Journal of Computational and Applied Mathematics 235 (2011) 2766–2774

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Restoration of images based on subspace optimization accelerating augmented Lagrangian approach

Dai-Qiang Chen *, Li-Zhi Cheng, Fang Su

Department of Mathematics, College of Science, National University of Defense and Technology, Changsha 410073, Hunan, People's Republic of China

ARTICLE INFO

Article history:
Received 1 July 2010
Received in revised form 18 November 2010

Keywords:
Total variation
Augmented Lagrangian method
Subspace optimization
Douglas–Rachford splitting

ABSTRACT

We propose a new fast algorithm for solving a TV-based image restoration problem. Our approach is based on merging subspace optimization methods into an augmented Lagrangian method. The proposed algorithm can be seen as a variant of the ALM (Augmented Lagrangian Method), and the convergence properties are analyzed from a DRS (Douglas–Rachford splitting) viewpoint. Experiments on a set of image restoration benchmark problems show that the proposed algorithm is a strong contender for the current state of the art methods.

1. Introduction

Restoration of digital images based on variational models and optimization techniques has been extensively studied for the last 20 years in many areas of image processing and computer vision, such as medical imaging [1], astronomy [2,3], and compressed sensing [4,5]. Image restoration in such fields of applications can often be formulated as linear inverse and ill-posed problems. Let $u_0 \in R^n$ denote the original image, and $f \in R^m$ denote the observed image (or measurements) with noise. The general model is given by

$$f = Ku_0 + \varepsilon \tag{1.1}$$

where ε is a white gaussian noise, and $K \in R^{m\times n}$ is a linear operator, typically a convolution operator in the image deconvolution problem or the composition of a sampling matrix and a specific transform matrix (such as Fourier transform) in the compressed sensing. These are two examples considered in this study.

By using an MAP (maximum a posteriori) estimator, we can derive the following variational problem for image restoration problem

$$\operatorname{argmin}_u \left\{ \frac{1}{2}\|Ku - f\|_2^2 + \tau\phi(u) \right\} \tag{1.2}$$

where $\phi(u)$ denotes the regularizer, which is usually convex but non-smooth, such as TV (total variation) or l_1 norms of frame coefficients, and $\tau > 0$ is the regularization parameter. In this paper, we mainly discuss TV-based image restoration problems. Thus for the image prior we adopt the TV function [6]. In other words, the regular term in the variational problem (1.2) is given by

$$\phi(u) = TV(u) = \|Du\|_1 = \sum_i |D_1u|_i + |D_2u|_i \tag{1.3}$$

where $D = (D_1; D_2)$ and $(D_1u)_i$, $(D_2u)_i$ denote the horizontal and vertical first order differences at pixel i (we use the definition of anisotropic total variation here, but the method proposed in the paper is also suited for isotropic total variation).

* Corresponding author. Tel.: +86 13467530489.
E-mail addresses: chener050@sina.com (D.-Q. Chen), clzcheng@vip.sina.com (L.-Z. Cheng).

0377-0427/$ – see front matter Crown Copyright © 2010 Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2010.11.028

The TV-based model has become one of the standard techniques known for preserving sharp discontinuities. Recently, various algorithms were developed on the basis of convex optimization techniques to compute the minimizer of the variational problem (1.2), such as SpaRSA (sparse reconstruction by separable approximation) [7], Nesterovs algorithm [8], FISTA (fast iterative shrinkage/thresolding algorithm) [9], TwIST (two-step IST) [10]. These methods were shown to be considerably faster than earlier methods, including interior-point methods [11,12] and iterative thresholding ideas [13,14]. Very recently, a new algorithm named SALSA (split augmented Lagrangian shrinkage algorithm) [15] was proposed and shown to be more efficient than TwIST, FISTA, and SpaRSA. SALSA can be seen as an application of variable splitting methods, and then the obtained constrained problem is tackled with an augmented Lagrangian (AL) scheme [16]. In other words, SALSA is a variant of the alternating direction method of multipliers (ADMM).

In this paper, we propose a new variant of augmented Lagrangian method using subspace optimization accelerating [17,18]. It also can be identified as a variant of split Bregman iteration [19]. Compared with SALSA applied to TV-based image restoration, our proposed algorithm is implemented without solving the sub-minimization problem of TV-denoising by Chambolle's projection algorithm [20]. Besides, by adopting subspace optimization accelerating technology, the proposed algorithm may be more efficient than SALSA, which belongs to the family of augmented Lagrangian algorithms. Experiments demonstrate that the proposed algorithm is efficient, and it is shown to be faster than SALSA.

The rest of this paper is organized as follows. In Section 2, we briefly review the augmented Lagrangian approach and propose our algorithm which can be regarded as a variant of the augmented Lagrangian method. In Section 3, the convergence properties of the proposed algorithm are researched from a DRS (Douglas–Rachford splitting) viewpoint. In Section 4, in the context of image deconvolution and compressed sensing, the proposed algorithm is implemented and compared with other available methods. The conclusion about the proposed algorithm is given in Section 5.

