Research Article

A Nonmonotone Gradient Algorithm for Total Variation Image Denoising Problems

Peng Wang,[1,2] Shifang Yuan,[1] Xiangyun Xie,[1] and Shengwu Xiong[2]

[1]*School of Mathematics and Computational Science, Wuyi University, Jiangmen 529020, China*
[2]*School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, China*

Correspondence should be addressed to Peng Wang; p_wong@126.com

The total variation (TV) model has been studied extensively because it is able to preserve sharp attributes and capture some sparsely critical information in images. However, TV denoising problem is usually ill-conditioned that the classical monotone projected gradient method cannot solve the problem efficiently. Therefore, a new strategy based on nonmonotone approach is digged out as accelerated spectral project gradient (ASPG) for solving TV. Furthermore, traditional TV is handled by vectorizing, which makes the scheme far more complicated for designing algorithms. In order to simplify the computing process, a new technique is developed in view of matrix rather than traditional vector. Numerical results proved that our ASPG algorithm is better than some state-of-the-art algorithms in both accuracy and convergence speed.

1. Introduction

Blur and noise are unavoidable during the acquisition and transmission of images, and this may be introduced by the poor lighting conditions, faulty camera or unperfect transmission channels, and so on. It is interesting and challenging to restore the ideal image from its degraded version. Several effective mathematical models and algorithms have emerged such as partial differential equations [1], sparse-land models [2, 3], and regularization [2].

This paper mainly elaborated on one of the most popular and effective tools for image denoising problem named TV [3–5], as it has been shown to preserve sharp edges both experimentally and theoretically. In view of the good performance in image processing, TV has been extensively studied in many image processing fields such as denoising [3, 4], deblurring [6], and inpainting [5, 7]. Meanwhile, various algorithms [3, 5–8] for solving TV are also proposed. Even though there are so many algorithms, most of them are merely related to vectorization [3, 6]. As images are two-dimensional signals, it is more intuitive and suitable to process by matrix than by vector. Furthermore, it is more efficient and simpler for image processing in a matrix view. In this

work, some universal and useful properties of TV are given in matrix version, while it also applies to other algorithms even without any modification.

Over the last few decades, there have been many different variations of the projected gradient method (PG) to solve TV denoising problems [9–11]. Generally, the classical PG method has complexity result of $O(1/k)$ as one of its drawbacks [12], where k is the iteration counter. Many authors devoted themselves into enhancing PG. Birgin et al. [13, 14] imposed a nonmonotone technique to classical PG and developed the spectral projected gradient (SPG) method. Nesterov [15] gave a concise and efficient strategy which improved the convergence speed of PG to $O(1/k^2)$. Recently, similar techniques were tactfully managed to different traditional algorithms [8, 10, 11, 16, 17] and corresponding optimal complexity results were obtained: Beck and Teboulle [10] proposed a fast duality-based gradient projection methods to solve TV; Chambolle and Pock [11] and Goldfarb et al. [16] accelerated the classical primal-dual algorithm and alternating direction augmented Lagrangian methods, respectively; more recently, Ouyang et al. [17] adopted the same technique to enhance the convergence of alternating direction method of multipliers (ADMM). Though so many studies have been done based

on Nesterov's theory, to the best of our knowledge, no work explores the possibility between Nesterov's scheme [15] and the spectral projected gradient algorithm. Although the SPG approach is interesting, the strong disturbance always occurs before its convergence and thus casts a shadow over the approach; meanwhile, the key thought in Nesterov's work [15] is that a combination of past iterates is used to compute the next iterate. This work is purposed to reduce the SPG's disturbance by Nesterov's method and propose an accelerated spectral projected gradient (ASPG) for TV denoising problem.

2. Primal-Dual Framework of Total Variation Denoising

In this section, we consider the TV denoising problem which is formulated as

$$\min_{u \in \Omega} \quad \|u - b\|^2 + 2\lambda \|u\|_{\mathrm{TV}}, \tag{1}$$

where u and b denote an ideal and the noised image, respectively, the box region Ω is practically selected as $[0, 1]$ or $[0, 255]$ to constrain the value of u, and $\|u\|_{\mathrm{TV}}$ is called TV seminorm usually divided into isotropic and anisotropic types which are, respectively, defined as the paper of Chambolle [9]. It is worth mentioning that our approach can be applied to TV_I and TV_A. However, for the sake of compactness, we will put emphasis on TV_I.

