

Research Article

A Weighted Two-Level Bregman Method with Dictionary Updating for Nonconvex MR Image Reconstruction

Qiegen Liu,¹ Xi Peng,^{2,3} Jianbo Liu,^{2,3} Dingcheng Yang,¹ and Dong Liang^{2,3}

¹ Department of Electronic Information Engineering, Nanchang University, Nanchang 330031, China

² Paul C. Lauterbur Research Centre for Biomedical Imaging, Institute of Biomedical and Health Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, Guangdong 518055, China

³ Shenzhen Key Laboratory for MRI, Chinese Academy of Sciences, Shenzhen 518055, China

Correspondence should be addressed to Dong Liang; dong.liang@siat.ac.cn

Received 6 July 2014; Revised 10 September 2014; Accepted 10 September 2014; Published 30 September 2014

Academic Editor: Jun Zhao

Copyright © 2014 Qiegen Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Nonconvex optimization has shown that it needs substantially fewer measurements than l_1 minimization for exact recovery under fixed transform/overcomplete dictionary. In this work, two efficient numerical algorithms which are unified by the method named weighted two-level Bregman method with dictionary updating (WTBMDU) are proposed for solving l_p optimization under the dictionary learning model and subjecting the fidelity to the partial measurements. By incorporating the iteratively reweighted norm into the two-level Bregman iteration method with dictionary updating scheme (TBMDU), the modified alternating direction method (ADM) solves the model of pursuing the approximated l_p -norm penalty efficiently. Specifically, the algorithms converge after a relatively small number of iterations, under the formulation of iteratively reweighted l_1 and l_2 minimization. Experimental results on MR image simulations and real MR data, under a variety of sampling trajectories and acceleration factors, consistently demonstrate that the proposed method can efficiently reconstruct MR images from highly undersampled k -space data and presents advantages over the current state-of-the-art reconstruction approaches, in terms of higher PSNR and lower HFEN values.

1. Introduction

Fast acquisition is an important issue in magnetic resonance imaging (MRI) for avoiding physiological effects and reducing scanning time on patients. Acquisition time is proportional to the number of acquired k -space samples [1]. Unfortunately, accelerating acquisition by reducing k -space samples leads to noise amplification, blurred object edges, and aliasing artifacts in MR reconstructions. Then improving reconstruction accuracy from highly undersampled k -space data becomes a complementary tool to alleviate the above side effects of reducing acquisition.

To cope with the loss of image quality, some prior information should be used in the reconstruction procedure. One process of introducing prior information in the reconstruction is known as “regularization” [2–11]. Tikhonov regularization, a commonly used method that pursues reconstructions by a l_2 -norm minimization, leads to a closed-form

solution that can be numerically implemented in an efficient way [6–8, 11]. Moreover, with the advent of compressed sensing (CS) theory, sparsity-promoting regularization has gained popularity in MRI (e.g., the l_1 -based regularization) [1–3, 5, 10, 12]. The CS theory states that sparse, or more generally, compressible signals, incoherently acquired in an appropriate sense, can be recovered from a reduced set of measurements that are largely below the Nyquist sampling rate. Exact reconstruction can be achieved by nonlinear algorithms, using l_1 minimization or orthogonal matching pursuit (OMP) [1–3, 5, 10, 12, 13]. Usually, nonconvex optimization like l_p ($0 < p < 1$) minimization will guarantee a better recovery by directly attacking the l_0 minimization problem [14–16]. Chartrand et al. investigated the nonconvex l_p ($0 < p < 1$) norm optimization as a relaxation of the l_0 -norm for MRI reconstruction [17, 18] and showed that it often outperforms the l_1 -minimization method. Inspired by viewing p as a scale parameter, Trzasko and Manduca

[15] developed a homotopic l_0 -minimization strategy to reconstruct MR image and generalized the nonconvex l_p ($0 < p < 1$) norm to a family of nonconvex functions as alternatives. Candès et al. presented the iteratively reweighted l_1 minimization (IRL1) algorithm [19], which amounts to linearly minimize the log penalty of the total variation. Wong et al. [16] further incorporated a seminonlocal priority into the homotopic l_0 -minimization for breast MRI.

Besides the prior-inducing penalty function, another important issue in CS-MRI reconstruction, is the choice of a sparsifying transform, since one of the important requirements in CS-MRI is that the image to be reconstructed has a sparse representation in a given transform domain. At the start, sparse representation was usually realized by total variation (TV) regularization and/or wavelet transform [1, 12, 13, 15, 20, 21]. As it is known, TV prior assumes the target images consist of piecewise constant areas which may not be valid in many practical MRI applications. Other analytically designed bases, such as wavelets and shearlets, also involve their intrinsic deficiencies [20]. Some advanced approaches were developed to address these issues. Knoll et al. introduced the total generalized variation (TGV) for MR image reconstruction problems [22]. Liang et al. [23] applied the nonlocal total variation (NLTV) regularization to improve the signal-to-noise ratio (SNR) in parallel MR imaging, by replacing the gradient functional in conventional TV using a weighted nonlocal gradient function to reduce the blocky effect of TV regularization. Since fixed bases might not be universally optimal for all images, some methods utilizing sparsity prior under adaptive transform/dictionary were developed recently [24–29]. The sparsity in this framework is enforced on overlapping image patches to emphasize local structures. Additionally, the global dictionary consisting of local atoms (prototype) is adapted to the particular image instance, which thereby provides better sparsity. Our work also adopts the similar concept as a penalty term for MRI reconstruction.

After reviewing two important components in CS, an interesting question is whether a better sparsity-inducing function (e.g., l_p ($0 < p < 1$) norm) combined with a better transform (e.g., adaptive dictionary) will lead to better reconstruction from the same set of measurements. To the best of our knowledge, there is only one paper by Shi et al. preliminarily studying the l_0 -approximation sparsity under the learned dictionary [14], where the l_0 penalty is relaxed by the nonconvex minimax concave (MC) penalty. The MC penalty approximates the l_0 -penalty by gradually increasing a scale parameter. However, their work focused on image processing and the homotopic strategy they used is computationally demanding [14].

In this paper, we propose a novel method to solve the nonconvex minimization problem in CS-MRI by incorporating the reweighting scheme with our recently developed two-level Bregman method with dictionary updating (TBMDU). An ultimate advantage of our proposed method is that, by incorporating the iteratively reweighted scheme into the two-level Bregman method/augmented Lagrangian (AL) with dictionary updating, the l_p -based updating step is conducted in each inner iteration of the AL scheme, and both the

minimization of the sparsity of image patches and the data-fidelity constraint can be solved in a unified formalism. The modified alternating direction method (ADM) efficiently solves the model of pursuing the approximated l_p ($0 < p < 1$)-norm penalty. Specifically, by applying iteratively reweighted l_2 or l_1 minimization scheme at each AL inner iteration, the proposed algorithm utilizes adaptive weights to encourage the large coefficients to be nonzero and the small ones to be zero under learned dictionary, hence resulting in more compact and discriminative representation for reconstructed image patches. Numerical experiments show that the performance of the derived method is superior to other existing methods under a variety of sampling trajectories and k -space acceleration factors. Particularly, it achieves better reconstruction results than those by using TBMDU without adding any computational complexity.