2. Augmented Lagrangian approach using Subspace Optimization accelerating

2.1. Augmented Lagrangian approach

Let us start with a general constrained optimization problem of the form

$$\min_{d\in R^m, u\in R^n} J(d) + H(u), \quad \text{s.t. } Bu = d \tag{2.1}$$

where $J : R^m \to R, H : R^n \to R$ are closed, proper and convex functions, and $B \in R^{m\times n}$ is a bounded linear operator.

The augmented Lagrangian function [15,16] for the problem (2.1) is given by

$$L^\mu(u, d, r) = J(d) + H(u) - \langle r, d - Bu\rangle + \frac{\mu}{2}\|Bu - d\|_2^2 \tag{2.2}$$

where $r \in R^m$ is a vector of Lagrange multipliers and $\mu > 0$ is the penalty parameter.

The so-called augmented Lagrangian method (ALM), also known as the method of multipliers (MM), is obtained by applying the general Uzawa algorithm [21] to the augmented Lagrangian formulation. More precisely, for fixed r we solve the minimizer of the Lagrangian function $L^\mu(u, d, r)$ with respect to $(u; d)$, and then for fixed $(u; d)$ we obtain r through one iterative step of gradient ascent method. By alternating these two steps we get a sequence (u^k, d^k, r^k) as follows:

$$\begin{cases} (u^{k+1}, d^{k+1}) = \text{argmin}_{u,d}\left(J(d) + H(u) - \langle r^k, d - Bu\rangle + \frac{\mu}{2}\|Bu - d\|_2^2\right) \\ r^{k+1} = r^k + \mu(Bu^{k+1} - d^{k+1}). \end{cases} \tag{2.3}$$

Let $r^k = \mu b^k$, then the ALM/MM can be rewritten as follows

$$\begin{cases} (u^{k+1}, d^{k+1}) = \text{argmin}_{u,d}\left(J(d) + H(u) + \frac{\mu}{2}\|Bu - d + b^k\|_2^2\right) \\ b^{k+1} = b^k + (Bu^{k+1} - d^{k+1}). \end{cases} \tag{2.4}$$

The first executing step in (2.4) is not trivial since it involves non-separable quadratic as well as non-smooth terms. Replacing it by the alternating minimization with respect to each vector leads to a variant of the so-called alternating direction method of multipliers (ADMM)

$$\begin{cases} d^{k+1} = \text{argmin}_d\left(J(d) + \frac{\mu}{2}\|Bu^k - d + b^k\|_2^2\right) \\ u^{k+1} = \text{argmin}_u\left(H(u) + \frac{\mu}{2}\|Bu - d^{k+1} + b^k\|_2^2\right) \\ b^{k+1} = b^k + (Bu^{k+1} - d^{k+1}). \end{cases} \tag{2.5}$$

2.2. Proposed image restoration algorithm

Let $l = (u; d)$. The first two executing steps in (2.5) can be interpreted as an application of block coordinate descent [22] for function $F^{b^k}(l) = J(d) + H(u) + \frac{\mu}{2}\|Bu - d + b^k\|_2^2$. The convergence rate of ADMM may suffer from the influence of the difference between l^k generated by (2.5) and $\tilde{l}^k = \text{argmin}_l F^{b^k}(l)$. Thus we consider to solve the sub-minimization problem $\tilde{l}^k = \text{argmin}_l F^{b^k}(l)$ by merging a recently developed subspace optimization method [17,18] into the alternating direction method, which has no additional complexity and may accelerate ADMM efficiently.

We now return to the unconstrained optimization formulation of regularized image restoration, as defined in (1.2). The corresponding constrained optimization problem is

$$\min_{d\in R^{2n},u\in R^n} \tau\|d\|_1 + \frac{1}{2}\|Ku-f\|_2^2, \quad \text{s.t. } Du = d \tag{2.6}$$

where $\|d\|_1 = \sum_i |d_i|$.

Let $F^b(l) = \tau\|d\|_1 + \frac{1}{2}\|Ku-f\|_2^2 + \frac{\mu}{2}\|Du-d+b\|_2^2$. Assume that the problem (2.6) is to be solved via an iterative algorithm. Instead of solving $\text{argmin}_l F^{b^k}(l)$ by the alternating direction method, we search for the solution of this problem within the subspace spanned by a set of directions $\{t_k^i\}_{i=1}^{L+1}$, that is

$$\alpha^k = \text{argmin}_\alpha F^{b^k}(l^k + S^k\alpha), \tag{2.7}$$

$$l^{k+1} = l^k + S^k\alpha^k \tag{2.8}$$

where the columns of S^k are $\{t_k^i\}_{i=1}^{L+1}$.

First, we discuss the definition of the direction t_k^i in (2.7). Assume that $\tilde{l}^k = (\tilde{u}^k; \tilde{d}^k)$ is generated by the first two executing steps of ADMM for problem (2.6), where $J(d) = \tau\|d\|_1$, $H(u) = \frac{1}{2}\|Ku-f\|_2^2$, $B = D$. We choose $s^k = \tilde{l}^k - l^k$ as the first direction included in S^k. Besides, the subspace may be enriched with L previous propagation directions. Specially we can define $S^k = [s^k, \{r_k^i\}_{i=1}^L]$, where $r_k^i = l^{k-i+1} - l^{k-i}$, $0 < i < L+1$. Then according to (2.7) and (2.7) we get: $F^{b^k}(l^{k+1}) \le F^{b^k}(\tilde{l}^k)$.