2.1. Notations and Outline. To smooth the process, some main notations are introduced here. The standard inner product in Euclidean space $R^{m \times n}$ is defined as $\langle A, B \rangle = \mathrm{tr}(A^T B)$. Denote the right linear space in Euclidean space $R^{(m-1) \times n} \times R^{m \times (n-1)} = \{[A, B] \mid A \in R^{(m-1) \times n}, B \in R^{m \times (n-1)}\}$. For convenience, we write $\mathscr{C} = R^{(m-1) \times n} \times R^{m \times (n-1)}$. Thus, we can define the inner product as

$$\langle [A_1, B_1], [A_2, B_1] \rangle = \mathrm{tr}\left(A_2^T A_1\right) + \mathrm{tr}\left(B_2^T B_1\right),$$
$$\forall [A_i, B_i] \in R^{(m-1) \times n} \times R^{m \times (n-1)}, \ (i = 1, 2). \tag{2}$$

The rest of this paper is organized as follows: In the following subsections, some important operators in matrix version for solving TV, followed by the algorithm for TV-based image denoising, are explained. In Section 3, the spectral projected gradient algorithm is first recapped and then the accelerated version ASPG is derived by it. To show the effectiveness of the proposed approach, numerical comparisons with existing state-of-the-art methods are carried out in Section 4. Finally, Section 5 contains the conclusion.

2.2. Some Preliminaries for TV Denoising. In this section, we will define some useful operators which are necessary for our algorithm. The operators are defined in matrix version instead of the traditional vector version [3, 6], which makes TV more understandable and concise.

(a) Subset. S is a subset in the inner product space $\mathscr{C}$ as follows:

$$S = \Big\{(P_1, P_2) \mid (P_1, P_2) \in \mathscr{C}, p_{1(i,j)}^2 + p_{2(i,j)}^2$$
$$\leq 1 \ \left(\text{for } i = 1 : (m-1), j = 1 : (n-1)\right), \ p_{1(i,n)}$$
$$\leq 1, \ \left(\text{for } i = 1 : (m-1)\right), \ p_{2(m,j)}$$
$$\leq 1, \ \left(\text{for } j = 1 : (n-1)\right)\Big\}, \tag{3}$$

where $p_{k,(i,j)}$ is the entry of P_k (for $k = 1, 2$).

(b) Projected Operator. In this paper, two kinds of projection operators $P_\Omega(u)$ and $P_S(\mathbf{P})$ are necessary: An image u projected onto a box set $\Omega \equiv [b_1, b_2]$ is easily computed:

$$(P_\Omega u)_{i,j} = \begin{cases} b_1 & \text{if } u_{i,j} < b_1 \\ u_{i,j} & \text{if } b_1 \leq u_{i,j} \leq b_2 \\ b_2 & \text{if } u_{i,j} > b_2. \end{cases} \tag{4}$$

Taking $\mathbf{P} = (P_1, P_2) \in \mathscr{C}$, for TV_I problem, $(P_S \mathbf{P})_{i,j} \in S$ is given by

$$(P_S P_1)_{i,j}$$
$$= \begin{cases} \dfrac{(P_1)_{i,j}}{\max\left\{1, \left|(P_1)_{i,n}\right|\right\}} & \text{if } i = 1 : (m-1) \\ \dfrac{(P_1)_{i,j}}{\max\left\{1, \sqrt{(P_1)_{i,j}^2 + (P_2)_{i,j}^2}\right\}} & \text{otherwise,} \end{cases}$$

$$(P_S P_2)_{i,j}$$
$$= \begin{cases} \dfrac{(P_2)_{i,j}}{\max\left\{1, \left|(P_2)_{i,n}\right|\right\}} & \text{if } i = 1 : (m-1) \\ \dfrac{(P_2)_{i,j}}{\max\left\{1, \sqrt{(P_1)_{i,j}^2 + (P_2)_{i,j}^2}\right\}} & \text{otherwise,} \end{cases} \tag{5}$$

while for TV_A problem the projected operator is defined as

$$(P_S P_1)_{i,j} = \frac{(P_1)_{i,j}}{\max\left\{1, \left|(P_1)_{i,j}\right|\right\}};$$
$$(P_S P_2)_{i,j} = \frac{(P_2)_{i,j}}{\max\left\{1, \left|(P_2)_{i,j}\right|\right\}}. \tag{6}$$

(c) Gradient Operator. If an image $u \in R^{m \times n}$, the gradient is an element in $\mathbf{D}u \in \mathscr{C}$ defined as $\mathbf{D}u = (D_x u, D_y u)$ with

$$\begin{aligned}
(D_x u)_{i,j} &= u_{i,j} - u_{i+1,j}, \quad i = 1 : (m-1), \; j = 1 : n; \\
(D_y u)_{i,j} &= u_{i,j} - u_{i,j+1}, \quad i = 1 : m, \; j = 1 : (n-1).
\end{aligned} \tag{7}$$

(d) Divergence Operator. The divergence of a vector $\mathbf{P} = (P_1, P_2) \in S$ is an array in $R^{m \times n}$ computed by

$$(\operatorname{div} \mathbf{P})_{i,j} = p_{1,(i,j)} - p_{1,(i-1,j)} + p_{2,(i,j)} - p_{2,(i,j-1)}, \tag{8}$$

which implies that $p_{1,(0,j)} = p_{1,(m,j)} = p_{2,(i,0)} = p_{2,(i,n)} \equiv 0$ for $i = 1 : m, \; j = 1 : n$.

