J Glob Optim
DOI 10.1007/s10898-012-9999-8

An improved first-order primal-dual algorithm with a new correction step

Xingju Cai · Deren Han · Lingling Xu

Received: 26 May 2012 / Accepted: 13 October 2012

Abstract In this paper, we propose a new correction strategy for some first-order primal-dual algorithms arising from solving, e.g., total variation image restoration. With this strategy, we can prove the convergence of the algorithm under more flexible conditions than those proposed most recently. Some preliminary numerical results of image deblurring support that the new correction strategy can improve the numerical efficiency.

Keywords Saddle-point problems · Primal-dual · Contraction step · Variational inequalities

1 Introduction

Let $\mathcal{X} \subseteq \mathcal{R}^L$, $\mathcal{Y} \subseteq \mathcal{R}^N$ be two nonempty, closed and convex sets, $b \in \mathcal{R}^N$, and $A \in \mathcal{R}^{N\times L}$, $B \in \mathcal{R}^{N\times N}$ are two matrices. We consider the saddle-point problem

$$\min_{y\in\mathcal{Y}} \max_{x\in\mathcal{X}} \Phi(x, y) := \langle y, Ax\rangle + \frac{\nu}{2}\|By - b\|^2, \tag{1.1}$$

where $\nu > 0$, $\langle\cdot,\cdot\rangle$ denotes the inner products of vectors and $\|\cdot\|$ is the Euclidean norm. This problem arises from a number of areas such as constrained optimization duality, zero-sum games, and general equilibrium theory. It has been attracted much attention of researchers recently, e.g., [4,6–9,11], just named a few.

For solving such a problem, the following first-order primal-dual iterative scheme has recently attracted most attentions of researchers, especially those from image processing:

X. Cai
Mathematical department, Nanjing university, Nanjing 210023, P. R. China
e-mail: caimaomao76@sina.com

X. Cai · D. Han (✉) · L. Xu
Jiangsu Key Laboratory for NSLSCS, School of Mathematical Sciences,
Nanjing Normal University, Nanjing 210023, P. R. China
e-mail: handeren@njnu.edu.cn

L. Xu
e-mail: xulingling@njnu.edu.cn

Published online: 21 October 2012

Algorithm 1.1: The primal-dual procedure for (1.1) Let $\tau > 0$, $\sigma > 0$ and $\theta \in \mathcal{R}$. For given (x^k, y^k), the new iterate (x^{k+1}, y^{k+1}) is generated by:

$$\begin{cases} x^{k+1} = \operatorname{Arg}\max_{x\in\mathcal{X}}\{\tau\Phi(x, y^k) - \frac{1}{2}\|x - x^k\|^2\}, & (1.2a) \\ \bar{x}^k = x^{k+1} + \theta(x^{k+1} - x^k), & (1.2b) \\ y^{k+1} = \operatorname{Arg}\min_{y\in\mathcal{Y}}\{\sigma\Phi(\bar{x}^k, y) + \frac{1}{2}\|y - y^k\|^2\}. & (1.2c) \end{cases}$$

The parameters τ, σ and θ play a crucial role for the efficiency of the algorithm. Thus, it is important to prove the convergence of the algorithm under conditions as flexible as possible. When $\theta = 0$, this procedure reduces to the primal-dual hybrid gradient method [12], which is the well-known Arrow-Hurwicz-Uzawa method [1]. With some restrictions on τ and σ, which are ensured when τ and σ are sufficiently small, the convergence of this procedure is guaranteed. Note that here τ and σ play the role of step sizes in the primal and dual steps, respectively, and small choice of them may cause slow convergence speed.