The rest of this paper is organized as follows. We start with a brief review on TBMDU and iteratively reweighting scheme for l_p -minimization in Section 2. Consequently, the proposed method WTBMDDU is presented in Section 3. Several numerical simulation results are illustrated in Section 4 to show the superiority of the proposed method. Finally, conclusions and discussions are given in Section 5.

2. Theory

Plenty of papers, with the theme of compressed sensing (CS) [2, 3], demonstrated that MR images can be reconstructed from very few linear measurements [1–3, 5, 10, 12]. These approaches take advantage of the sparsity inherent in MR images. The well-known example is the success of TV regularization which implies that MR images can be approximated by those having sparse gradients. Generally, if a sparsifying transform ψ can be readily defined, ideally one could choose the estimation with the sparsest representation in ψ that still matches the limited measurements. That is, by minimizing transform-domain sparsity-promoting energy $J(u) = \|\psi u\|_0$ that is subject to the data consistency $f = F_p u + n$, where n is noise, it yields

$$\begin{aligned} \min_u \quad & J(u) \\ \text{s.t.} \quad & \|F_p u - f\|_2 \leq \varepsilon, \end{aligned} \tag{1}$$

where $u \in \mathbb{C}^N$ is a vector representing the N -pixel 2D complex image. $F_p \in \mathbb{C}^{Q \times N}$ denotes the partially sampled Fourier encoding matrix and $f \in \mathbb{C}^Q$ represents the acquired data in k -space. $\varepsilon = \sqrt{2}\sigma$ is the standard deviation of the zero-mean complex Gaussian noise, and σ is the standard deviation of both real and imaginary parts of the noise. Practically, the l_0 quasi-norm $\|\cdot\|_0$ in the regularization term of (1) is usually relaxed. In this paper, we propose some efficient algorithms for nonconvex relaxation such as l_p quasi-norm $\|\cdot\|_p$, $0 < p < 1$ [17, 18].

The choice of the sparsifying transform ψ is an important consideration in minimizing functional (1). Besides the fixed transforms such as finite-difference and wavelet transform, modeling image patches (signals) by sparse and redundant

representations have been drawing considerable attention in recent years [24]. Considering the image patches set $Ru = [R_1u, R_2u, \dots, R_Lu]$ consisting of L signals, $R_lu \in \mathbb{C}^M$ denotes a vectored form of the $\sqrt{M} \times \sqrt{M}$ patch extracted from the image u of size $\sqrt{N} \times \sqrt{N}$. The sparseland model for image patches suggests that each image patch, R_lu , could be sparsely represented over a learned dictionary D^* [27]; that is,

$$\alpha_l^* = \arg \min_{\alpha_l} \left(\|\alpha_l\|_0 + \frac{\lambda}{2} \|D^* \alpha_l - R_lu\|_2^2 \right), \quad l = 1, 2, \dots, L, \quad (2)$$

where α_l stands for the sparse coefficient of the l th patch. The combination of sparse and redundant representation modeling of signals, together with a learned dictionary using signal examples, has shown its promise in a series of image processing applications [24–29]. For example, in [25, 26], the authors used reference MRI image slices to train the dictionary and achieved limited improvement. Considering that a dictionary learnt from a reference image would not be able to effectively sparsify new features in the current scan; Ravishankar and Bresler [29] suggested learning dictionary from the target image itself and developed a two-step alternative method to remove aliasing artifacts and noise in one step and subsequently fills in the k -space data in the other step.

In the following, to make the paper self-contained, we, respectively, review the TBMDU method for MRI reconstruction and reweighted scheme for solving nonconvex relaxation of the l_p -quasi-norm minimization.

2.1. The TBMDU Framework for MRI Reconstruction. Recently, we proposed a series of dictionary learning methods [30–32] based on augmented Lagrangian/Bregman iterative scheme [33]. Particularly, by using a two-level Bregman iterative method and the relaxed version of sparseland model (i.e., $J(u) = \sum_l (\|\alpha_l\|_1 + (\lambda/2) \|D\alpha_l - R_lu\|_2^2)$) as the regularization term for MRI reconstruction [31], the TBMDU method solves the objective function equation (1) by iterating the following:

$$u^{k+1} = \arg \min_u \left\{ \min_{D, \Gamma} \sum_l \left(\|\alpha_l\|_1 + \frac{\lambda}{2} \|D\alpha_l - R_lu\|_2^2 \right) + \frac{\mu}{2} \|F_p u - f^k\|_2^2 \right\}, \quad (3)$$

$$f^{k+1} = f^k + f - F_p u^{k+1},$$

where $D = [d_1, d_2, \dots, d_J] \in \mathbb{C}^{M \times J}$ and $\Gamma = [\alpha_1, \alpha_2, \dots, \alpha_L] \in \mathbb{C}^{J \times L}$. λ stands for the sparse level of the image patches in the “optimal” dictionary. For many natural or medical images, the value of λ can be determined empirically with robust performance in our work. $J = K \cdot M$, in which K measures the degree of the overcompleteness of the dictionary.

The proposed TBMDU method consists of a two-level Bregman iterative procedure, where (3) is the outer-level of the method. On the other hand, the inner-level Bregman iterative is employed to solve the subproblem of sparse

representation in (3), that is, adding auxiliary variables z_l to convert the unconstrained problem (2) into a constrained formulation:

$$\begin{aligned} \min_{D, \alpha_l} \quad & \|\alpha_l\|_1 + \frac{\lambda}{2} \|z_l\|_2^2 \\ \text{s.t.} \quad & z_l = D\alpha_l - R_lu, \quad l = 1, 2, \dots, L. \end{aligned} \quad (4)$$

Then split Bregman method/augmented Lagrangian is used to solve problem (4) as follows:

$$\begin{aligned} & (D^{i+1}, \alpha_l^{i+1}, z_l^{i+1}) \\ & = \arg \min_{D, \alpha_l, z_l} \sum_l \|\alpha_l\|_1 + \frac{\lambda}{2} \|z_l\|_2^2 + \frac{\beta}{2} \left\| z_l + D\alpha_l - R_lu - \frac{y_l^i}{\beta} \right\|_2^2, \end{aligned} \quad (5)$$

$$Y^{i+1} = Y^i + \beta (Ru - Z^{i+1} - D^{i+1} \Gamma^{i+1}). \quad (6)$$

We minimize the corresponding AL function (5) alternatively with respect to one variable at a time. This scheme decouples the minimization process and simplifies the optimization task. At the dictionary updating step, by updating (5) with respect to D along the gradient descent direction and constraining the norm of each atom to be unit [24], we can get the following update rule:

$$\begin{aligned} D^{i+1} &= D^i + \xi Y^{i+1} (\Gamma^{i+1})^T, \\ d_j^{i+1} &= \frac{d_j^{i+1}}{\|d_j^{i+1}\|_2}, \quad j = 1, 2, \dots, J. \end{aligned} \quad (7)$$

At the sparse coding stage of updating α_l , the optimization problem for each α_l is derived:

$$\alpha_l^{i,m+1} = \arg \min_{\alpha_l} \left\| D^i \alpha_l - b_l - \frac{y_l^i}{\beta} \right\|_2^2 + 2 \frac{\lambda + \beta}{\lambda \beta} \|\alpha_l\|_1. \quad (8)$$