It should be mentioned that the operators (b), (c), and (d) are defined with the assumption of periodic boundary which is adopted in the rest of this paper.

2.3. The Primal Dual Solution of TV.

As the denoising problem (1) is not differentiable, the traditional optimal methods are not applicable. Though there are so many techniques to overcome the difficulty, we employ the scheme of Chambolle [9, 18] for its effectiveness and simpleness.

Now, we give the TV-based image denoising algorithm in primal-dual view which mainly contains two steps.

Algorithm 1 (TV-based image denoising (TVID)).

    Input:

    b: observed image

    λ: regularization parameter

    Step 0: Take $\mathbf{P}_0 = (\mathbf{0}_{(m-1) \times n}, \mathbf{0}_{m \times (n-1)})$.

    Step 1: $u = P_\Omega(b - \lambda \operatorname{div} \mathbf{P})$ Compute

    Step 2: $\mathbf{P}^* = \operatorname{argmin}_{\mathbf{P} \in S} \|b - \lambda \operatorname{div} \mathbf{P}\|_F^2 - \|u^* - b + \lambda \operatorname{div} \mathbf{P}\|_F^2$

    Output: $u^* = P_\Omega(b - \lambda \operatorname{div} \mathbf{P}^*)$

Theorem 2. *The TVID algorithm can obtain an analytic solution of the TV image denoising problem (1).*

Proof. According to the primal-dual approach [9, 18], the TV seminorm can be written as follows:

$$\|u\|_{\mathrm{TV}} = \sup_{\mathbf{P} \in S} \langle u, \operatorname{div} \mathbf{P} \rangle. \tag{9}$$

Substituting (9) into (1), then we get

$$\min_{u \in \Omega} \max_{\mathbf{P} \in S} \quad \|u - b\|^2 + 2\lambda \langle u, \operatorname{div} \mathbf{P} \rangle. \tag{10}$$

By the min-max theorem, the above equation is equivalent to

$$\max_{\mathbf{P} \in S} \min_{u \in \Omega} \quad \|u - b\|^2 + 2\lambda \langle u, \operatorname{div} \mathbf{P} \rangle. \tag{11}$$

Obviously, the inner part of the above equation,

$$\min_{u \in \Omega} \quad \|u - b\|^2 + 2\lambda \langle u, \operatorname{div} \mathbf{P} \rangle, \tag{12}$$

has an analytic solution:

$$u^* = P_\Omega(b - \lambda \operatorname{div} \mathbf{P}), \tag{13}$$

where $P_\Omega(\zeta)$ denotes the projection of ζ onto the convex set Ω. $\square$

3. Accelerated Spectral Projected Gradient Based TV Denoising

In this section, first there will be a brief on spectral projected gradient method [14]. Next an accelerated strategy will be used to enhance SPG.

3.1. Introduction to Spectral Projected Gradient Method.

By combining Barzilai-Borwein nonmonotone [19] ideas with traditional projected gradient algorithm, Birgin et al. [13, 14] developed a new gradient algorithm that is SPG, which is suitable for convex optimization problems with projection easy to be computed on feasible set. By incorporating a spectral step length and a nonmonotone globalization strategy, SPG speeds up the traditional PG and maintains it to be simple and easy to code.

The SPG method aims to minimize f on a closed and convex set Ω; that is,

$$\min_{x \in \Omega} \quad f(x). \tag{14}$$

As the traditional projection method, the SPG method has the form

$$x^{k+1} = x^k + \alpha_k d_k, \tag{15}$$

where α_k is the step length, and the search direction d_k has been defined as [14]

$$d_k = P_\Omega \left[x^k - \lambda_k \nabla f(x^k) \right] - x^k. \tag{16}$$

In the classical SPG method, the nonmonotone decrease criterion is realized by an integer parameter $M \geq 1$ which guarantees the object function decrease every M iterations. In the line search stage, a sufficient decrease parameter $\gamma \in (0, 1)$ is used to satisfy Armijo criterion. Besides, some other parameters are also required: safeguarding parameters $0 < \alpha_{\min} < \alpha_{\max} < \infty$ for the spectral step length and safeguarding parameters $0 < \sigma_1 < \sigma_2 < 1$ for the quadratic interpolation. For later reference, we now recall the theorem by Birgin et al. [14] in which convergence of classical SPG is shown.