Recently, Chambolle and Pock [3] suggest to relax the choice of θ from $\theta = 0$ to $\theta \in [0, 1]$. This flexible choice for θ make them to be able to prove the global convergence of the procedure under conditions less restrictive than those in [1,12]. Specially, when $\theta = 1$, Chambolle and Pock prove the global convergence of the primal-dual procedure under the condition that the step sizes satisfy

$$\tau\sigma < \frac{1}{\|A^T A\|}. \tag{1.3}$$

Here and throughout the paper, for a matrix B, $\|B\|$ denotes its norm

$$\|B\| := \max\{\|Bx\| : \ \|x\| \le 1\}.$$

The condition (1.3) avoids the difficulty of choosing very small step sizes and can enhance the numerical performance of the algorithm.

Most recently, He and Yuan [10] analyzed the convergence of the above primal-dual procedure and proposed a new modification. Under the help of an additional simple correction step, they proved that

- the range of θ can be enlarged to $[-1, 1]$;
- When $\theta = -1$, the step size τ and σ can be arbitrary positive numbers. When $\theta \in (-1, 1]$, they proved the convergence of their new procedure under the requirement

$$\tau\sigma\frac{(1+\theta)^2}{4} < \frac{1}{\|A^T A\|}. \tag{1.4}$$

In this paper, we further relax the requirements of the parameters θ, τ and σ. Our results are

- the parameter θ can be arbitrary;
- When $\theta = -1$, the step size τ and σ can be arbitrary positive numbers; otherwise τ and σ should satisfy

$$\tau\|A^T A\|\frac{(1+\theta)^2}{4}\frac{\sigma}{1+\sigma\nu\lambda_{\min}(B^T B)} < 1. \tag{1.5}$$

Note that our requirements are much flexible than those in [10], which are important from numerical point of view.

This paper is organized as follows. In Section 2, we summarize some necessary preliminaries. In Section 3, we describe our method formally and present its convergence. We reported some numerical results on our algorithm in Section 4 and conclude the paper with Section 5.

2 Preliminaries

In this section, we describe some useful preliminaries for our method and subsequent convergence analysis.

Let (x^*, y^*) be a solution of the saddle point problem (1.1). Then,

$$\begin{aligned}\max_{x\in\mathcal{X}}\{\langle y^*, Ax\rangle + \frac{\nu}{2}\|By^* - b\|^2\} \le \langle y^*, Ax^*\rangle \\ +\frac{\nu}{2}\|By^* - b\|^2 \le \min_{y\in\mathcal{Y}}\{\langle y, Ax^*\rangle + \frac{\nu}{2}\|By - b\|^2\},\end{aligned}$$

which is equivalent to the following variational inequalities

$$\begin{cases} x^* \in \mathcal{X},\ \langle x - x^*, -A^T y^*\rangle \ge 0, & \forall x \in \mathcal{X}, \\ y^* \in \mathcal{Y},\ \langle y - y^*, Ax^* + \nu B^T(By^* - b)\rangle \ge 0, & \forall y \in \mathcal{Y}. \end{cases}$$

Denote

$$u := \begin{pmatrix} x \\ y \end{pmatrix},\ F(u) := \begin{pmatrix} -A^T y \\ Ax + \nu B^T(By - b) \end{pmatrix}, \quad \text{and} \quad \Omega := \mathcal{X} \times \mathcal{Y},$$

the above variational inequalities can be rewritten in the compact form of finding $u^* \in \Omega$, such that

$$\langle u - u^*, F(u^*)\rangle \ge 0, \quad \forall u \in \Omega. \tag{2.1}$$

The mapping F is monotone, i.e.,

$$\langle u - v, F(u) - F(v)\rangle \ge 0, \quad \forall u, v.$$

Let $P_\Omega(\cdot)$ denote the projection onto Ω under the Euclidean norm, i.e.,

$$P_\Omega(v) := \arg\min\{\|u - v\| :\ u \in \Omega\}.$$

Then it follows that for any $w \in \Omega$,

$$\langle u - P_\Omega(u), w - P_\Omega(u)\rangle \le 0. \tag{2.2}$$

The following lemma is due to He and Yuan [10].