By applying the iterative shrinkage/thresholding algorithm (ISTA) [33, 34] (i.e., performing a gradient descent of the first functional term in (8) and then solving a proximal operator problem) and using the immediate variable y_l^{m+1} to represent the updating scheme compactly (i.e., the formulation $y_l^{m+1} = (\lambda\beta/\lambda + \beta)(-D^i \alpha_l^{i,m} + b_l + y_l^i/\beta)$ derived from (6)), it attains the solution of α_l as follows:

$$\alpha_l^{i,m+1} = \text{Shrink} \left(\alpha_l^{i,m} + (\lambda + \beta) (D^i)^T \frac{y_l^{m+1}}{\gamma \lambda \beta}, \frac{(\lambda + \beta)}{\gamma \lambda \beta} \right), \quad (9)$$

where $\gamma > \text{eig}((D^i)^T D^i)$ and $x = \text{Shrink}(g, \mu) = (g \cdot \max(|g| - \mu, 0)) / (\max(|g| - \mu, 0) + \mu)$ is the solution of $x = \arg \min_x \|x\|_1 + (1/2\mu)\|x - g\|_2^2$.

```

(1) Initialization:  $\Gamma^0 = 0, C^0 = 0, D^0, u^0 = F_p^T f, f^0 = f, T_0 = \lambda\beta / (\lambda + \beta)$ 
(2) While For  $\|F_p u^k - f\|_2 > \sigma$  do
(3)    $i = 1$ 
(4)   While  $i \leq P$  do
(5)     While stop criterion not satisfied (loop in  $m$ ) do
(6)        $Y^{m+1} = T_0 (-D^i \Gamma^{i,m} + R u^{k,i} + C^i / \beta)$ 
(7)        $\Gamma^{i,m+1} = \text{Shrink}(\Gamma^{i,m} + (D^i)^T Y^{m+1} / \gamma T, 1 / \gamma T_0)$ 
(8)     End (While)
(9)      $C^{i+1} = Y^{m+1}, \Gamma^{i+1,0} = \Gamma^{i,m+1}, D^{i+1} = D^i + \xi C^{i+1} (\Gamma^{i+1,0})^T, d_j^{i+1} = d_j^{i+1} / \|d_j^{i+1}\|_2, j = 1, 2, \dots, J$ 
(10)    update  $u^{k+1}$  by frequency interpolation using
         $S_1 = FF_p^T f^k, S_2 = F \sum_l R_l^T (D^{k+1} \alpha_l^{k+1} + z_l^{k+1} - y_l^{k+1} / \beta) / \omega$ 
        and  $Fu(k_x, k_y) = \begin{cases} S_2(k_x, k_y), & (k_x, k_y) \notin \Omega \\ [\mu S_1(k_x, k_y) + \beta S_2(k_x, k_y)] / (\mu + \beta), & (k_x, k_y) \in \Omega \end{cases}$ 
(11)     $i \leftarrow i + 1$ 
(12)  End (While)
(13)   $f^{k+1} = f^k + f - F_p u^{k+1}$ 
(14) End (While)

```

ALGORITHM 1: TBMDU.

The corresponding procedure of TBMDU is summarized in Algorithm 1. Line 10 is the frequency interpolation step for updating the image u . AL scheme and ADM applied to the constrained problem result in an efficient two-step iterative mechanism, which alternatively updates the solution u and image-patch related coefficients (Γ , Y , and D). For a more detailed description of the derivations, the interested readers can refer to [31].

2.2. Reviews of Reweighted Scheme on Solving Nonconvex Relaxation of the l_p Quasi-Norm. In (1), the l_0 penalty can be relaxed into several tractable alternatives. For example, some nonconvex relaxations such as l_p quasi-norm $\|\cdot\|_p, 0 < p < 1$ [17, 18], log penalty [19], and smoothly clipped absolute deviation (SCAD) penalty [35] were usually used. Compared to l_1 norm, l_p function with $0 < p < 1$ is a smooth relaxation of l_0 and the geometry of l_p ball gives better approximation on sparsity. Recent numerical results showed that adapting nonconvex optimization technique can reduce the required number of measurements for reconstruction [14–16].

Most current approaches for solving nonconvex l_p -quasi-norm optimization can be classified into two categories: reweighted l_2 and reweighted l_1 [36–41]. The common idea of these methods is to utilize an optimization transfer technique to iteratively surrogate the nonconvex potential function by a convex function [37, 39, 41]. Usually, the convex function is chosen as a local quadratic approximation or linear local approximation. A technique drawing much interest is the iteratively reweighted least squares (IRLS) method which is widely used in compressed sensing and image processing [36, 40]. Another technique closely related to IRLS is the iteratively reweighted norm (IRN) approach introduced by Rodríguez and Wohlberg for generalized total variation functional [42]. The IRN method pursues to minimize the l_p

norm $F(\alpha) = (1/p)\|\alpha\|_p^p$ for $p \leq 2$ by using a weighted l_2 norm in an iterative manner. At iteration $k + 1$, the solution α^{k+1} is the minimizer of

$$G(\alpha) = \tilde{F}(\alpha, \alpha^k) = \frac{1}{2} \|\sqrt{W^k} \alpha\|^2, \quad \sqrt{W^k} = \text{diag}(|\alpha^k|^{p-2}). \quad (10)$$

The sequence of solution $\{\alpha^k\}$ converges to the minimizer of $F(\alpha)$ as $k \rightarrow +\infty$. It has been proven that IRN is one kind of majorization-minimization (MM) method [42], which involves good property of $G(\alpha^{k+1}) \leq G(\alpha^k)$. In order to extend the iterative reweighting approach IRLS to the general nonconvex function, Mourad and Reilly developed the quadratic local approximation of the nonconvex function as follows [39]:

$$F(\alpha_{j,l}) = \varphi(|\alpha_{j,l}|) = \phi(|\alpha_{j,l}^k|) + \phi'(|\alpha_{j,l}^k|)(|\alpha_{j,l}| - |\alpha_{j,l}^k|) + \zeta_0(|\alpha_{j,l}| - |\alpha_{j,l}^k|)^2, \quad (11)$$

where $\zeta_0 \geq 0$ is a suitably chosen parameter and ϕ' is the first derivative of ϕ . j stands for the j th index number of vector α_j .

A similar work conducted by Elad and Aharon is the iteratively reweighted l_1 minimization (IRL1) algorithm [27], which iteratively minimizes the linearization of a quasi-logarithmic potential function. Zou and Li [41] presented the linear local approximation of the general nonconvex function and the upper bound of

$$F(\alpha_{j,l}) = \varphi(|\alpha_{j,l}|) = \phi(|\alpha_{j,l}^k|) + \phi'(|\alpha_{j,l}^k|)(|\alpha_{j,l}| - |\alpha_{j,l}^k|). \quad (12)$$

Equation (12) aims to approximate the penalty function with its first-order Taylor expansion at the current iteration. Its

majorization properties can be easily obtained by exploiting the concavity of the function ϕ , which always lies below its tangent. The numerical tests in [17] by Chartrand and Yin demonstrated that IRLS is comparable with IRL1. In [38], the authors presented weighted nonlinear filter for compressed sensing (WNFCS) for general nonconvex minimization by integrating the IRL1 algorithm shown in (12) and their previously developed NFCS framework [43].

3. Materials and Methods

As discussed in previous section, dictionary updating and sparse coding to (5) are performed sequentially. Therefore, an interesting question is that will the better sparsity inducing function in the coefficient matrix lead to better image reconstruction under the learned dictionary? In the following, we investigate two variations at the sparse coding step. By employing the l_p -seminorms ($0 < p \leq 1$), it is feasible to expect that utilizing approximations that are closer to the l_0 -quasi-norm than l_1 will correspondingly reduce the required number of measurements for accurate reconstruction.