Theorem 3 (see [14]). *The spectral projected algorithm is well defined, and any accumulation point of the sequence $\{x_k\}$ that it generates is a constrained stationary point of (14).*

3.2. Accelerated SPG for TVID.

In this section, we will elaborate our algorithm on TV denoising problem (1). From algorithm TVID, problem (1) can be solved if and only if Step 2 is solved. Thus, we proposed an accelerated spectral projected gradient (ASPG) for solving Step 2. The proposed ASPG mainly consists of two parts: a nonmonotone projected

gradient for finding the feasible solutions and a combination of the former two iterations for reducing the turbulence of the algorithm and accelerating the convergence speed. In order to solve Step 2 in algorithm TVID, substituting (13) into (11) and simplifying yield

$$\min_{\mathbf{P} \in S} \quad G(\mathbf{P}) = \|b - \lambda \operatorname{div} \mathbf{P}\|_F^2 - \|u^* - b + \lambda \operatorname{div} \mathbf{P}\|_F^2. \quad (17)$$

And the gradient of (17) is

$$\nabla G(\mathbf{P}) = -2\lambda D(P_\Omega(b - \lambda \operatorname{div} \mathbf{P})), \quad (18)$$

where $D(\cdot)$ can be computed by the gradient operator in Section 2.2. Substituting (18) into (16) and simplifying them together with (13), we get the search direction of ASPG:

$$d_k = P_S(\mathbf{P}_k + 2\lambda \alpha_k D(u_k)) - \mathbf{P}_k, \quad (19)$$

where $P_S(\cdot)$ is defined as the projected operator in Section 2.2.

The nonmonotone linear search and the spectral step length are updated just as the classical SPG [14]. Following nonmonotone line search, ASPG adopts Nesterov's [15] technique to improve the performance of classical SPG. Finally, it should be mentioned that the ASPG algorithm is used for the dual step in TVID algorithm; that is, the variation $\mathbf{P}$ in ASPG is an element of the inner product space $\mathscr{C}$. Consequently, the inner product involving ASPG should be computed as (2).

Algorithm 4 (accelerated spectral projected gradient (ASPG)).

Input: $\mathbf{Q}_0 \in \mathscr{C}$, integer $M \geq 1$, $\alpha_0 \in [\alpha_{\min}, \alpha_{\max}]$

   the regularization parameter λ

   an integer Iter and the observed image b

Step 0: Take $\mathbf{P}_1 = \mathbf{Q}_0 = (0_{(m-1) \times n}, 0_{m \times (n-1)})$, $t_1 = 1$

Step 1: $(k = 1, \ldots, \text{Iter})$ Compute

Step 2: Iterate

   $u_k = P_\Omega(b - \lambda \operatorname{div} \mathbf{P}_k)$
   $d_k = P_S(\mathbf{P}_k + 2\lambda \alpha_k D(u_k)) - \mathbf{P}_k$;
   Step 2.1: Nonmonotone Line Search
   Compute
   $G_{\max} = \max\{G(\mathbf{P}_{k-j}) \mid 0 \leq j \leq \min(k, M-1)\}$
   and set $\tau \leftarrow 1$

      Step 2.1.1:

   If $G(\mathbf{P}_k + \tau d_k) \leq G_{\max} - 2\gamma\tau\lambda\langle d_k, D(u_k)\rangle$
   then set $\tau_k \leftarrow \tau$ and stop

      Step 2.1.2: Compute

   $\tau_t = \lambda\tau^2\langle d_k, D(u_k)\rangle / 2[G(\mathbf{P}_k + \tau d_k) - G(\mathbf{P}_k) + 2\lambda\tau\langle d_k, D(u_k)\rangle]$
   If $\tau_t \in [\sigma_1, \sigma_2\tau]$ then set $\tau \leftarrow \tau_t$

   Otherwise, set $\tau \leftarrow \tau/2$.
   Go to Step 2.1.1
   Step 2.2: $\mathbf{Q}_k = P_S(\mathbf{P}_k + \alpha d_k)$

Step 3:

   $t_{k+1} = (1 + \sqrt{1 + 4t_k^2})/2$;
   $\mathbf{P}_{k+1} = \mathbf{Q}_k + ((t_k - 1)/t_{k+1})(\mathbf{Q}_k - \mathbf{Q}_{k-1})$;

Step 4: Compute the spectral step length

   $S_k = \mathbf{P}_{k+1} - \mathbf{P}_k$;
   $Y_k = -2\lambda(D(u_{k+1}) - D(u_k))$
   If $\langle Y_k, S_k \rangle \leq 0$ then set $\alpha_{k+1} \leftarrow \alpha_{\max}$
   Otherwise
   $\alpha_{k+1} = \max\{\alpha_{\min}, \min\{\langle S_k, S_k\rangle / \langle Y_k, S_k\rangle, \alpha_{\max}\}\}$

Output: $u^* = P_\Omega(b - \lambda \operatorname{div} \mathbf{P}_{\text{Iter}})$

3.3. Convergence Analysis of ASPG. Since our method can be seen as an accelerated version of classical SPG in the dual space, its convergence can be deduced from Theorem 3. Now, we summarize the convergence of ASPG in the following theorem.