Lemma 2.1 *For given $u^k = (x^k, y^k)$, let $u^{k+1} = (x^{k+1}, y^{k+1})$ be generated by the primal-dual procedure (1.2a)–(1.2c). Then, we have*

$$u^{k+1} \in \Omega,\ \langle u - u^{k+1}, F(u^{k+1}) + M(u^{k+1} - u^k)\rangle \ge 0, \quad \forall u \in \Omega, \tag{2.3}$$

where

$$M := \begin{pmatrix} \frac{1}{\tau} I & A^T \\ \theta A & \frac{1}{\sigma} I \end{pmatrix}_{(L+N)\times(L+N)}.$$

To prove the convergence of the primal-dual procedure under flexible requirements on the parameters θ, τ and σ, He and Yuan [10] introduce a simple correction step to the algorithm. In other words, for any given $u^k = (x^k, y^k)$, they denote the point generated by (1.2a) and (1.2c) by $\tilde{x}^k$ and $\tilde{y}^k$ and let $\tilde{u}^k = (\tilde{x}^k, \tilde{y}^k)$, and suggest to generate the next iterate via

$$u^{k+1} = u^k - \alpha_k M(u^k - \tilde{u}^k), \tag{2.4}$$

where

$$\alpha_k := \gamma \frac{\langle u^k - \tilde{u}^k, M(u^k - \tilde{u}^k)\rangle}{\|M(u^k - \tilde{u}^k)\|^2}, \tag{2.5}$$

and $\gamma \in (0, 2)$ is a relaxation parameter. When $\theta = -1$ and τ and σ are any positive numbers, or when $\theta \in (-1, 1]$ and τ and σ satisfied (1.4), they prove that the sequence $\{u^k\}$ generated by the recursion (2.4) converges globally to a solution of the saddle point problem (1.1).

3 The algorithm and its convergence

As we have stated in the introduction, the purpose of this paper is to introduce a new correction step other than (2.4), and prove its global convergence with weaker assumptions than those made in [10]. To this end, we first note that since $\tilde{u}^k$ is generated by (1.2a)–(1.2c) and $u^* \in \Omega$, it follows from (2.3) that

$$\langle u^* - \tilde{u}^k, F(\tilde{u}^k) + M(\tilde{u}^k - u^k)\rangle \geq 0. \tag{3.6}$$

On the other hand, since u^* is a solution of (2.1) and $\tilde{u}^k \in \Omega$,

$$\langle \tilde{u}^k - u^*, F(u^*)\rangle \geq 0. \tag{3.7}$$

Adding (3.6) and (3.7), we obtain

$$\langle u^* - \tilde{u}^k, (F(\tilde{u}^k) - F(u^*)) + M(\tilde{u}^k - u^k)\rangle \geq 0. \tag{3.8}$$

Recall that u and F are respectively defined as

$$u := \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{and} \quad F(u) := \begin{pmatrix} -A^T y \\ Ax + \nu B^T(By - b) \end{pmatrix},$$

we have

$$\begin{aligned} &\langle u^* - \tilde{u}^k, F(\tilde{u}^k) - F(u^*)\rangle \\ &\quad = \langle x^* - \tilde{x}^k, -A^T\tilde{y}^k + A^T y^*\rangle + \langle y^* - \tilde{y}^k, A\tilde{x}^k \\ &\qquad + \nu B^T(B\tilde{y}^k - b) - (Ax^* + \nu B^T(By^* - b))\rangle \\ &\quad = \langle y^* - \tilde{y}^k, \nu B^T B(\tilde{y}^k - y^*)\rangle. \end{aligned} \tag{3.9}$$

Inserting (3.9) into (3.8), we obtain

$$\langle u^* - \tilde{u}^k, M(\tilde{u}^k - u^k)\rangle + \langle y^* - \tilde{y}^k, \nu B^T B(\tilde{y}^k - y^*)\rangle \geq 0. \tag{3.10}$$

We have the following lemma.