Next we consider the subspace optimization problem defined in (2.7). Define the set $C = \{(u; d) \in R^{3n} \,|\, |d_i| \ne 0, 1 \le i \le 2n\}$. F^{b^k} is a smooth function on the set C, and for any $l \in C$, $\nabla^2 F^{b^k}$ is given by

$$\nabla^2 F^{b^k} = \begin{pmatrix} \mu D^T D + K^T K & -\mu D^T \\ -\mu D & \mu I \end{pmatrix}. \tag{2.9}$$

For the subspace optimization problem (2.7) we have

$$(S^k)^T \nabla F^{b^k}(l^k + S^k\alpha^k) = 0. \tag{2.10}$$

If $l^k \in C$, then function $\nabla^2 F^{b^k}$ is independent of variables u and d according to the formula (2.9). Through Taylor series expansion we get

$$\nabla F^{b^k}(l^k + S^k\alpha^k) = \nabla F^{b^k}(l^k) + (\nabla^2 F^{b^k})S^k\alpha^k. \tag{2.11}$$

Together (2.10) with (2.11) we have

$$\alpha^k = -\left((S^k)^T\nabla^2 F^{b^k} S^k\right)^{-1}\left((S^k)^T\nabla F^{b^k}(l^k)\right). \tag{2.12}$$

If $l^k \notin C$, then let $\hat{l}^k = l^k + [0; \varepsilon c]$, where $0 < \varepsilon \ll 1$ and $c = [c_i]_{1\le i\le 2n}^T$ satisfies

$$c_i = \begin{cases} 0, & d_i \ne 0 \\ 1, & d_i = 0. \end{cases}$$

Let $c_0 = [0; c]$. Through Taylor series expansion at the point $\hat{l}^k$ we have

$$\nabla F^{b^k}(l^k + S^k\alpha^k) = \nabla F^{b^k}(\hat{l}^k) + (\nabla^2 F^{b^k})(S^k\alpha^k - \varepsilon c_0). \tag{2.13}$$

Together (2.10) with (2.13) we get

$$\alpha^k = -\left((S^k)^T\nabla^2 F^{b^k} S^k\right)^{-1}\left((S^k)^T(\nabla F^{b^k}(\hat{l}^k) - \varepsilon\nabla^2 F^{b^k} c_0)\right). \tag{2.14}$$

As $\varepsilon \to 0$, we have $\alpha^k \approx -\left((S^k)^T\nabla^2 F^{b^k} S^k\right)^{-1}\left((S^k)^T\nabla F^{b^k}(l^k)\right)$, which is just the value defined in the formula (2.12).

Let $\varphi(d) = \tau\|d\|_1$, and the whole augmented Lagrangian approach using Subspace Optimization accelerating (AL_SOP) can be written as follows

Algorithm 1. AL_SOP for image restoration

Step 1: Initialize: $u^k = f$, $b^k = 0$, $k = 0$;

Step 2: Iteration:

$$\tilde{d}^k = \mathrm{argmin}_d \left(\varphi(d) + \frac{\mu}{2}\|Du^k - d + b^k\|_2^2\right);$$

$$\tilde{u}^k = \mathrm{argmin}_u \left(\frac{1}{2}\|Ku - f\|_2^2 + \frac{\mu}{2}\|Du - \tilde{d}^k + b^k\|_2^2\right);$$

update subspace $S^k = [s^k, \{r_k^i\}_{i=1}^L]$;

$$\alpha^k = -\left((S^k)^T \nabla^2 F^{b^k} S^k\right)^{-1}\left((S^k)^T(\nabla F^{b^k}(\hat{l}^k) - \varepsilon \nabla^2 F^{b^k} c_0)\right);$$

$$l^{k+1} = l^k + S^k\alpha^k;$$

$$b^{k+1} = b^k + Du^{k+1} - d^{k+1};$$

$$k = k+1;$$

until some stopping criterion is satisfied.

3. Convergence analysis

The question arising now is whether the iterative sequence generated by the algorithm proposed in Section 2.2 converges to the solution of problem (2.6). In this section we analyze the convergence properties of the iterative algorithm from the DRS viewpoint. First, we introduce the DRS algorithm proposed in [23].

Considering the minimization problem:

$$\min_{p\in R^n} h_1(p) + h_2(p). \tag{3.1}$$

Define proximity operator $\mathrm{prox}_{\lambda h}(z) = \mathrm{argmin}_x\{\frac{1}{2\lambda}\|x - z\|^2 + h(x)\} = (I + \lambda\partial h)^{-1}(z)$, $A = \partial h_1$, $B = \partial h_2$. We have the following convergence result (for more details we refer to Theorem 7 in [23] and Theorem 19 in [24]).