Theorem 5. *The ASPG for TVID is well defined, and any accumulation point of the sequence $\{\mathbf{P}_k\}$ is a constrained stationary point of (17); furthermore, the output u^* of ASPG is a stationary point of TVID problem (1).*

Proof. By (13), u^* generated by ASPG is a stationary point of (1) if and only if $\mathbf{P}$ is a stationary of (17).

Equation (18) shows that the kth gradient of (17) can be gotten by u_k; that is, $\nabla G(\mathbf{P}_k) = -2\lambda D(P_\Omega(u_k))$. Substituting it into (16), we get the search direction of ASPG that is (19). This means that the search direction of ASPG is the same as that of the SPG direction except that (19) is in the dual space $\mathscr{C}$. Furthermore, the definition of gradient operator $D(\cdot)$ shows that $D(u_k)$ is an element of $\mathscr{C}$. Therefore, the inner product $\langle d_k, D(u_k)\rangle$ can be computed by (2). Moreover, both the nonmonotone line search and the spectral step length mainly involve the inner products $\langle d_k, D(u_k)\rangle$, $\langle S_k, S_k\rangle$, and $\langle S_k, Y_k\rangle$. Consequently, the nonmonotone line search and the spectral step length of ASPG are just different from those of the classical SPG in the inner product.

Step 3 in ASPG is an additional step besides the classical SPG. It can be known that the step is a linear combination of the previous two results, which will not change the convergence of the whole algorithm. Therefore, from Theorem 3, it follows that $\mathbf{P}$ is a stationary of (17). This implies the conclusion. □
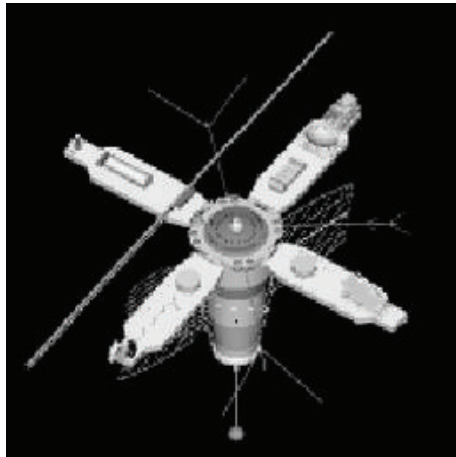
4. Numerical Experiments

In this section, the proposed ASPG algorithm (ASPG_Proposed) is employed to solve the TVID problem, followed by the comparison with some state-of-the-art algorithms: the classical projected gradient given by Chambolle

(a) Barbara

(b) Boats

(c) Satellite

(d) Pepper

Figure 1: Original image.

(PG_Chambolle) [9], Birgin's spectral projected gradient [13, 14] (SPG_Birgin), and the fast gradient projection by Beck and Teboulle [10] (FPG_Beck). The effect of these experiments is measured by relative error, peak signal-to-noise ratio (PSNR), and structural similarity index measurement (SSIM). The test images are 256-by-256 images as shown in Figure 1: (a) Barbara, (b) Boats, (c) Satellite, and (d) Peppers. All the image pixel values are scaled to the interval [0, 1].

In the following, the applicability of the proposed algorithm on image denoising problems is illustrated. The degraded images are generated by adding Gaussian noise with zero mean and standard deviation of size 0.05 to the original images (Figure 1). Throughout the experiments, the parameters involved in SPG and ASPG are set as follows: $\alpha_{\min} = 10^{-30}$, $\alpha_{\max} = 10^{30}$, $\sigma_1 = 0.1$, $\sigma_2 = 0.9$, $\gamma = 10^{-4}$, and $M = 3$. For the value of the regularization parameter λ for all schemes, tune it manually and choose the one that gives the most satisfactory results for all four images.

The PSNR and the relative error of the four images with different algorithms are shown in Table 1. To exclude the random fact to the results, the last column of Table 1 shows the average metrics of the four tested images, and the metrics of the best method for each test case are highlighted in boldface. Similar to FPG_Beck improving PG_Chambolle by Nesterov's strategy, the proposed ASPG improves the classical SPG_Birgin. Table 1 illustrates that ASPG and FPG_Beck are better than SPG_Birgin and PG_Chambolle, respectively. This fact shows that the accelerated scheme is valid for the ASPG as well as the FPG_Beck. Furthermore, we can find that ASPG is superior to FPG_Beck in that the former is based on the nonmonotonic line search. Table 1 shows that, among all the algorithms, the proposed ASPG algorithm performs the best in both PSNR and relative error.