Lemma 3.1 *For given $u^k = (x^k, y^k)$, let $\tilde{u}^k = (\tilde{x}^k, \tilde{y}^k)$ be generated by the primal-dual procedure* (1.2a)–(1.2c). *Then, we have*

$$\langle u^k - u^*, \widehat{M}(u^k - \tilde{u}^k)\rangle \geq \langle u^k - \tilde{u}^k, \widetilde{M}(u^k - \tilde{u}^k)\rangle, \tag{3.11}$$

where

$$\widehat{M} := \begin{pmatrix} \frac{1}{\tau}I & A^T \\ \theta A & \frac{1}{\sigma}I + 2\nu B^T B \end{pmatrix}_{(L+N)\times(L+N)} \quad \text{and} \quad \widetilde{M} := \begin{pmatrix} \frac{1}{\tau}I & A^T \\ \theta A & \frac{1}{\sigma}I + \nu B^T B \end{pmatrix}_{(L+N)\times(L+N)}.$$

Proof Note that

$$\begin{aligned} &\langle y^* - \tilde{y}^k, \nu B^T B(\tilde{y}^k - y^*)\rangle \\ &\quad = -\nu\|B(y^k - y^*)\|^2 + 2\nu\langle y^* - y^k, B^T B(\tilde{y}^k - y^k)\rangle - \nu\|B(y^k - \tilde{y}^k)\|^2 \\ &\quad \le 2\nu\langle y^* - y^k, B^T B(\tilde{y}^k - y^k)\rangle - \nu\|B(y^k - \tilde{y}^k)\|^2. \end{aligned}$$

Substituting it into (3.10) and rearranging terms, we have

$$\begin{aligned} &\langle u^k - u^*, M(u^k - \tilde{u}^k)\rangle + \langle y^k - y^*, 2\nu B^T B(y^k - \tilde{y}^k)\rangle \ge \langle u^k - \tilde{u}^k, M(u^k - \tilde{u}^k)\rangle \\ &\quad +\nu\|B(y^k - \tilde{y}^k)\|^2, \end{aligned} \tag{3.12}$$

which is exactly the inequality we need. □

The results in the above lemma inspire us to use $-\widehat{M}(u^k - \tilde{u}^k)$ as descent direction and generate the next iterate via

$$u^{k+1} = u^k - \alpha_k \widehat{M}(u^k - \tilde{u}^k), \tag{3.13}$$

where

$$\alpha_k := \gamma \frac{\langle u^k - \tilde{u}^k, \widetilde{M}(u^k - \tilde{u}^k)\rangle}{\|\widehat{M}(u^k - \tilde{u}^k)\|^2}. \tag{3.14}$$

The following result indicates the rationality in choosing the step size α_k

Theorem 3.1 *Suppose that*

$$\tau\|A^T A\|\frac{(1+\theta)^2}{4}\frac{\sigma}{1+\sigma\nu\lambda_{\min}(B^T B)} < 1. \tag{3.15}$$

Then the step size $\{\alpha_k\}$ *generated by* (3.14) *is uniformly bounded away from zero, i.e., there is* $\alpha_{\min} > 0$, *such that*

$$\alpha_k \ge \alpha_{\min} > 0, \quad \forall k \ge 0. \tag{3.16}$$

Proof From the definition of $\widetilde{M}$, it follows

$$\begin{aligned} &\langle u^k - \tilde{u}^k, \widetilde{M}(u^k - \tilde{u}^k)\rangle \\ &\quad = \frac{1}{\tau}\|x^k - \tilde{x}^k\|^2 + (1+\theta)\langle x^k - \tilde{x}^k, A^T(y^k - \tilde{y}^k)\rangle + \frac{1}{\sigma}\|y^k - \tilde{y}^k\|^2 + \nu\|B(y^k - \tilde{y}^k)\|^2 \\ &\quad \ge \frac{1}{\tau}\|x^k - \tilde{x}^k\|^2 + (1+\theta)\langle x^k - \tilde{x}^k, A^T(y^k - \tilde{y}^k)\rangle + (\frac{1}{\sigma} + \nu\lambda_{\min}(B^T B))\|y^k - \tilde{y}^k\|^2. \end{aligned} \tag{3.17}$$