Proposition 3.1 (*DRS Algorithm [23,24]*)**.** *Let A, B be maximal monotone operators and assume that a solution of (3.1) exists. Let $\lambda > 0$, and $\{\tau_k\}$ be one sequence such that $0 < \inf_{k\in N}\tau_k \le \sup_{k\in N}\tau_k < 2$, $\sum_{k\in N}\tau_k(2-\tau_k) = +\infty$, and $\{\xi^k\}$ and $\{\eta^k\}$ be sequences in R^n. Suppose that $0 \in \mathrm{Ran}(A + B)$, and $\sum_{k\in N}\tau_k(\|\xi^k\|_2 + \|\eta^k\|_2) < +\infty$. Then, for any initial elements t^0, the following DRS algorithm converges to an element t^*:*

$$p^k = \mathrm{prox}_{\lambda h_2}(t^k) + \xi^k, \tag{3.2}$$

$$t^{k+1} = t^k + \tau_k(\mathrm{prox}_{\lambda h_1}(2p^k - t^k) + \eta^k - p^k). \tag{3.3}$$

Furthermore, the sequence $\{p^k\}$ converges to p^* satisfying $0 \in A(p^*) + B(p^*)$.

Considering the constrained optimization problem (2.6). Let $l = (u; d)$, $F(l) = \varphi(d) + \frac{1}{2}\|Ku - f\|_2^2$, $G = (D, -I)$. Then the ALM/MM for (2.6) can be rewritten as follows

$$\begin{cases} l^{k+1} = \mathrm{argmin}_l \left(F(l) + \dfrac{\mu}{2}\|Gl + b^k\|_2^2\right) \\ b^{k+1} = b^k + Gl^{k+1}. \end{cases} \tag{3.4}$$

Based on the idea of Theorem 2.4 in [25] we have the following result.

Theorem 3.2. *Let $t^k = \mu b^k$, where b^k is given by (3.4). Then the iterative sequence $\{t^k\}$ satisfies:*

$$t^{k+1} = \mathrm{prox}_{\mu(F^*\circ(-G^*))}(t^k). \tag{3.5}$$

Proof. First, we show that the following relation holds true

$$\hat{p} = \mathrm{argmin}_{p\in R^n}\left\{\frac{\gamma}{2}\|Ap - q\|_2^2 + h(p)\right\} \Rightarrow \gamma(A\hat{p} - q) = \mathrm{prox}_{\gamma(h^*\circ(-A^*))}(-\gamma q) \tag{3.6}$$

where $h : R^m \to R \cup \{+\infty\}$ is proper, convex and lower semi-continuous (lsc), and $A \in R^{m\times n}$ is a bounded linear operator, and '$*$' denotes the dual operator.

Though the relation (3.6) has been proved in [25], we present the proof here for completeness. The left-hand side of (3.6) is equivalent to

$$0 \in \gamma A^*(A\hat{p} - q) + \partial h(\hat{p}) \tag{3.7}$$

which implies

$$\hat{p} \in \partial h^*(-\gamma A^*(A\hat{p} - q)). \tag{3.8}$$

Applying $-\gamma A$ on both sides of (3.8) and then adding $\gamma(A\hat{p}-q)$ we obtain

$$-\gamma q \in (I+\gamma\partial(h^*\circ(-A^*)))(\gamma(A\hat{p}-q)) \tag{3.9}$$

which is, by definition of proximity operator $\mathrm{prox}_{\lambda h}$, equivalent to the right-hand side of (3.6).

Applying (3.6) to (3.4) with $h=F$, $\hat{p}=l^{k+1}$, $A=G$, $q=-b^k$, $\gamma=\mu$ we get

$$t^{k+1}=\mathrm{prox}_{\mu(F^*\circ(-G^*))}(t^k). \quad \square \tag{3.10}$$

Finally, we analyze the convergence properties of Algorithm 1 based on the above results.

Theorem 3.3. *Let $\{(l^k=(u^k;d^k),b^k)\}$ be sequences generated by Algorithm 1, and let $\alpha^k=l^k-\tilde{l}^k$, where $\tilde{l}^k=\mathrm{argmin}_l(F(l)+\frac{\mu}{2}\|Gl+b^{k-1}\|_2^2)$. Suppose that $\sum_{k\in N}(\|\alpha^k\|_2)<+\infty$, then $\{b^k\}$ converges. Furthermore, if $\{l^k\}$ converges, then the limit of $\{l^k\}$ is the solution of (2.6).*

Proof. Let $\tilde{b}^{k+1}=b^k+G\tilde{l}^{k+1}$. Then $\beta^{k+1}=b^{k+1}-\tilde{b}^{k+1}=G(l^{k+1}-\tilde{l}^{k+1})=G\alpha^{k+1}$. Thus we have $\|\beta^{k+1}\|_2^2=(\alpha^{k+1})^TG^TG\alpha^{k+1}$. Let $c=\rho(G^TG)$ denote the spectral radius of G^TG. Thus we have

$$\|\beta^{k+1}\|_2^2\le c\|\alpha^{k+1}\|_2^2. \tag{3.11}$$

Since $\sum_{k\in N}(\|\alpha^k\|_2)<+\infty$, thus $\sum_{k\in N}(\|\beta^k\|_2)<+\infty$.