The following goes with a deep comparison of the mentioned algorithms. To illustrate this more clearly, in Figure 2, relative error, PSNR, and SSIM value are introduced
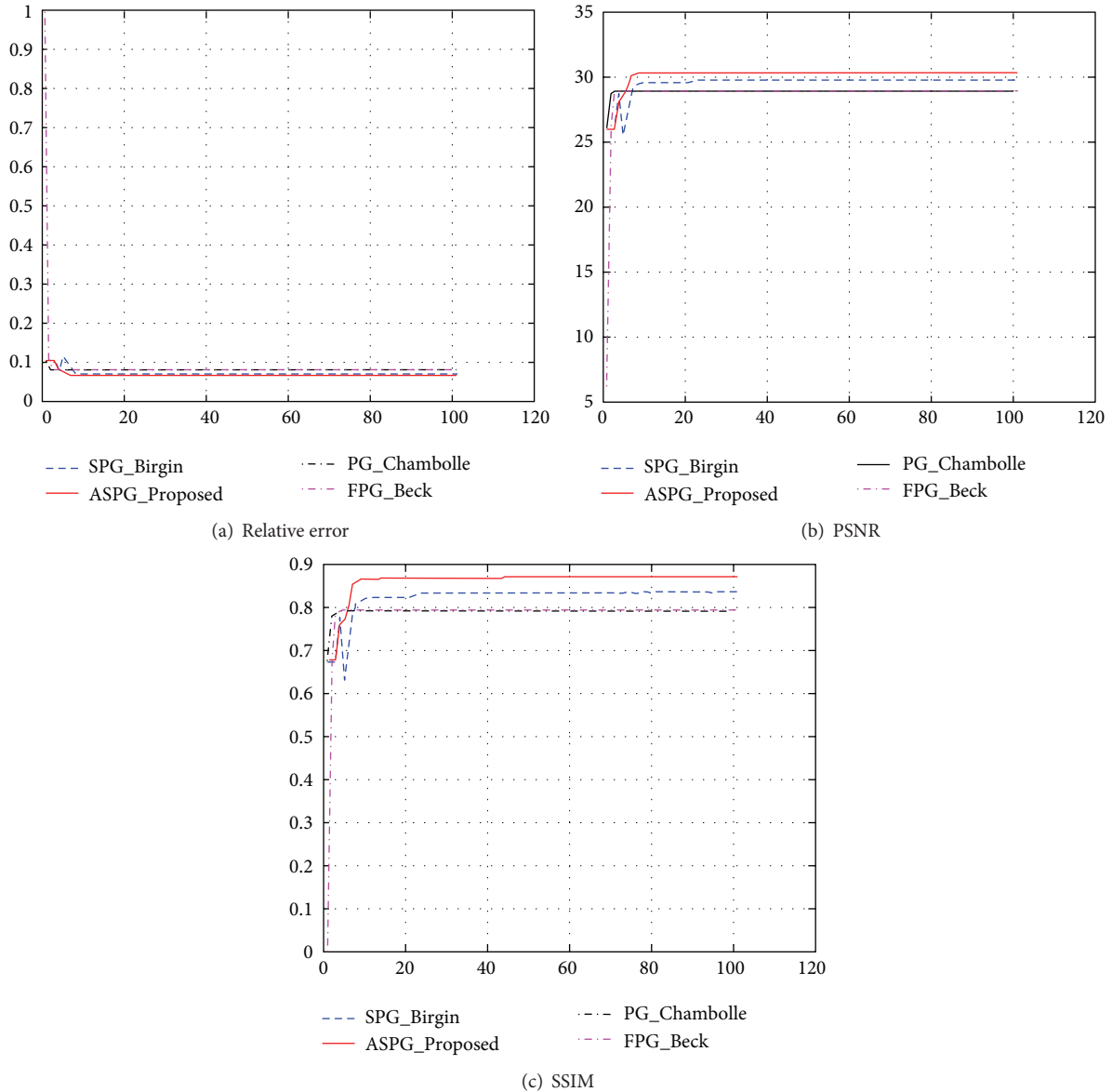
(a) Relative error

(b) PSNR

(c) SSIM

FIGURE 2: The evaluation index.

TABLE 1: Numerical comparison for images (a)–(d) in Figure 1.

	Barbara		Boats		Satellite		Pepper		Average	
	PSNR	Error	PSNR	Error	PSNR	Error	PSNR	Error	PSNR	Error
PG_Chambolle	28.8919	0.0735	29.1882	0.0654	29.5731	0.1077	29.4148	0.0662	29.2670	0.0782
SPG_Birgin	29.7996	0.0662	30.4961	0.0562	32.4344	0.0775	30.8096	0.0564	30.8849	0.0641
FPG_Beck	28.9254	0.0732	29.2779	0.0647	31.2876	0.0884	29.4227	0.0661	29.7284	0.0731
ASPG	**30.2991**	**0.0625**	**31.3927**	**0.0507**	**32.6485**	**0.0756**	**32.0317**	**0.0490**	**31.5930**	**0.0594**

here, which are related to the iterative times for noised Barbara (generated from Figure 1(a)). From Figure 2, we can see that the classical SPG has strong disturbance before its convergence while the ASPG reduces the disturbance and accelerates the convergence speed. Meanwhile, Figure 2 also illustrates that the nonmonotone line search improves all the metrics PSNR, relative error, and SSIM. One more point to mention is that here we only show the relative error, PSNR, and SSIM value which are generated from the degraded version of Figure 1(a), while, for the other figures

(a) Barbara, noised

(b) Denoised, SPG

(c) Denoised, FPG

(d) Denoised, ASPG

FIGURE 3: Denoised by different algorithms.