Define the parameter *delta* to be the positive scalar satisfying the equality

$$\tau(1+\delta)\|A^T A\|\frac{(1+\theta)^2}{4} = \left[\frac{1}{\sigma} + \nu\lambda_{\min}(B^T B)\right]\frac{1}{(1+\delta)},$$

or equivalently,

$$\delta := \frac{2}{1+\theta}\sqrt{\frac{1+\sigma\nu\lambda_{\min}(B^T B)}{\tau\sigma\|A^T A\|}} - 1.$$

Since for any two vectors a and b and a positive parameter μ,

$$2\langle a, b\rangle \geq -\mu\|a\|^2 - \frac{1}{\mu}\|b\|^2,$$

we have

$$\begin{aligned}
&(1+\theta)\langle x^k - \tilde{x}^k, A^T(y^k - \tilde{y}^k)\rangle \\
&\geq -(\tau(1+\delta)\|A^TA\|)\frac{(1+\theta)^2}{4}\|y^k - \tilde{y}^k\|^2 - \frac{1}{\tau(1+\delta)\|A^TA\|}\|A(x^k - \tilde{x}^k)\|^2 \\
&= -\left[\frac{1}{\sigma} + \nu\lambda_{\min}(B^TB)\right]\frac{1}{1+\delta}\|y^k - \tilde{y}^k\|^2 - \frac{1}{\tau(1+\delta)\|A^TA\|}\|A(x^k - \tilde{x}^k)\|^2 \\
&\geq -\frac{1}{(1+\delta)}\left(\frac{1}{\tau}\|x^k - \tilde{x}^k\|^2 + \left[\frac{1}{\sigma} + \nu\lambda_{\min}(B^TB)\right]\|y^k - \tilde{y}^k\|^2\right),
\end{aligned} \tag{3.18}$$

where the equality and the last inequality follow from the definition of δ. Note that the condition (3.15) ensures that $\delta > 0$. Combining (3.17) and (3.18), we get

$$\begin{aligned}
&\langle u^k - \tilde{u}^k, \widetilde{M}(u^k - \tilde{u}^k)\rangle \\
&\geq \frac{\delta}{1+\delta}\left(\frac{1}{\tau}\|x^k - \tilde{x}^k\|^2 + \left[\frac{1}{\sigma} + \nu\lambda_{\min}(B^TB)\right]\|y^k - \tilde{y}^k\|^2\right).
\end{aligned} \tag{3.19}$$

On the other hand, using the Cauchy-Schwarz inequality, we can see that there is a constant C, such that

$$\|\widehat{M}(u^k - \tilde{u}^k)\|^2 \leq C\left(\frac{1}{\tau}\|x^k - \tilde{x}^k\|^2 + \left[\frac{1}{\sigma} + \nu\lambda_{\min}(B^TB)\right]\|y^k - \tilde{y}^k\|^2\right).$$

Consequently, for any k,

$$\alpha_k \geq \frac{\gamma\delta}{(1+\delta)C},$$

which completes the proof. □

Theorem 3.2 *Suppose that (3.15) holds. Then the sequence $\{u^k\}$ generated by (3.13) converges to a solution of the saddle point problem (1.1) globally.*