3.1. Proposed Method: WTBMDU. By replacing l_1 norm with nonconvex l_p norm, (3) can be rewritten as

$$\begin{aligned} u^{k+1} = \arg \min_u \left\{ \min_{D, \Gamma} \sum_l \left(\frac{1}{p} \|\alpha_l\|_p + \frac{\lambda}{2} \|D\alpha_l - R_l u\|_2^2 \right) \right. \\ \left. + \frac{\mu}{2} \|F_p u - f^k\|_2^2 \right\}. \\ f^{k+1} = f^k + f - F_p u^{k+1}. \end{aligned} \quad (13)$$

After some manipulations similar to those in Section 2.1, updating coefficients at the sparse coding stage are deduced as follows:

$$\alpha_l^{i,m+1} = \arg \min_{\alpha_l} \left\| D^i \alpha_l - b_l - \frac{y_l^i}{\beta} \right\|_2^2 + 2 \frac{\lambda + \beta}{\lambda \beta} \frac{1}{p} \|\alpha_l\|_p. \quad (14)$$

When other variables are fixed, numerical strategies are proposed in the following to solve the corresponding nonconvex minimization problems with regard to α , by iteratively minimizing a convex majorization of the nonconvex objective function.

3.1.1. Algorithm 1: Reweighed l_1 . By inserting the local linear approximation in the sparse coding step, it is possible to transform the nonconvex unconstrained minimization problem into a convex unconstrained problem, which can be solved iteratively. In the penalty functional of (14), like the well-known method ISTA, a proximal operator is

employed to approximate the first term, and the reweighted l_1 -approximation is employed to the second term:

$$\alpha_l^{i,m+1} = \arg \min_{\alpha_l} \underbrace{\left\| D^i \alpha_l - b_l - \frac{y_l^i}{\beta} \right\|_2^2}_{\text{ISTA}} + 2 \frac{\lambda + \beta}{\lambda \beta} \frac{1}{p} \underbrace{\sum_j |\alpha_{j,l}|^p}_{\text{Re } L1}, \quad (15)$$

where j stands for the j th index number of vector α_l . Following the definition in (12), where $\phi(|\alpha_{j,l}|) = (1/p)|\alpha_{j,l}|^p$ and $\phi'(|\alpha_{j,l}|) = 1/(|\alpha_{j,l}|^{1-p} + \varepsilon)$, it attains the solution of α_l :

$$\begin{aligned} \alpha_l^{i,m+1} = \arg \min_{\alpha_l} \gamma \left\| \alpha_l - \left[\alpha_l^{i,m} + \frac{(\lambda + \beta)(D^i)^T y_l^{m+1}}{\gamma \lambda \beta} \right] \right\|_2^2 \\ + 2 \frac{\lambda + \beta}{\lambda \beta} \|W_l^{i,m} \alpha_l\|_1 \\ = \text{Shrink} \left(\alpha_l^{i,m} + \frac{(\lambda + \beta)(D^i)^T y_l^{m+1}}{\gamma \lambda \beta}, \frac{(\lambda + \beta) W_l^{i,m}}{\gamma \lambda \beta} \right), \end{aligned} \quad (16)$$

where $W_l^{i,m} = 1/(|\alpha_{j,l}^{i,m}|^{1-p} + \varepsilon)$, $j = 1, 2, \dots, J$. $\varepsilon > 0$ is a small parameter to prevent division by zeros.

3.1.2. Algorithm 2: Reweighed l_2 . Compared to the previous reweighted l_1 strategy, the only difference is that the reweighted l_2 -approximation in (10) is employed to the second term:

$$\alpha_l^{i,m+1} = \arg \min_{\alpha_l} \underbrace{\left\| D^i \alpha_l - b_l - \frac{y_l^i}{\beta} \right\|_2^2}_{\text{ISTA}} + 2 \frac{\lambda + \beta}{\lambda \beta} \frac{1}{p} \underbrace{\|\alpha_l\|_p^p}_{\text{Re } L2}. \quad (17)$$

Then it attains the solution of α_l :

$$\begin{aligned} \alpha_l^{i,m+1} = \arg \min_{\alpha_l} \gamma \left\| \alpha_l - \left[\alpha_l^{i,m} + \frac{(\lambda + \beta)(D^i)^T y_l^{m+1}}{\gamma \lambda \beta} \right] \right\|_2^2 \\ + \frac{\lambda + \beta}{\lambda \beta} \left\| \sqrt{W_l^{i,m}} \alpha_l \right\|_2^2. \end{aligned} \quad (18)$$

The minimizer of this least-square problem is given by

$$\alpha_l^{i,m+1} = \frac{\gamma \alpha_l^{i,m} + (\lambda + \beta)(D^i)^T y_l^{m+1} / \lambda \beta}{\gamma + (\lambda + \beta) W_l^{i,m} / \lambda \beta}, \quad (19)$$

where $W_l^{i,m} = 1/(|\alpha_{j,l}^{i,m}|^{2-p} + \varepsilon)$, $j = 1, 2, \dots, J$. Similar to that in (14), $\varepsilon > 0$ is a small parameter to prevent numerical instabilities.

Now, we summarize our proposed method for MRI reconstruction here, which we call *weighted TBMDU*. The detailed description of the proposed method is listed in

Algorithm 2. Similar to TBMDU, the proposed *WTBMDU* method alternatively updates the target solution u , image patch related coefficients (Γ , Y , and D), and auxiliary variables. The difference between the plain TBMDU and the weighted TBMDU mainly lies on the weights used in updating the sparse coefficients. In *WTBMDU*, the weights are obtained by evaluating the function of variables at the solution of the previous step. Specifically, whether the $W_{j,l}^{i,m} = 1/(|\alpha_{j,l}^{i,m}|^{1-p} + \varepsilon)$ in (16) or $W_{j,l}^{i,m} = 1/(|\alpha_{j,l}^{i,m}|^{2-p} + \varepsilon)$ in (19), the weighting scheme $W_{j,l}^{i,m}$ decreases as the absolute value of $\alpha_{j,l}^{i,m}$ increases, indicating that it penalizes more on the coefficients with small magnitude value. Therefore, this operation strongly encourages the large coefficients to be nonzero and the small ones to be zero. In line 8 for reweighted l_1 and line 10 for reweighted l_2 , the formulations denote operating every element in the matrix component-wise. In the following content of this paper, we denote the reweighted l_1 and reweighted l_2 in *WTBMDU* as *WTBMDU-L1* and *WTBMDU-L2*, respectively.

The strategy we used in *WTBMDU* is similar to that used in *WNFCS* [38], that is, combining the penalized proximal splitting strategy and reweighting strategy. The difference is that *WNFCS* focus on updating u in the simple CS domain and the proximal splitting strategy is employed to the Fourier-related data-fidelity term, while our *WTBMDU* devotes to updating the coefficients in the adaptive dictionary and the proximal splitting strategy is employed to the dictionary-related sparse representation error term. Additionally, we also derive the updating scheme based on reweighted l_2 .