(i.e., Figures 1(b), 1(c), and 1(d)), they are just different in value compared with that of Figure 1(a), but the performance priority of PG_Chambolle, SPG_Birgin, FPG_Beck, and ASPG is the same, which all verify that ASPG is the best. Finally, Figure 3(a) is corrupted from Figure 1(a). Figures 3(b)–3(d) illustrate the denoised results which processed by SPG_Birgin, FPG_Beck, and ASPG.

Remote sensing images are widely used in so many fields [20–23]. However, noise is unavoided during the acquisition and transmission of images. To restore the the ideal image from its degraded version is an interesting field of remote sensing image processing. Here, we employ the proposed ASPG algorithm to denoise a remote sensing image (Figure 4). The results are shown as that from Figures 4(a)–4(f), and a selected area is magnified to better prove them in more detail. Both the original and magnified areas are highlighted in green.

5. Conclusions

In this paper, an accelerated scheme for spectral project gradient is proposed for TV denoising problems. To handle TV in matrix space, a concise version for total variation model is introduced. Numerical examples have been done to verify the performance of algorithm, which show that it is much better than that of some existing state-of-the-art methods.

Competing Interests

There are no competing interests related to this paper.

Acknowledgments

(a) Ground truth

(b) Noised

(c) PG (PSNR = 28.3365)

(d) SPG (PSNR = 28.8468)

(e) FPG (PSNR = 28.2593)

(f) ASPG (PSNR = 28.8547)
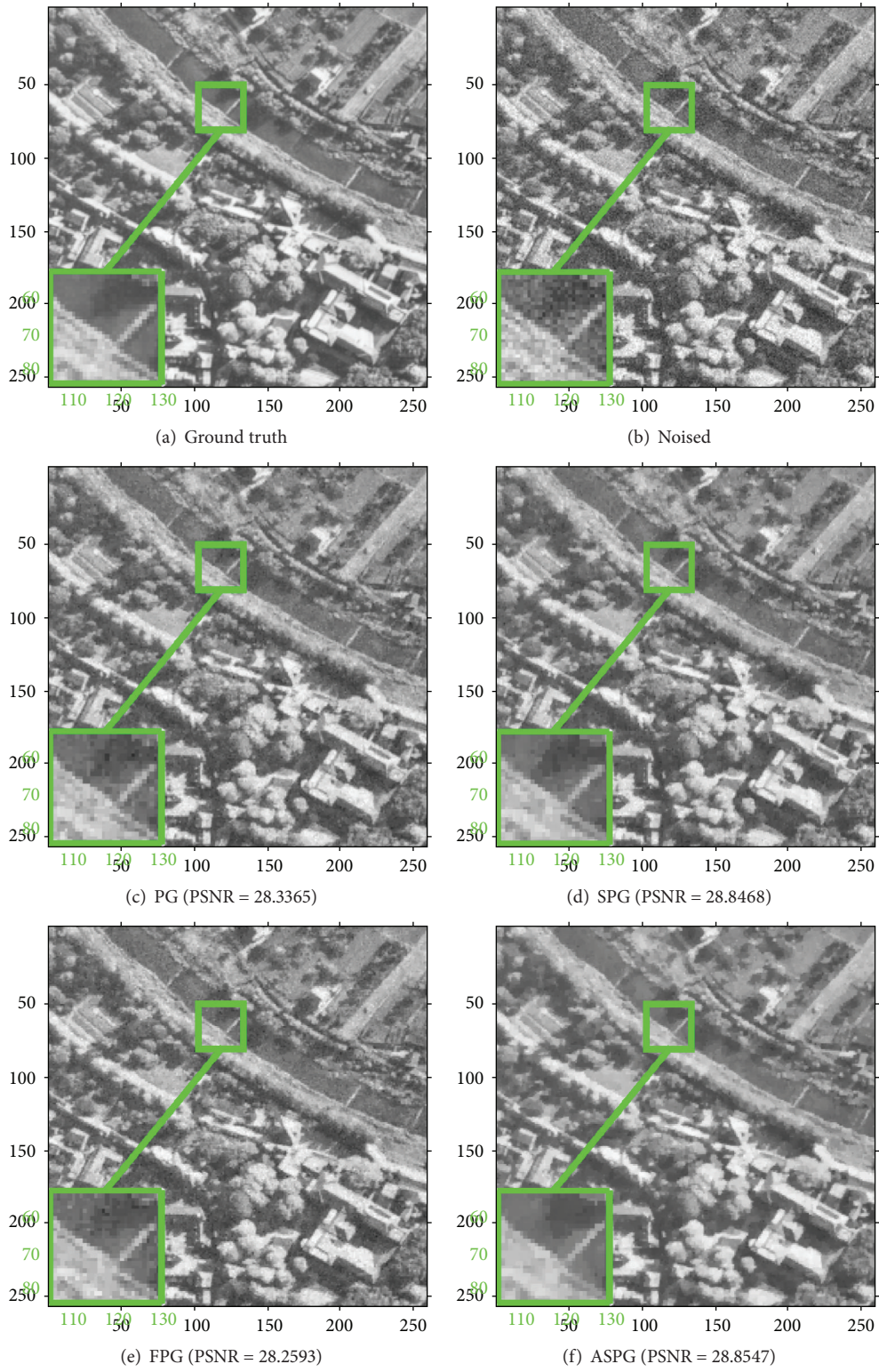
FIGURE 4: Remote sensing image denoising.