Proof It follows from (3.13) that

$$\begin{aligned}
&\|u^{k+1} - u^*\|^2 \\
&= \|u^k - u^* - \alpha_k\widehat{M}(u^k - \tilde{u}^k)\|^2 \\
&= \|u^k - u^*\|^2 - 2\alpha_k\langle u^k - u^*, \widehat{M}(u^k - \tilde{u}^k)\rangle + \alpha_k^2\|\widehat{M}(u^k - \tilde{u}^k)\|^2 \\
&\leq \|u^k - u^*\|^2 - 2\alpha_k\langle u^k - \tilde{u}^k, \widetilde{M}(u^k - \tilde{u}^k)\rangle + \alpha_k^2\|\widehat{M}(u^k - \tilde{u}^k)\|^2 \\
&= \|u^k - u^*\|^2 - (2-\gamma)\alpha_k\langle u^k - \tilde{u}^k, \widetilde{M}(u^k - \tilde{u}^k)\rangle \\
&\leq \|u^k - u^*\|^2 - \alpha_{\min}(2-\gamma)\frac{\delta}{1+\delta}\left(\frac{1}{\tau}\|x^k - \tilde{x}^k\|^2 + \left[\frac{1}{\sigma} + \nu\lambda_{\min}(B^TB)\right]\|y^k - \tilde{y}^k\|^2\right),
\end{aligned}$$

where the last inequality follows from (3.16) and (3.19). The above inequality means the $\{u^k\}$ is Fejér monotone and the assertion follows immediately from the results in [5,2]. □

4 Numerical experiments

In this section, we apply the improved primal-dual algorithm (denoted as "IPDA") to solve some TV image deblurring problems, and report some preliminary numerical results. For this application, the underlying linear operator A in (1.1) is the matrix representation of the discrete gradient operator, and the matrix B in (1.1) is a deconvolution operator. Since the proposed method is an improved version of He and Yuan's [10] Algorithm 1 (see (2.4)–(2.5) and denote it as "HYPD" for short), we only compare "IPDA" method with "HYPD" to demonstrate the improvements of our algorithm.

Without loss of generality, the quality of restored images is measured by the value of signal-to-noise ratio (SNR) defined by

$$\text{SNR} := 20\log_{10}\frac{\|\hat{y}\|}{\|\bar{y}-\hat{y}\|},$$

where $\bar{y}$ is the image restored by certain algorithm and $\hat{y}$ represent the original image. To report the numerical results, we mainly list the plots of evolutions of SNR value with respect to iterations.

All the proposed algorithms were coded by MATLAB 2010b and all the numerical experiments were run on a HP desktop with Intel-Pentium Dual Core CPU 2.6 GHz and 2G memory.

In our numerical experiments, we test the images 'House.png' (128×128), 'Cameraman.png' (256×256) and 'Barbara.png' (512×512) as shown in Fig. 1. These images are then degraded by convolution and the zero-mean Gaussian noise with variance 4.0×10^{-6}. The blur operator and the additive noise are generated by scripts `fspecial` and `imnoise` in MATLAB Image Processing Toolbox, respectively. Specifically, we set the `motion` blur

Fig. 1 The *first row* lists the original images and the *second row* lists the degraded images. From left to right: 'House.png', 'Cameraman.png' and 'Barbara.png'

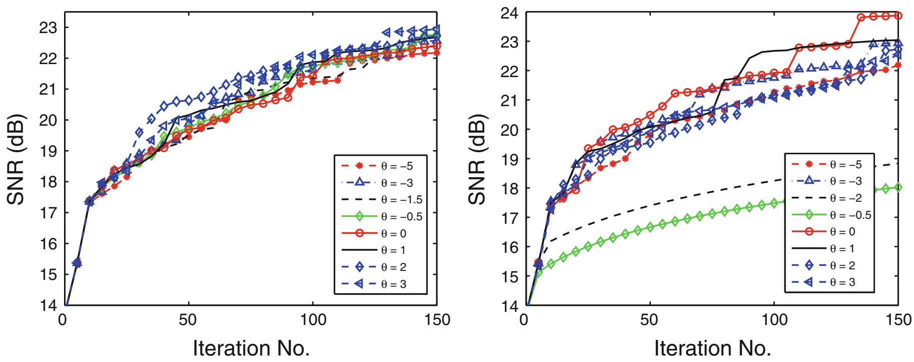

Fig. 2 Sensitivity analysis of parameter θ in "IPDA" for 'Cameraman.png'