3.2. Parameter Values, Continuation Strategy, and Algorithm Convergence. The proposed method involves four parameters: β , μ , λ , and ζ . The setting of these parameters is similar to that in TBMDU. Firstly, both parameters β and μ are the Bregman (or augmented Lagrangian) positive parameters associated with the small image patches and the whole image itself, respectively. One is for the overlapping image patches and the other is for the image solution itself. It has been mathematically proven that the choice of the Bregman parameter has little effect on the final reconstruction quality as long as it is sufficiently small [31, 44, 45]. In our work, the smaller the value of the Bregman parameter is, the more iterations the Bregman method need to reach the stopping condition. Moreover, since β and μ are with different orders of magnitude, we set $\mu \approx \rho N \beta / M$ in the AL formalism for the balance between various “penalty” terms. Secondly, λ stands for the sparse level of the image patches and can be determined empirically. Finally, the step size ζ in the dictionary updating stage can be set to be a small positive number, for example, 0.01.

In summary, there is only the parameter β that should be carefully chosen in order to enable the efficiency of the algorithm. In our method, by taking advantage of the typical structure of the problem in this paper, we propose a similar rule to that used in [21]. Specifically, we set β and μ so as to achieve condition numbers $\kappa(1 + W_l^{k,m}/\gamma T_0)$ and $\kappa(FF_p^T F_p F^T + (\beta/\mu)F \sum_l R_l^T R_l F^T)$ that result in fast convergence of the algorithm.

Since $T_0 = \lambda\beta/(\lambda + \beta)$ and $\lambda \gg \beta$, $T_0 \approx \beta$. Consequently $\kappa(1 + W_l^{k,m}/\gamma T_0)$ is a decreasing function of β . Choosing β such that $\kappa(1 + W_l^{k,m}/\gamma T_0) \rightarrow 1$ would require a large β and accordingly, the influence of the weight $W_l^{k,m}$ diminishes and the solution of $\alpha_l^{k,m+1}$ trends to the least-square solution ($q = 2$). In our experiments, we observed that this phenomenon would result in an “immature” dictionary that is not learned enough. On the other hand, taking $\beta \rightarrow 0$ would increase $\kappa(1 + W_l^{k,m}/\gamma T_0)$ which makes $(1 + W_l^{k,m}/\gamma T_0)^{-1}$ numerically unstable. The same trend also applies to $\kappa(FF_p^T F_p F^T + (\beta/\mu)F \sum_l R_l^T R_l F^T)$ as a function of β/μ . We found that empirically choosing β such that $E(\kappa(W_l^{k,m})) \approx \sum_l \kappa(W_l^{k,m})/L \in [2, 6]$ generally provided good convergence speed in all experiments. Additionally, we can estimate β in a more practical way. Specifically, since the number of data samples is huge, we can set $E(W_l^{k,m}) \approx \gamma T_0$, where $E(W_l^{k,m}) \in [0.2, 1]$ is a robust estimate. Together with the fact that $\gamma \approx 100$, it leads to $\beta \in [1/500, 1/100]$. As for the parameter μ , since $\kappa(FF_p^T F_p F^T + (\beta/\mu)F \sum_l R_l^T R_l F^T) = (\mu + \beta\omega)/\beta\omega$, where $\omega = 64$ when the patch overlapping is $r = 1$, it concludes that setting $\mu \approx \rho N \beta / M$ such that $\kappa(FF_p^T F_p F^T + (\beta/\mu)F \sum_l R_l^T R_l F^T) = 64\rho + 1 \in [4, 20]$ is a reasonable choice, under the assumption that the k -space data is significantly undersampled (e.g., $\rho \in [5\%, 30\%]$).

As for the convergence, because of the unconvexity and nonlinearity of the problem in the case of updating dictionary, the global solution may not to be found easily like in TBMDU. Nevertheless, our dictionary is updated by a gradient descent of the AL scheme which leads to a monotonic decrease in the cost function. At the sparse coding stage, since *WTBMDU* merges the reweighted and penalization strategies, it still maintains the descent property of the latter and local minimum must be attained as demonstrated in [38]. Therefore, both the value of the objective function and the norm of the reconstruction difference between successive iterations can be chosen as the stopping criterion. The convergence property of the algorithm will be presented in the numerical section.

4. Experiment Results

In this section, we evaluate the performance of the proposed method using a variety of sampling schemes, with different undersampling factors. Sampling schemes used in our experiments include 2D random sampling [15], Cartesian approximation of multishot variable-density spiral sampling [15], Cartesian sampling with random phase encodings (1D random) [1, 15], and pseudo radial sampling [15, 29]. Reconstruction results on simulated MRI data, a complex phantom, and real MRI data were presented. The MR images tested in the synthetic experiments are from in vivo MR scans of size 512×512 (many of which are courtesy (2009, American Radiology Services [Online]. Available: <http://www3.americanradiology.com/pls/web1/wwimggal.vmg/>) and used in [29]), and the real MRI data examples

```

(1) Initialization:  $\Gamma^0 = 0, C^0 = 0, D^0, u^0 = F_p^T f, f^0 = f, T_0 = \lambda\beta / (\lambda + \beta)$ 
(2) While For  $\|F_p u^k - f\|_2 > \sigma$  do
(3)    $i = 1$ 
(4)   While  $i \leq P$  do
(5)     While stop criterion not satisfied (loop in  $m$ ) do
(6)        $Y^{m+1} = T_0 (D^i + \Gamma^{i,m} + Ru^{k,i} + C^i / \beta)$ 
(7)       If Reweighted  $l_1$ 
(8)          $\Gamma^{i,m+1} = \text{Shrink}(\Gamma^{i,m} + (D^i)^T Y^{m+1} / \gamma T_0, W^{i,m} / \gamma T_0)$ 
(9)       ElseIf Reweighted  $l_2$ 
(10)         $\Gamma^{i,m+1} = \frac{\gamma \Gamma^{i,m} + (D^i)^T Y^{m+1} / T_0}{\gamma + W^{i,m} / T_0}$ 
(11)      End (If)
(12)    End (While)
(13)     $C^{i+1} = Y^{m+1}, \Gamma^{i+1,0} = \Gamma^{i,m+1}, D^{i+1} = D^i + \xi C^{i+1} (\Gamma^{i+1,0})^T, d_q^{i+1} = d_q^{i+1} / \|d_q^{i+1}\|_2, q = 1, 2, \dots, Q$ 
(14)    update  $u^{k+1}$  by frequency interpolation using  $S_1 = FF_p^T f^k$ ,
       $S_2 = F \sum_l R_l^T (D^{k+1} \alpha_l^{k+1} + z_l^{k+1} - y_l^{k+1} / \beta) / \omega$ 
      and  $Fu(k_x, k_y) = \begin{cases} S_2(k_x, k_y), & (k_x, k_y) \notin \Omega \\ [\mu S_1(k_x, k_y) + \beta S_2(k_x, k_y)] / (\mu + \beta), & (k_x, k_y) \in \Omega \end{cases}$ 
(15)     $i \leftarrow i + 1$ 
(16)  End (While)
(17)   $f^{k+1} = f^k + f - F_p u^{k+1}$ 
(18) End (While)