References

[1] W. Tian, T. Ma, Y. Zheng et al., "Weighted curvature-preserving PDE image filtering method," *Computers & Mathematics with Applications*, vol. 70, no. 6, pp. 1336–1344, 2015.

[2] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1838–1857, 2011.

[3] C. A. Micchelli, L. Shen, Y. Xu, and X. Zeng, "Proximity algorithms for the L1/TV image denoising model," *Advances in Computational Mathematics*, vol. 38, no. 2, pp. 401–426, 2013.

[4] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1–4, pp. 259–268, 1992.

[5] Q. Cheng, H. Shen, L. Zhang, and P. Li, "Inpainting for remotely sensed images with a multichannel nonlocal total variation model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 175–187, 2014.

[6] C. Wu, J. Zhang, and X.-C. Tai, "Augmented Lagrangian method for total variation restoration with non-quadratic fidelity," *Inverse Problems and Imaging*, vol. 5, no. 1, pp. 237–261, 2011.

[7] Y.-W. Wen, R. H. Chan, and A. M. Yip, "A primal-dual method for total-variation-based wavelet domain inpainting," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 106–114, 2012.

[8] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[9] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 89–97, 2004.

[10] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419–2434, 2009.

[11] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, 2011.

[12] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, vol. 87, Springer, Berlin, Germany, 2013.

[13] E. G. Birgin, J. M. Martínez, and M. Raydan, "Spectral projected gradient methods: review and perspectives," *Journal of Statistical Software*, vol. 60, no. 3, pp. 1–21, 2014.

[14] E. G. Birgin, J. M. Martínez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex sets," *SIAM Journal on Optimization*, vol. 10, no. 4, pp. 1196–1211, 2000.

[15] Y. Nesterov, "A method of solving a convex programming problem with convergence rate o $(1/k^2)$," *Soviet Mathematics Doklady*, vol. 27, pp. 372–376, 1983.

[16] D. Goldfarb, S. Ma, and K. Scheinberg, "Fast alternating linearization methods for minimizing the sum of two convex functions," *Mathematical Programming*, vol. 141, no. 1-2, pp. 349–382, 2013.

[17] Y. Ouyang, Y. Chen, G. Lan, and J. Pasiliao Jr., "An accelerated linearized alternating direction method of multipliers," *SIAM Journal on Imaging Sciences*, vol. 8, no. 1, pp. 644–681, 2015.

[18] A. Chambolle, "Total variation minimization and a class of binary MRF models," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, A. Rangarajan, B. Vemuri, and A. L. Yuille, Eds., vol. 3757 of *Lecture Notes in Computer Science*, pp. 136–152, Springer, New York, NY, USA, 2005.

[19] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988.

[20] R. Solimene, A. Cuccaro, A. Dell'Aversano, I. Catapano, and F. Soldovieri, "Ground clutter removal in GPR surveys," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 3, pp. 792–798, 2014.

[21] E. Izquierdo-Verdiguier, R. Jenssen, L. Gómez-Chova, and G. Camps-Valls, "Spectral clustering with the probabilistic cluster kernel," *Neurocomputing*, vol. 149, no. 5, pp. 1299–1304, 2015.

[22] J. C. Jiménez-Muñoz, J. A. Sobrino, and A. R. Gillespie, "Surface emissivity retrieval from airborne hyperspectral scanner data: insights on atmospheric correction and noise removal," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 2, pp. 180–184, 2012.

[23] D. Tuia, M. Volpi, L. Copa, M. Kanevski, and J. Muñoz-Marí, "A survey of active learning algorithms for supervised remote sensing image classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 606–617, 2011.