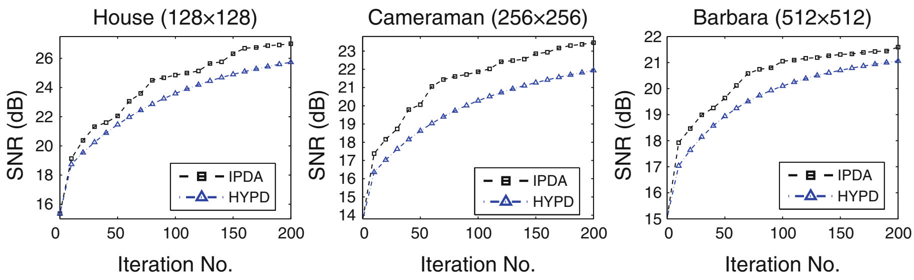

Fig. 3 Evolutions of SNRs with respect to iterations for the tested images. From *left* to *right* corresponding to: 'House.png', 'Cameraman.png' and 'Barbara.png', respectively

with `len` $= 21$ and `theta` $= 135$. The corresponding corrupted images are listed in Fig. 1. Throughout these experiments, we take $\nu = 80$ for the data-fidelity term in (1.1).

For one of our contributions is that we can choose the parameter θ arbitrarily, we first analyze the sensitivity of θ for the improved algorithm. In Fig. 2, the left plot shows the effects of θ with fixed $\tau = 0.02$ and $\sigma = 2$, and the right one lists the effect of θ with $\tau = 0.02$ and $\sigma = \frac{1}{2\tau(1+\theta)^2}$ for the image 'Cameraman.png'. From Fig. 2, we can see that our method is very stable for different θ in many cases.

Secondly, in order to demonstrate the improvements of our new algorithm, we plotted the evolutions of SNRs with respect to iterations for "IPDA" and "HYPD" in Fig. 3. Throughout the following experiments, we set $\tau = 0.02$, $\gamma = 1.5$ and $\sigma = 2$ for both of "IPDA" and "HYPD". In addition, we took $\theta = 1.5$ and $\theta = 0.8$ in "IPDA" and "HYPD", respectively. According to Fig. 3, we can see that the proposed "IPDA" can reach a higher SNR value than "HYPD" with the same iterations. In other words, the proposed method requires less iterations to get the same SNR value than "HYPD". Therefore, this figure supports our improvements numerically.

In practice, we usually do not know the original images. Thus, we use the following rule

$$\frac{\|y^{k+1} - y^k\|}{\|y^{k+1}\|} \leq \varepsilon, \tag{4.20}$$

to be the stopping criterion, where $\{y^k\}$ is the sequence generated by one of the tested algorithms and ε is the error tolerance. Figure 4 lists the corresponding restored images by "IPDA"

