F$, then let it decrease gradually until the target value $\bar{\eta} = \rho\eta_0$ is reached. We take $\rho = 10^{-3}$ and use the following relation to update η_k :

$$\eta_{k+1} = \max\left(\left(0.9 + 0.1\left(\frac{2k-2}{2k-1}\right)^{\frac{1}{8}}\right)\eta_k, \bar{\eta}\right). \quad (4.14)$$

Below, we summed up the algorithm for solving the TV_{aniso} image deblurring:

Algorithm 3:

Input: $b, \lambda, \eta_0, A, \Phi_v, \Phi_h$

$$x_0 = P_{B_{l,u}}(b), z_0 = P_{B_{l,u}}(b), \phi_v^0 = \Phi_v x_0, \phi_h^0 = \Phi_h x_0, \theta_0 = 1, \sigma = 1, \bar{\eta} = \rho \eta_0,$$

$$x_x^1 = x_0, x_{\phi_v}^1 = \phi_v^0, x_{\phi_h}^1 = \phi_h^0, z_x^1 = z_0, z_{\phi_v}^1 = \phi_v^0, z_{\phi_h}^1 = \phi_h^0.$$

for $k = 1, 2, \dots$ **and do**

Update $y_x^k, y_{\phi_v}^k, y_{\phi_h}^k$:

$$y_x^k = (1 - \theta_k)x_x^k + \theta_k z_x^k,$$

$$y_{\phi_v}^k = (1 - \theta_k)x_{\phi_v}^k + \theta_k z_{\phi_v}^k,$$

$$y_{\phi_h}^k = (1 - \theta_k)x_{\phi_h}^k + \theta_k z_{\phi_h}^k.$$

Update $z_x^{k+1}, z_{\phi_v}^{k+1}, z_{\phi_h}^{k+1}$:

$$z_x^{k+1} = P_{B_{l,u}} \left\{ z_x^k - \frac{1}{\theta_k \lambda_{\max}} [\eta A^T (A y_x^k - b) + \Phi_v^T (\Phi_v y_x^k - y_{\phi_v}^k) + \Phi_h^T (\Phi_h y_x^k - y_{\phi_h}^k)] \right\},$$

$$z_{\phi_v}^{k+1} = T_{\lambda \eta_k / \sigma} \left(z_{\phi_v}^k - \frac{1}{\theta_k \sigma} (y_{\phi_v}^k - \Phi_v y_x^k) \right),$$

$$z_{\phi_h}^{k+1} = T_{\lambda \eta_k / \sigma} \left(z_{\phi_h}^k - \frac{1}{\theta_k \sigma} (y_{\phi_h}^k - \Phi_h y_x^k) \right).$$

Update $x_x^{k+1}, x_{\phi_v}^{k+1}, x_{\phi_h}^{k+1}$:

$$x_x^{k+1} = (1 - \theta_k)x_x^k + \theta_k z_x^{k+1},$$

$$x_{\phi_v}^{k+1} = (1 - \theta_k)x_{\phi_v}^k + \theta_k z_{\phi_v}^{k+1},$$

$$x_{\phi_h}^{k+1} = (1 - \theta_k)x_{\phi_h}^k + \theta_k z_{\phi_h}^{k+1}.$$

end for

(4) In a similar way, the problem (4.3) can be divide into the problem (4.10) and the following subproblem:

$$(z_{\phi_\nu}^{k+1}, z_{\phi_h}^{k+1}) = \arg \min_{\phi_\nu, \phi_h} \{ \lambda \eta \|(\phi_\nu, \phi_h)\|_{l_2} + \frac{\sigma}{2} \|\phi_\nu - \tilde{\phi}_\nu^k\|_F^2 + \frac{\sigma}{2} \|\phi_h - \tilde{\phi}_h^k\|_F^2 \}, \quad (4.19)$$

where $\tilde{\phi}_\nu^k, \tilde{\phi}_h^k$ are defined in (4.14) and (4.15).

For the above subproblem, using ([15], Lemma 3.3) again, we have

$$(z_{\phi_\nu}^{k+1}, z_{\phi_h}^{k+1}) = S_{\lambda\eta/\sigma}(\tilde{\phi}_\nu^k, \tilde{\phi}_h^k), \quad (4.20)$$

where the shrinkage $S_{\lambda\eta/\sigma}(\cdot)$ is defined as

$$S_{\lambda\eta/\sigma}(\phi) = \frac{\phi}{\|\phi\|} \max\{\|\phi\| - \lambda\eta/\sigma, 0\}. \quad (4.21)$$

We can see that our APG_s -based algorithm for isotropic TV-based image deblurring is almost the same as Algorithm 3, so we omit it.