Let $t^{k+1}=\mu b^{k+1}$, $\tilde{t}^{k+1}=\mu\tilde{b}^{k+1}$. Then

$$t^{k+1}=\tilde{t}^{k+1}+\mu\beta^{k+1}. \tag{3.12}$$

Since $\tilde{l}^{k+1}=\mathrm{argmin}_l(F(l)+\frac{\mu}{2}\|Gl+b^k\|_2^2)$, and $\tilde{b}^{k+1}=b^k+G\tilde{l}^{k+1}$, applying (3.6) to these equations with $h=F$, $\hat{p}=\tilde{l}^{k+1}$, $A=G$, $q=-b^k$, $\gamma=\mu$ we obtain

$$\tilde{t}^{k+1}=\mathrm{prox}_{\mu(F^*\circ(-G^*))}(t^k). \tag{3.13}$$

Together (3.12) with (3.13) we get

$$t^{k+1}=\mathrm{prox}_{\mu(F^*\circ(-G^*))}(t^k)+\mu\beta^{k+1}. \tag{3.14}$$

The above formula coincides with the DRS algorithm (3.2) and (3.3) for $h_2=0$, $\lambda h_1=\mu(F^*\circ(-G^*))$, $\tau_k=1$, $\xi^k=0$, $\eta^k=\mu\beta^{k+1}$. As $\sum_{k\in N}(\|\beta^k\|_2)<+\infty$, then according to Proposition 3.1 the sequence $\{t^k\}$ converges to an element t^* satisfying

$$t^*=\mathrm{argmin}_t F^*\circ(-G^*)(t). \tag{3.15}$$

Thus the sequence $\{b^k\}$ converges to an element $b^*=\frac{t^*}{\mu}$.

Besides, considering the equation $b^{k+1}=b^k+Gl^{k+1}$ included in Algorithm 1, by passing to the limit we get

$$\lim_{k\to+\infty} Gl^k=0. \tag{3.16}$$

Furthermore, if the sequence $\{l^k\}$ converges to an element l^*, then by the Eq. (3.16) we get $Gl^*=0$. As $\sum_{k\in N}(\|\alpha^k\|_2)<+\infty$, thus $\lim_{k\to+\infty}\|\alpha^k\|_2=0$, and hence the sequence $\{\tilde{l}^k\}$ also converges to l^*. By the definition of $\{\tilde{l}^k\}$ we get

$$\partial F(\tilde{l}^k)+\mu G^T(G\tilde{l}^k+b^{k-1})=0. \tag{3.17}$$

By passing to the limit we get

$$\partial F(l^*)+\mu G^T(Gl^*+b^*)=\partial F(l^*)+G^Tt^*=0. \tag{3.18}$$

Let $L(l,t)=F(l)+<t,Gl>$: that is the Lagrangian function for problem (2.6). Then according to (3.15), (3.18) and Karush–Kuhn–Tucker (KKT) conditions [26] we conclude: $\{l^*\}$ is the solution of (2.6). $\square$

4. Experiment

In this section, we present some numerical results using Algorithm 1 and compare it with that of the current state of the art methods (all of which are freely available online): TWIST[1] [10], FISTA[2] [9] and SALSA[3] [15]. Our programs are performed under Windows XP and MATLAB 7.0 running on a Lenovo laptop with a Dual Intel Pentium CPU 1.8 G and 1 GB of memory.

1 http://www.lx.it.pt/bioucas/code/TwIST_v1.zip.

2 http://iew3.technion.ac.il/becka/papers/wavelet_FISTA.zip.

3 http://cascais.lx.it.pt/mafonso/SALSA_v1.0.zip.

Table 1
TV-based image restoration: CPU times (in seconds).

Experiment	FISTA	TWIST	SALSA	Our algorithm
1	30.453	21.876	11.01	3.6406
2	44.813	20.06	12.0156	2.6875
3	120.77	22.864	9.1406	6.6402
4	35.328	21.025	12.719	6.5156

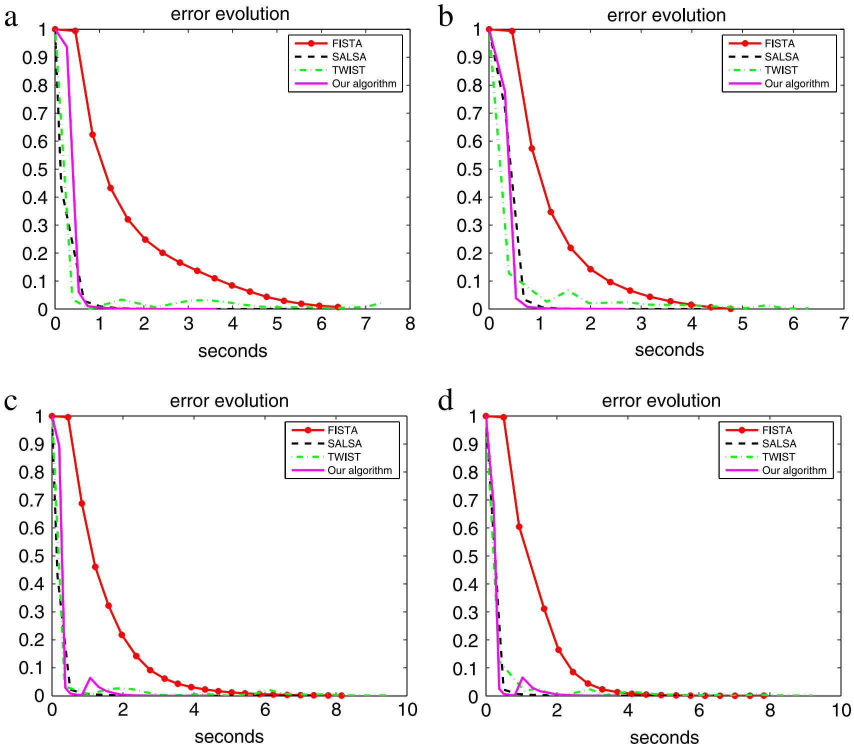

Fig. 1. Error evolution: (a) experiment 1; (b) experiment 2; (c) experiment 3; (d) experiment 4.