```

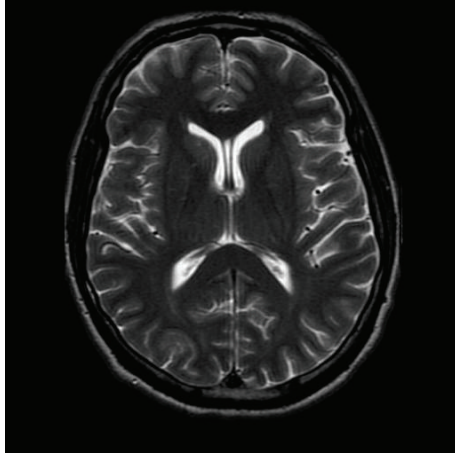
ALGORITHM 2: WTBMDU.

reported here are of size 256×256 (except in Figures 6 and 7 where a phantom of size 512×512 was used). According to many prior work on CS-MRI [1, 13, 29], the CS data acquisition was simulated by subsampling the 2D discrete Fourier transform of the MR images (except in the second subsection where real acquired data was used). Our proposed method WTBMDU was compared with the leading DLMRI (the code is available in <https://netfiles.uiuc.edu/ravisha3/www/DLMRICODE.zip>) [29] and TBMDU methods [31], which have been shown to substantially outperform other CS-MRI methods such as LDP (Matlab codes are available in <http://www.stanford.edu/~mlustig/>) [1], and the zero-filling reconstruction. DLMRI directly solves the l_0 -minimization by OMP while TBMDU is devoted to the l_1 -induced sparse minimization. In each given example, the parameters for the DLMRI method were set to be default values.

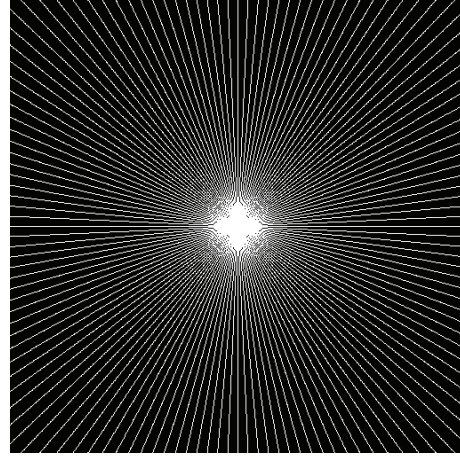
In the experiments, the nominal values of various parameters were set as patch size $\sqrt{M} = 6$, the over-completeness of the dictionary $K = 1$ (correspondingly $J = 36$), and the patch overlap $r = 1$; thereby the number of data samples were $L = 65536$ for $\sqrt{N} = 256$ and $L = 262144$ for $\sqrt{N} = 512$. $\omega = (\sqrt{M}/r)^2 = 36$, $\beta = 0.0056$, $\mu = N\beta/M$, $\zeta = 0.01$, $\lambda = 12$, and $m = 3$. To avoid dividing by zero, the parameter ε used in the weighted matrix $W_{q,l}^{i,m}$ decreased by 2% after each inner iteration with an initial value of 5. The setting of parameters m and λ is very similar to that in, [31, 32] and not discussed here due to the limit of paper space. Real-valued dictionaries were used for the simulated experiments with real-valued images, where the over-complete discrete

cosine transform (DCT) was chosen as the initial dictionary. Complex-valued dictionaries were used for real MR data, where both the real and imaginary parts were the same DCT matrix. The quality of the reconstruction was quantified using the peak signal-to-noise ratio (PSNR (the PSNR is defined as $\text{PSNR} = 20 \log_{10} 255/\text{RMSE}$, where the RMSE is the root mean error estimated between the ground truth and the reconstructed image)) and high-frequency error norm (HFEN) [29]. All algorithms were implemented in MATLAB 7.1 on a PC equipped with AMD 2.31 GHz CPU and 3 GByte RAM.

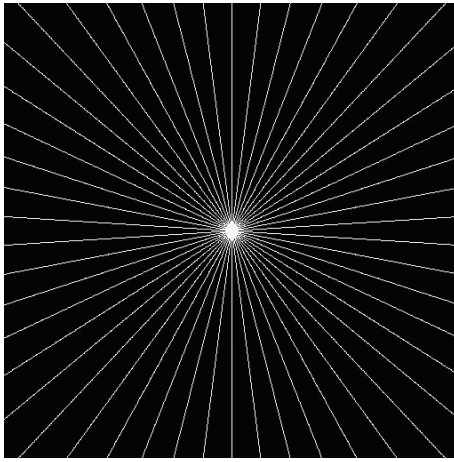
4.1. Reconstruction of Simulated Data (Real-Valued). We first investigate the performance of WTBMDU with the noiseless measurements. Figure 1 involves an axial T2-weighted reference image of the brain with pseudo radial sampling under 85% and 95% undersampling percentages, respectively (i.e., only acquiring 15% and 5% k -space data with corresponding acceleration factors of 6.67 and 20). The plots of PSNR and HFEN values as functions of iteration number under the undersampling percentage of 85% are presented in Figures 1(d) and 1(e). It can be observed that the quantitative measures of both TBMDU and WTBMDU change quickly during the first few iterations. In other words, these measure values only need less iterations to reach the convergence zone and hence the iterative convergence property of our method is better than that of DLMRI. The higher PSNR values and lower HFEN values after convergence also confirm the superiority of our method to DLMRI and TBMDU. The reconstructed results shown in Figures 1(f), 1(g), 1(h), and 1(i)



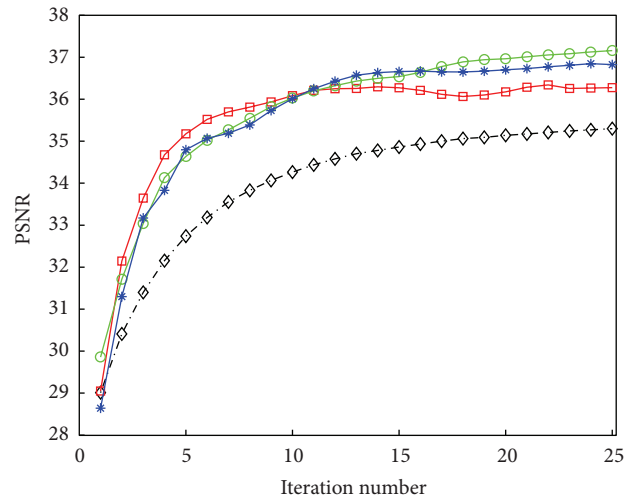
(a)



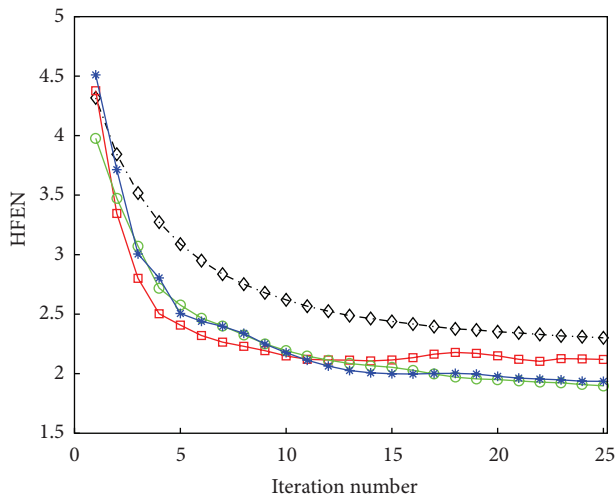
(b)



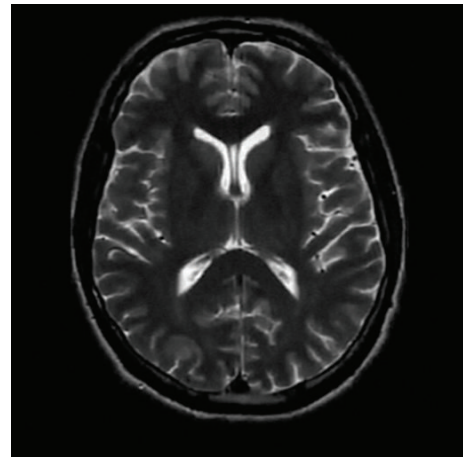
(c)



(d)



(e)



(f)

FIGURE 1: Continued.