5. Numerical Result

In this section, we apply our algorithm and GAPG proposed by [1] for image deblurring from blurred and noisy images and compare their performances. Here, we only test the algorithms for solving isotropic TV-based image deblurring problem. We mainly compare the computational time and SNR (PSNR) of the two method and the original APG method.

In our numerical test, we focus on the following image deblurring problem:

$$b = x * \kappa + \omega, \quad (5.1)$$

where κ is a blurring kernel, ω is an additive normally distributed noise, and $*$ is the convolution operator.

All the numerical experiments are performed in Matlab 2016a on a 64-bit PC with an Intel (R) Core i5-7th Gen CPU (3.60GHz) and 128GB of RAM.

Here, we consider two 256×256 images, i.e., Lena and Cameraman that go through a 9×9 Gaussian blurring kernel mean zero with standard deviation 4, and then added with normally distributed noise with mean zero and standard deviation 10^{-3} . Besides, the regularization parameter was chosen to be $\lambda = 10^{-4}$ due to the chosen can give the best SNR. We use original APG, GAPG, and $GMAPG_s$ with 150 iterations on problem (5.1), and obtain the two reconstructions given in Figure 1 and Figure 2. The reconstruction of $GMAPG_s$ is slightly better than GAPG. This also reflected in the objective function values, the SNR values and the CPU times given in the following Table 5.1:

Table 5.1. Algorithms for Lena and Cameraman

| Algorithm | PSNR(dB) | | CPU time (s) | |
|--------------|----------|-----------|--------------|-----------|
| | Lena | Cameraman | Lena | Cameraman |
| Original APG | 23.93 | 21.48 | | |
| GAPG | 29.83 | 27.73 | 1.34 | 1.48 |
| $GMAPG_s$ | 29.96 | 27.87 | 1.77 | 1.86 |



Figure 1. Algorithm for Lena.

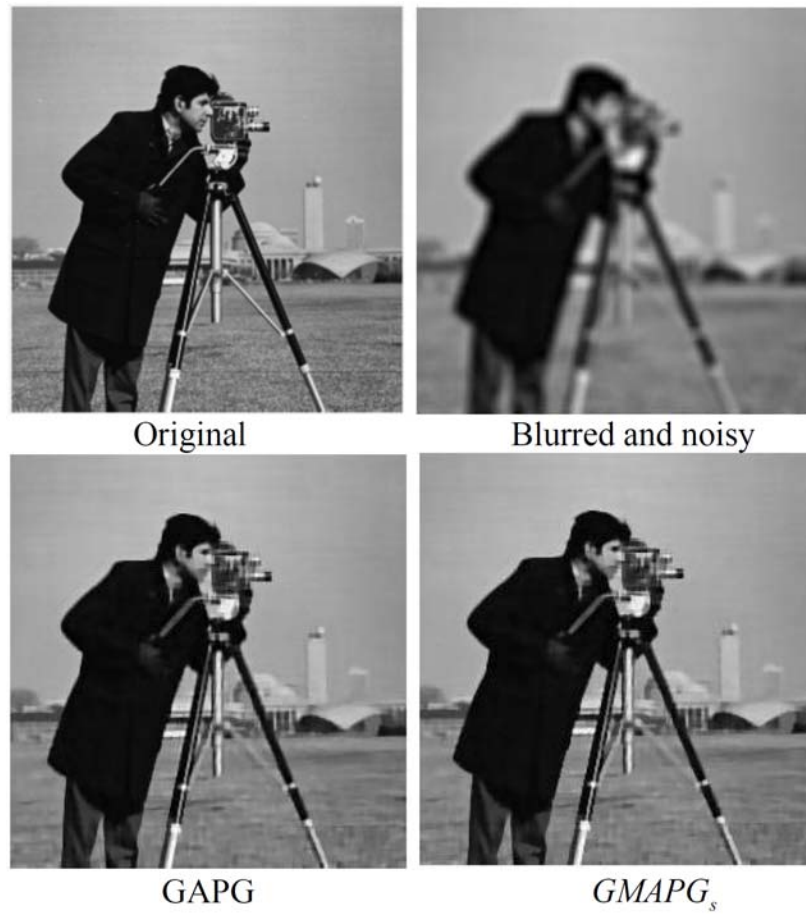


Figure 2. Algorithm for Cameraman.

From the above table, we see that the PSNR values of $GMAPG_s$ and GAPG are greater than the original APG, and the value of $GMAPG_s$ is slightly larger than GAPG.

6. Conclusion

In this paper, based on the second APG method proposed by Nesterov [12], we construct $GMAPG_s$ for TV-based image deblurring problem. Our algorithm framework allows a wide range of choice of the proximal parameters $\{\theta_k\}$, which including those used in FISTA with fixed restart [11]. First, we extend the second APG method with a suitable positive matrix instead of the original Lipschitz constant, so we can simplify the computation. Then, we use the auxiliary variables introduced in [10] to decouple the problems so that we need not to solve correlative subproblem in each iteration. Finally, in order to gradually reduce the relaxation parameter, we combine the continuation technique (introduced in [15]) in our algorithms, then we can obtain a good initial solution to make the rate of convergence faster. Numerical experiments show that our method performs well.