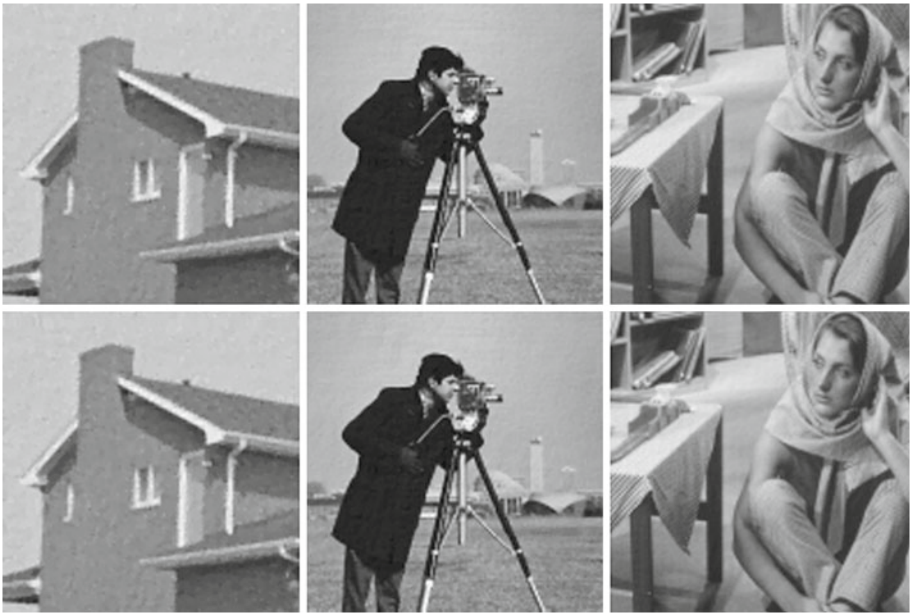

Fig. 4 Restored images by "IPDA" (the *first row*) and "HYPD" (the *second row*) under stopping criterion (4.20) with error tolerance $\varepsilon = 10^{-5}$

and "HYPD" under stopping criterion (4.20) with error tolerance $\varepsilon = 10^{-5}$. It is clear that both of "IPDA" and "HYPD" methods can restore the corrupted images with high quality.

5 Conclusion

We have presented a new correction scheme in the first-order primal-dual algorithms for solving the saddle point problem (1.1). We prove the global convergence of the new scheme under more flexible conditions than the existing literature. That is, we have no restriction of the parameter θ, and the requirement on τ and σ is (3.15). This flexible in choosing the parameters is crucial from numerical point of view. Some preliminary numerical experiments on image deblurring supports our improvements for He and Yuan's Algorithm 1 [10].

Note that in He and Yuan [10], the authors extended their algorithm to generic saddle point problems. Since Lemma 3.1 depends on the special structure of the problem, it seems not possible to extend our algorithm to generic saddle point problems.

Acnowledgments The authors thank the anonymous referee for the constructive comments, which help us improve the presentation of the paper. The research is supported by the NSFC grants 11071122 and 11171159.

References

1. Arrow, K.J., Hurwicz, L., Uzawa, H.: With contributions by H.B. Chenery, S.M. Johnson, S. Karlin, T. Marschak, and R.M. Solow. Studies in Linear and Non-Linear Programming, volume II of Stanford Mathematical Studies in the Social Science. Stanford Unversity Press, Stanford, (1958)
2. Censor, Y., Elfving, T.: New methods for linear inequalities. Linear Algebra It's Appl. **42**, 199–211 (1982)

3. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. **40**, 120–145 (2011)
4. Esser, E., Zhang, X., Chan, T.: A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science. SIAM J. Imaging Sci. **3**, 1015–1046 (2010)
5. Gubin, L.G., Polyak, B.T., Raik, E.V.: The method of projections for finding the common point of convex sets. USSR Comput. Math. Math. Phys. **7**, 1–24 (1967)
6. Han, D., Xu, W., Yang, H.: An operator splitting method for variational inequalities with partially unknown mappings. Numer. Math. **111**(2), 207–237 (2008)
7. He, B., Liao, L.Z., Han, D., Yang, H.: A new inexact alternating directions method for monotone variational inequalities. Math. Progr. **92**(1), 103–118 (2002)
8. He, B., Tao, M., Yuan, X.: Alternating direction method with gaussian back substitution for separable convex programming. SIAM J. Optim. **22**(2), 313–340 (2012)
9. He, B., Yuan, X.: On the $O(1/n)$ convergence rate of the douglas-rachford alternating direction method. SIAM J. Numer. Anal. **50**(2), 700–709 (2012)
10. He, B.S., Yuan, X.M.: Convergence analysis of primal-dual algorithms for a saddle-point problem: From contraction perspective. SIAM J. Imaging Sci. **5**, 119–149 (2012)
11. Nemirovski, A.: Prox-method with rate of convergence O(1/t) for variational inequalities with lipschitz continuous monotone operators and smooth convex-concave saddle point problems. SIAM J. Optim. **15**(1), 229–251 (2004)
12. Zhu, M.Q., Chan, T.: An efficient primal-dual hybrid gradient algorithm for total variation image restoration. CAM reports 08–34, UCLA, (2008)