4.1. Image deconvolution result

In this subsection we use image Cameraman as the test image, and evaluate the algorithms considered in this paper in four typical image deconvolution scenarios: strong blur with low noise; strong blur with medium noise; mild blur with low noise; mild blur with medium noise, which are summarized below:

- Experiment 1: PSF (point-spread function): $v(x_1, x_2) = 1/(1 + x_1^2 + x_2^2)$, $-7 \le x_1, x_2 \le 7$, gaussian noise with $\sigma^2 = 2$;
- Experiment 2: PSF: $v(x_1, x_2) = 1/(1 + x_1^2 + x_2^2)$, $-7 \le x_1, x_2 \le 7$, gaussian noise with $\sigma^2 = 8$;
- Experiment 3: PSF: 7×7 Gaussian PSF with standard deviation 2, gaussian noise with $\sigma^2 = 2$;
- Experiment 4: PSF: 7×7 Gaussian PSF with standard deviation 2, gaussian noise with $\sigma^2 = 8$;

Set $J(u) = \varphi(Du) + \frac{1}{2}\|Ku - f\|_2^2$, and the stopping criterion we used in these experiments is

$$\|J(u^{k+1}) - J(u^k)\|/\|J(u^k)\| < 10^{-5}. \tag{4.1}$$

In our experiments, we adjust the regularization parameter τ to achieve the best SNR (signal-noise-ratio), and use $\mu = 0.1\tau$ in all experiments following [15]. Besides, we define $\partial|d_i|(0) = 0$ and choose $\varepsilon = 0$ in Algorithm 1. For experiments 1 and 2, we adopt $S^k = [s^k]$ for the subspace optimization problem; for experiments 3 and 4, we use the ADMM algorithm for the first 5 iteration steps and then adopt Algorithm 1 with the subspace $S^k = [s^k]$.

The CPU times taken by FISTA, SALSA, TwIST and our proposed algorithm are listed in Table 1. The evolutions of errors defined by (4.1) are plotted in Fig. 1.

We can conclude from Table 1 and Fig. 1, for a TV-based image restoration problem, our proposed algorithm is clearly faster than the other competing algorithms. Referring to Sections 2 and 3, our proposed algorithm uses the Hessian matrix of the data fidelity term of (1.2) and can be seen as an approximation to (3.5), which converges to an element with an exponential convergence rate assuming $\|\text{prox}_{\mu(F^* \circ (-G^*))}\| < 1$; while the above mentioned algorithms, i.e. FISTA and TWIST, are essentially only variants of gradient descent methods. Furthermore, our proposed algorithm is implemented