References

- [1] W. Zuo and Z. Lin, A generalized accelerated proximal gradient approach for total-variation-based image deblurring, *IEEE Transaction Image Processing* 20(10) (2011), 2748-2759.
DOI: <https://doi.org/10.1109/TIP.2011.2131665>
- [2] H. Takeda, S. Farsiu and P. Milanfar, Deblurring using regularized locally adaptive kernel regression, *IEEE Transactions on Image Processing* 17(4) (2008), 550-563.
DOI: <https://doi.org/10.1109/TIP.2007.918028>
- [3] A. Beck and M. Teboulle, Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems, *IEEE Transactions on Image Processing* 11(18) (2009), 2419-2434.
DOI: <https://doi.org/10.1109/TIP.2009.2028250>
- [4] J. Huang, T. Huang, X. Zhao et al., Two soft-thresholding based iterative algorithms for image deblurring, *Information Sciences* 271 (2014), 179-195.
DOI: <https://doi.org/10.1016/j.ins.2014.02.089>

- [5] Y. Wen, M. Ng and W. Ching, Iterative algorithm based on decoupling of deblurring and denoising for image restoration, *SIAM Journal on Scientific Computing* 30(5) (2008), 2655-2674.
DOI: <https://doi.org/10.1137/070683374>
- [6] A. Chambolle, An algorithm for total variation minimization and applications, *Journal of Mathematical Imaging and Vision* 20(1-2) (2004), 89-97.
DOI: <https://doi.org/10.1023/B:JMIV.0000011325.36760.1e>
- [7] A. Chambolle, Total variation minimization and a class of binary MRF models, In *Lecture Notes in Computer Sciences* 3757 (2005), 136-152.
DOI: https://doi.org/10.1007/11585978_10
- [8] I. Daubechies, M. Defrise and C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, *Communications on Pure and Applied Mathematics* 57(11) (2004), 1413-1457.
DOI: <https://doi.org/10.1002/cpa.20042>
- [9] K. Bredies and D. Lorenz, Linear convergence of iterative soft-thresholding, *J. Fourier Anal. Appl.* 14(5-6) (2008), 813-837.
DOI: <https://doi.org/10.1007/s00041-008-9041-1>
- [10] J. Bioucas-Dias and M. Figueiredo, A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration, *IEEE Transactions on Image Processing* 16(12) (2007), 2992-3004.
DOI: <https://doi.org/10.1109/TIP.2007.909319>
- [11] A. Beck and M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM Journal on Imaging Sciences* 2(1) (2009), 183-202.
DOI: <https://doi.org/10.1137/080716542>
- [12] Y. Nesterov, *Introductory Lectures on Convex Optimization*, Kluwer Academic Publisher, Dordrecht, the Netherlands, 2004.
- [13] P. Tseng, Approximation accuracy, gradient methods, and error bound for structured convex optimization, *Mathematical Programming* 125(2) (2010), 263-295.
DOI: <https://doi.org/10.1007/s10107-010-0394-2>
- [14] N. Parikh and S. Boyd, Proximal algorithms, *Foundations and Trends in Optimization* 1(3) (2013), 123-231.
DOI: <http://dx.doi.org/10.1561/2400000003>
- [15] J. Yang, W. Yin, Y. Zhang and Y. Wang, A fast algorithm for edge-preserving variational multichannel image restoration, *SIAM Journal on Imaging Sciences* 2(2) (2009), 569-592.

- [16] M. Afonso, J. Bioucas-Dias and M. Figueiredo, Fast image recovery using variable splitting and constrained optimization, *IEEE Transactions on Image Processing* 19(9) (2010), 2345-2356.

DOI: <https://doi.org/10.1109/TIP.2010.2047910>

- [17] T. Goldstein and S. Osher, The split Bregman method for l_1 -regularized problems, *SIAM Journal on Imaging Sciences* 2(2) (2009), 323-343.

DOI: <https://doi.org/10.1137/080725891>

- [18] D. Lin and C. G. Liu, The modified second APG method for DC optimization problems, preprint.

- [19] Z. J. Jiang and C. J. Duanmu, Using cartoon texture image decomposition and GAPG for total-variation-based image restoration (by Chinese), *Computer Engineering and Applications* 50(2) (2014), 162-169.

DOI: <https://doi.org/10.3778/j.issn.1002-8331.1203-0312>

- [20] D. Donoho and I. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, *J. Amer. Statis. Assoc.* 90(432) (1995), 1200-1224.