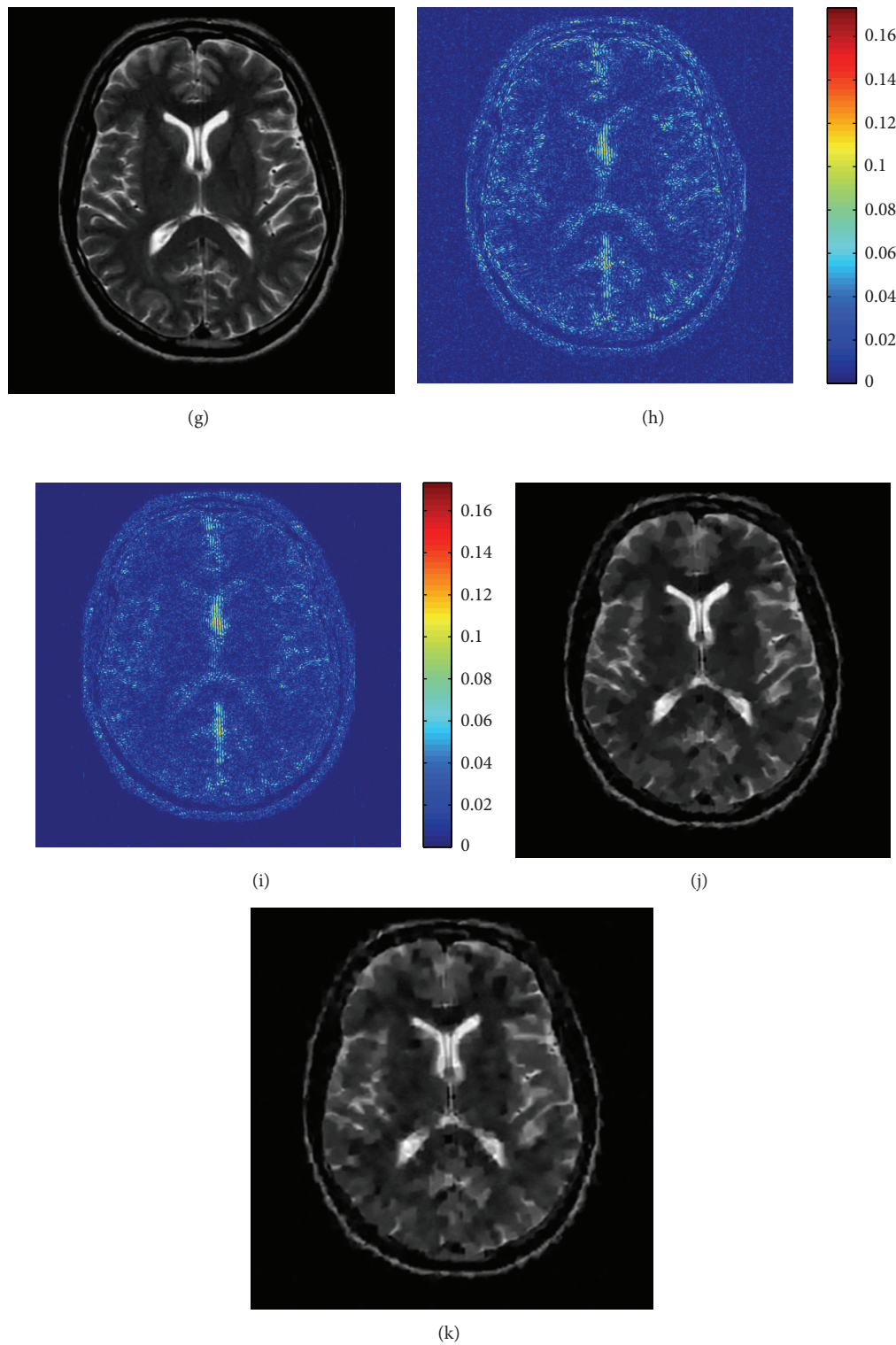


FIGURE 1: Performance of various algorithms at different undersampling ratios. ((b) and (c)) Pseudo radial sampling with 85% and 95% undersampling. ((d) and (e)) PSNR and HFEN versus the number of iterations of DLMRI, TBMDU, and WTBM DU. ((f), (g), (h), and (i)) The reconstruction results and the corresponding error magnitudes of DLMRI and WTBM DU-L1 with $p = 0.5$ at 85% undersampling. ((j) and (h)) The reconstruction results of DLMRI and WTBM DU-L1 with $p = 0.5$ at 95% undersampling.