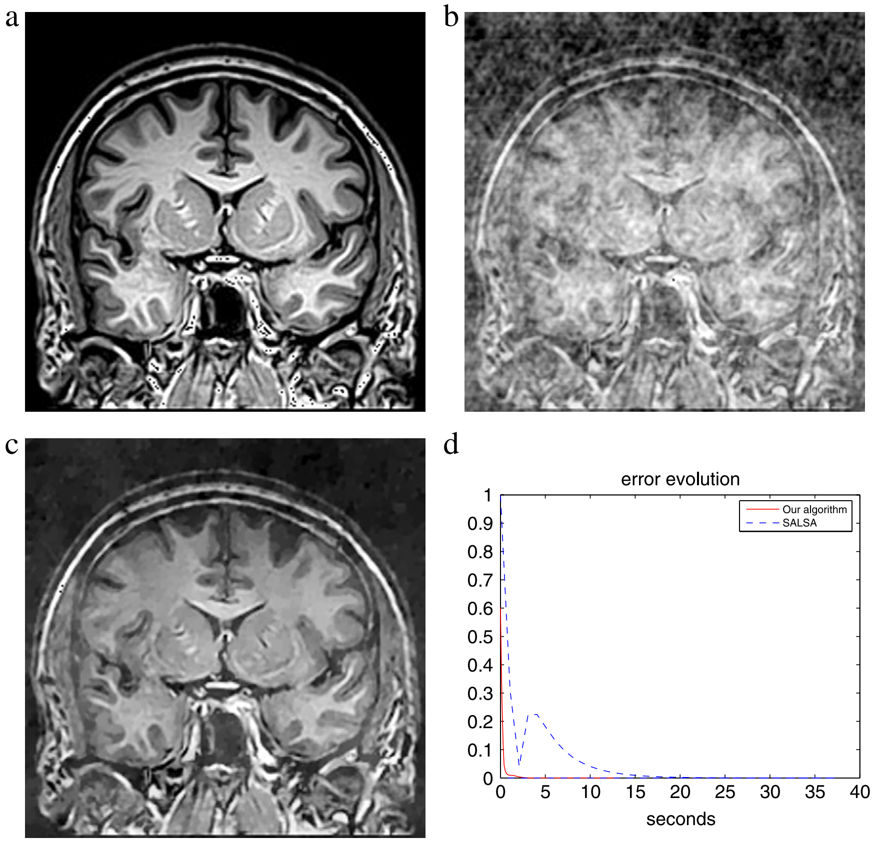

Fig. 2. Compressive sensing example 1: (a) image Brain; (b) 50% randomly chosen Fourier coefficients with Gaussian noise $\sigma = 0.001$, SNR $= 5.62$ dB; (c) recovered result using Algorithm 1, SNR $= 10.06$ dB; (d) error evolution.

without solving the nested minimization problem of TV-denoising compared with SALSA. Besides, by adopting subspace optimization accelerating technology, the proposed algorithm may be faster than SALSA which belongs to the family of augmented Lagrangian algorithms. The analysis of these algorithms is verified by our experiments above.

4.2. Compressed sensing result

We now discuss the magnetic resonance image (MRI) restoration from partial Fourier observations, which has been the focus of much recent interest due to its connection with compressed sensing and the fact that it models MRI acquisition [27,15]. For this example, the linear operator K in model (1.1) is given by

$$K = P\mathcal{F} \tag{4.2}$$

where $\mathcal{F} \in \mathbb{C}^{n\times n}$ represents a Fourier transform matrix, and $P \in R^{p\times n}$ is a sub-sampling matrix containing p rows of a $n \times n$ identity matrix.

In the experiments, we test our algorithm on two real MRI images: Brain and Chest. The CS (Compressed sensing) data is obtained from random sampled Fourier coefficients of MRI images. In other words, the matrix P is a random sampling matrix. Note that, while ideal for compressed sensing, this type of sampling is not practical for most MRI applications. For more detailed information about the selection of sampling matrices we refer to [1,27]. Because the focus of this paper is on numerics, and not on the details of image acquisition, we choose uniform random sampling for simplicity. In the experiments, we adjust the parameters τ and μ following Section 4.1. For Algorithm 1, we adopt $R^k = [s^k]$ for the subspace optimization problem. Figs. 2 and 3 show the results of compressed sensing for images of the Brain and Chest. Figs. 2(d) and 3(d) show the the evolutions of errors defined in (4.1). Again, we may conclude our algorithm is faster than SALSA, which is also the fastest of the other competing algorithms mentioned above.

5. Conclusion

In this article, by merging a recently developed subspace optimization method into the alternating direction method, we propose a new iterative algorithm for image restoration. Then we analyze the convergence properties of the iterative algorithm from a DRS viewpoint. Experiments demonstrate our algorithm is clearly faster than recently proposed methods.

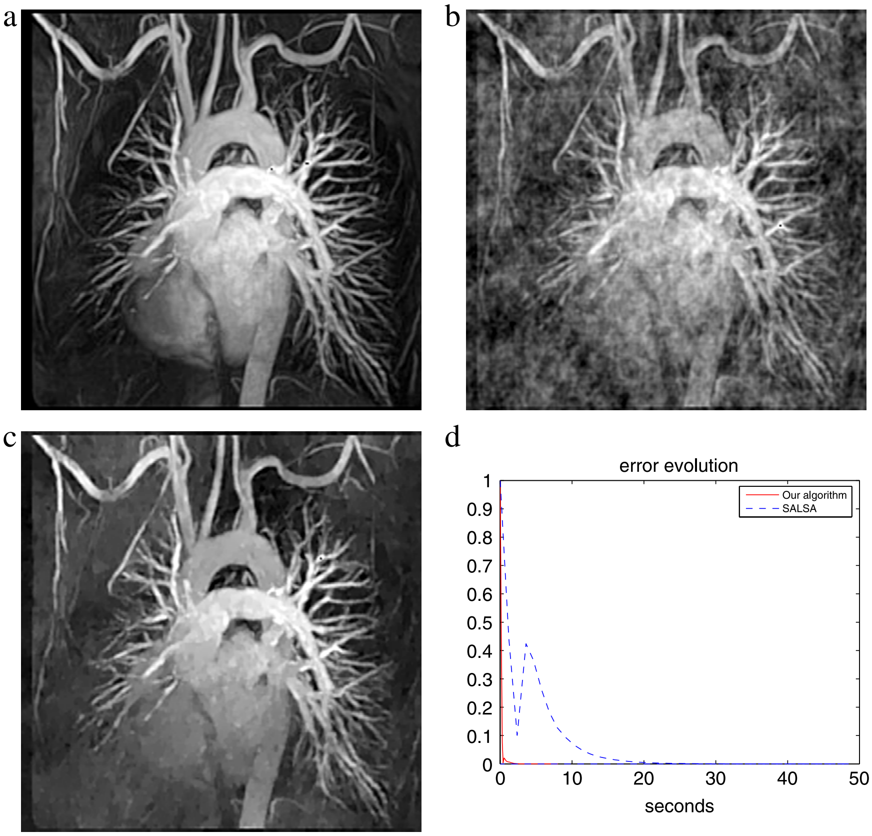

Fig. 3. Compressive sensing example 2: (a) image Chest; (b) 50% randomly chosen Fourier coefficients with Gaussian noise $\sigma = 0.001$, SNR = 5.29 dB; (c) recovered result using Algorithm 1, SNR = 7.82 dB; (d) error evolution.

Acknowledgements

We thank Hui Zhang and Bo Chen for their helpful discussions regarding image restoration. The research was supported by National Science Foundation of China under grants 10601068.

References

[1] S.Q. Ma, W.T. Yin, Y. Zhang, A. Chakraborty, An efficient algorithm for compressed MR imaging using total variation and wavelets, in: IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1–8.
[2] B. Anconelli, M. Bertero, P. Boccacci, M. Carbillet, H. Lanteri, Iterative methods for the reconstruction of astronomical images with high dynamic range, J. Comput. Appl. Math. 198 (2) (2007) 321–331.
[3] J.-F. Cai, R.H. Chan, L. Shen, Z. Shen, Restoration of chopped and nodded images by framelets, SIAM J. Sci. Comput. 30 (3) (2008) 1205–1227.
[4] E.J. Candes, M. Wakin, An introduction to compressive sampling, IEEE Sig. Proc. Mag. 25 (2) (2008) 21–30.
[5] E.J. Candes, M. Wakin, S. Boyd, Enhancing sparsity by reweighted l1 minimization, J. Fourier Anal. Appl. 14 (5) (2008) 877–905.
[6] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1–4) (1992) 259–268.
[7] S. Wright, R. Nowak, M. Figueiredo, Sparse reconstruction by separable approximation, IEEE Trans. Signal Process. 57 (7) (2009) 2479–2493.
[8] S. Becker, J. Bobin, E. Candes, Nesta: a fast and accurate first order method for sparse recovery, SIAM J. Imaging Sci. (2010) (revised).
[9] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imaging Sci. 2 (1) (2009) 183–202.
[10] J. Bioucas-Dias, M. Figueiredo, A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration, IEEE Trans. Image Process. 16 (12) (2007) 2992–3004.
[11] T.F. Chan, G.H. Golub, P. Mulet, A nonlinear primal–dual method for total variation based image restoration, SIAM J. Sci. Comput. 20 (6) (1999) 1964–1977.
[12] Silvia Bonettini, Thomas Serafini, Non-negatively constrained image deblurring with an inexact interior point method, J. Comput. Appl. Math. 231 (1) (2009) 236–248.
[13] I. Daubechies, M. Defrise, C. De Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Commun. Pure Appl. Math. 57 (11) (2004) 1413–1457.
[14] P.L. Combettes, V.R. Wajs, Signal recovery by proximal forward–backward splitting, SIAM Multiscale Model. Simul. 4 (4) (2005) 1168–1200.
[15] M. Afonso, J. Bioucas-Dias, M. Figueiredo, Fast image recovery using variable splitting and constrained optimization, IEEE Trans. Image Process. 19 (9) (2010) 2345–2356.
[16] A.N. Iusem, Augmented Lagrangian methods and proximal point methods for convex optimization, Investigación Oper. 8 (1999) 11–49.
[17] M. Elad, B. Matalon, M. Zibulevsky, Coordinate and subspace optimization methods for linear least squares with non-quadratic regularization, Appl. Comput. Harmon. Anal. 23 (3) (2007) 346–367.
[18] M. Elad, M. Zibulevsky, Iterative shrinkage algorithms and their acceleration for L1-L2 signal and image processing applications, IEEE Sig. Proc. Mag. 27 (3) (2010) 78–88.

[19] T. Goldstein, S. Osher, The split Bregman method for L1 regularized problems, SIAM J. Imaging Sci. 2 (2) (2009) 323–343.
[20] A. Chambolle, An algorithm for total variation minimization and applications, J. Math. Imaging Vision 20 (1–2) (2004) 89–97.
[21] K.J. Arrow, L. Hurwicz, H. Uzawa, Studies in Linear and Non-Linear Programming, Stanford Univeristy Press, 1958.
[22] M.Q. Zhu, Fast numerical algorithms for total variation based image restoration, Ph.D. Thesis, 2008.
[23] J. Eckstein, D. Bertsekas, On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math. Program. 55 (3) (1992) 293–318.
[24] P.L. Combettes, J.-C. Pesquet, A Douglas–Rachford splitting approach to nonsmooth convex variational signal recovery, IEEE J. Sel. Top. Signal Process. 1 (2007) 564–574.
[25] S. Setzer, Operator splittings, Bregman methods and frame shrinkage in image processing, Int. J. Comput. Vis., 2010, in press (doi:10.1007/s11263-010-0357-3), published online (16 July 2010).
[26] I. Ekeland, R. Temam, Convex analysis and variational problems, SIAM Classics Appl. Math. (1999) 68–69.
[27] M. Lustig, D. Donoho, J. Pauly, Sparse MRI: the application of compressed sensing for rapid MR imaging, Magn. Reson. Med. 58 (2007) 1182–1195.