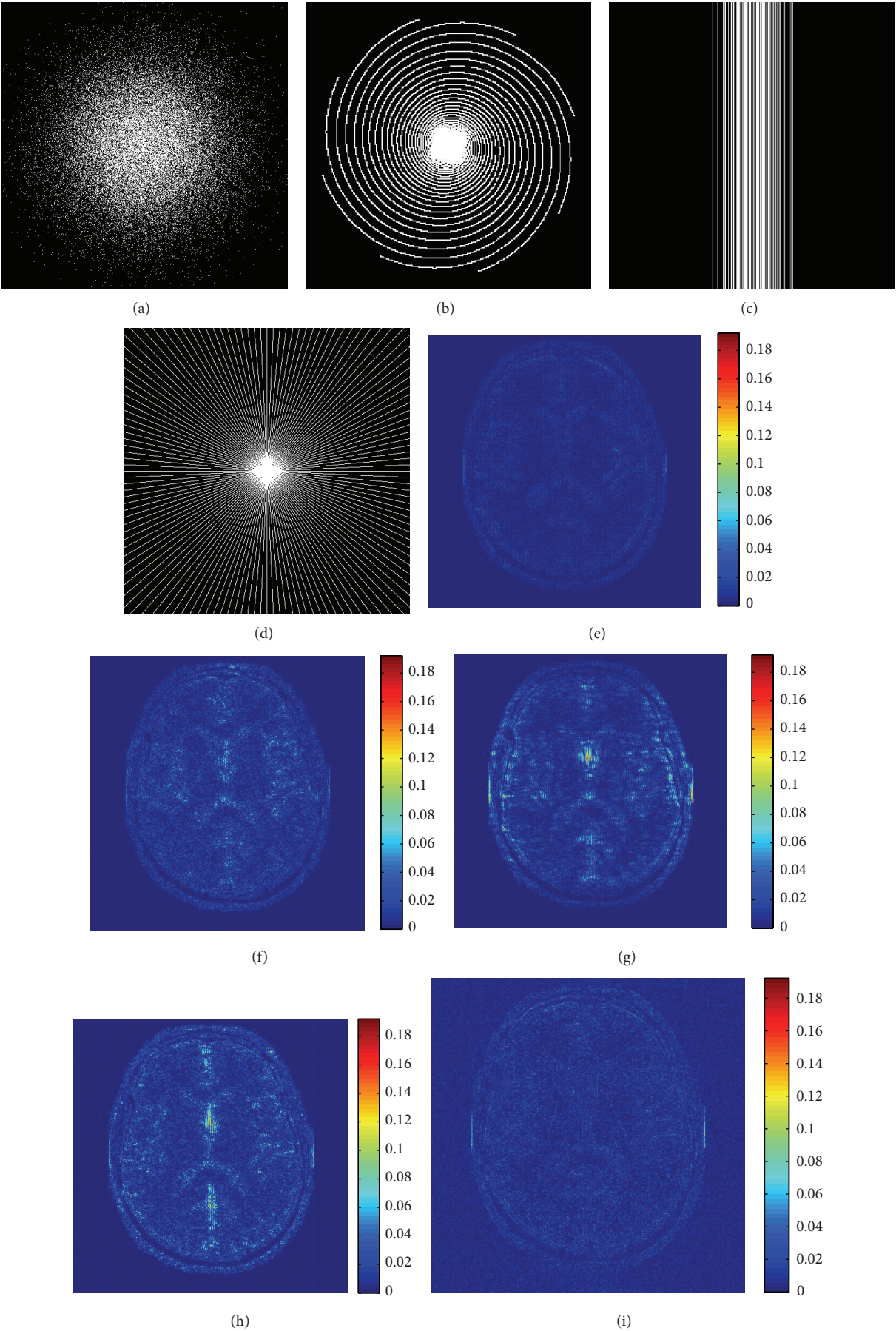


FIGURE 2: Continued.

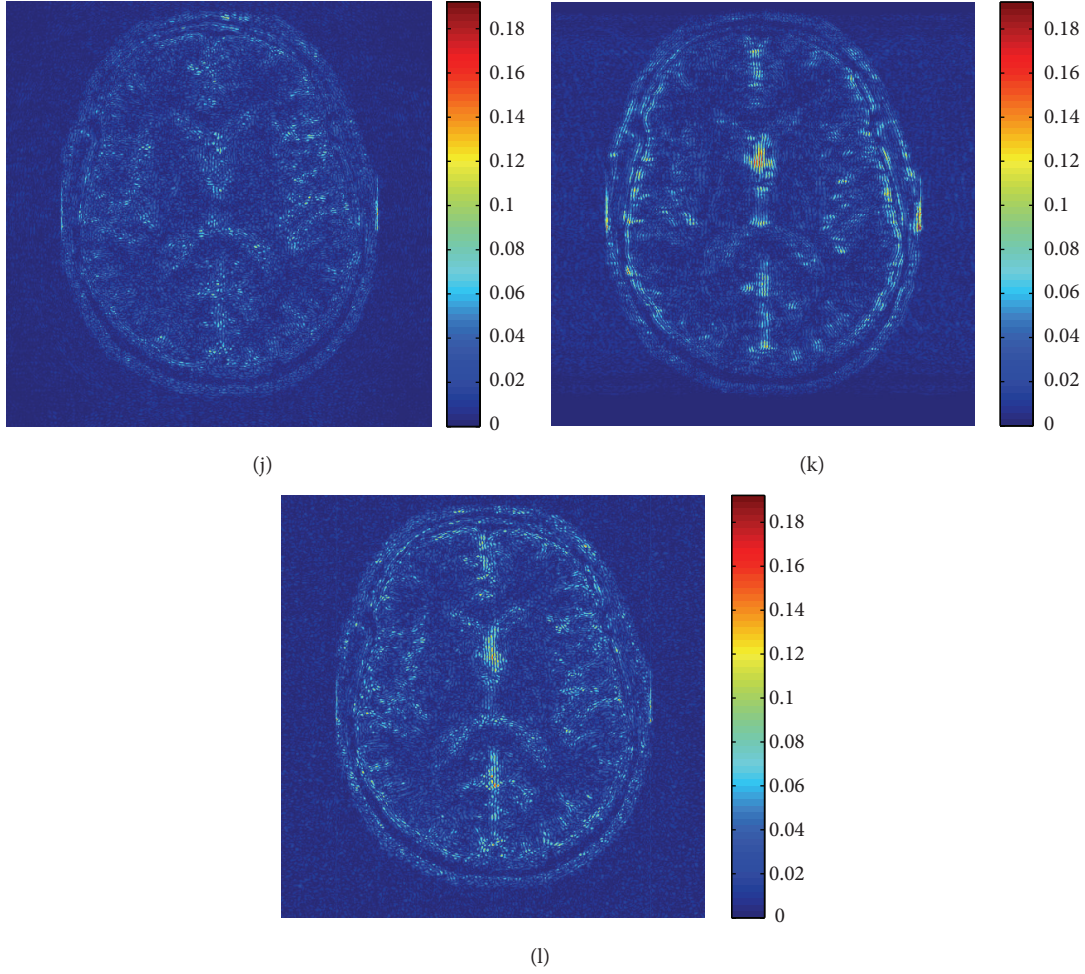


FIGURE 2: Performance of algorithms at different sampling trajectories with the same undersampling ratio. The first line ((a), (b), (c), and (d)): the variable density 2D random sampling, Cartesian approximation of multishot spiral sampling, Cartesian sampling, and pseudo radial sampling with 86% acceleration factor, respectively. The second and third lines are the reconstruction errors of DLMRI and WTBMDU-L1 with $p = 0.7$ for the four sampling masks.

reveal that the method WTBMDU-L1 with $p = 0.5$ provides a more accurate reconstruction on image contrast and sharper anatomical depiction. Compared to DLMRI, the magnitude images of the reconstruction error shown in Figures 1(h) and 1(i) indicate that our method exhibits crisper reconstruction of object edges (the large anatomical structure in the middle region) and preserves finer texture information (the gray matter regions in the bottom-right of the reconstruction). In general, our proposed method provides better intensity fidelity to the fully sampled image.

More obvious differences in visual quality can be observed in the case of 95% undersampling as shown in Figures 1(j) and 1(k). The obtained PSNRs of DLMRI and WTBMDU-L1 with $p = 0.5$ are 26.30 dB and 28.13 dB, respectively. The DLMRI reconstruction in Figure 1(j) based on k-SVD dictionary updating and greedy pursuit of coefficients shows a large number of spurious oscillations, although it gives much improvement than zero-filling and LDP (not shown in this paper). In contrast, WTBMDU shown in Figure 1(k) results in much fewer spurious oscillations and

better preservation of edges through iterative weighting of the coefficients under the data-adaptive dictionary. This is especially noticeable for the brain's fissures with sharper anatomical edges.

Figure 2 compares the results generated by DLMRI and WTBMDU using four sampling trajectories roughly under the same undersampling percentage: variable density random with 87% undersampling [15], Cartesian approximation of multishot spiral with 86% undersampling [15], 1D Cartesian trajectory [29], and pseudo radial sampling with 86% undersampling (i.e., 7.11-fold acceleration). The test image is the axial T2-weighted brain image shown in Figure 1(a). As can be observed from the error images, WTBMDU-L1 performs better in reducing aliasing artifacts and maintaining fine details than DLMRI with all verified sampling trajectories.

Table 1 lists the PSNR values of the axial T2-weighted brain image at different sampling trajectories with the same undersampling percentage using DLMRI, TBMDU, and WTBMDU. Generally, the improvements gained by WTBMDU over other methods are different for four kinds

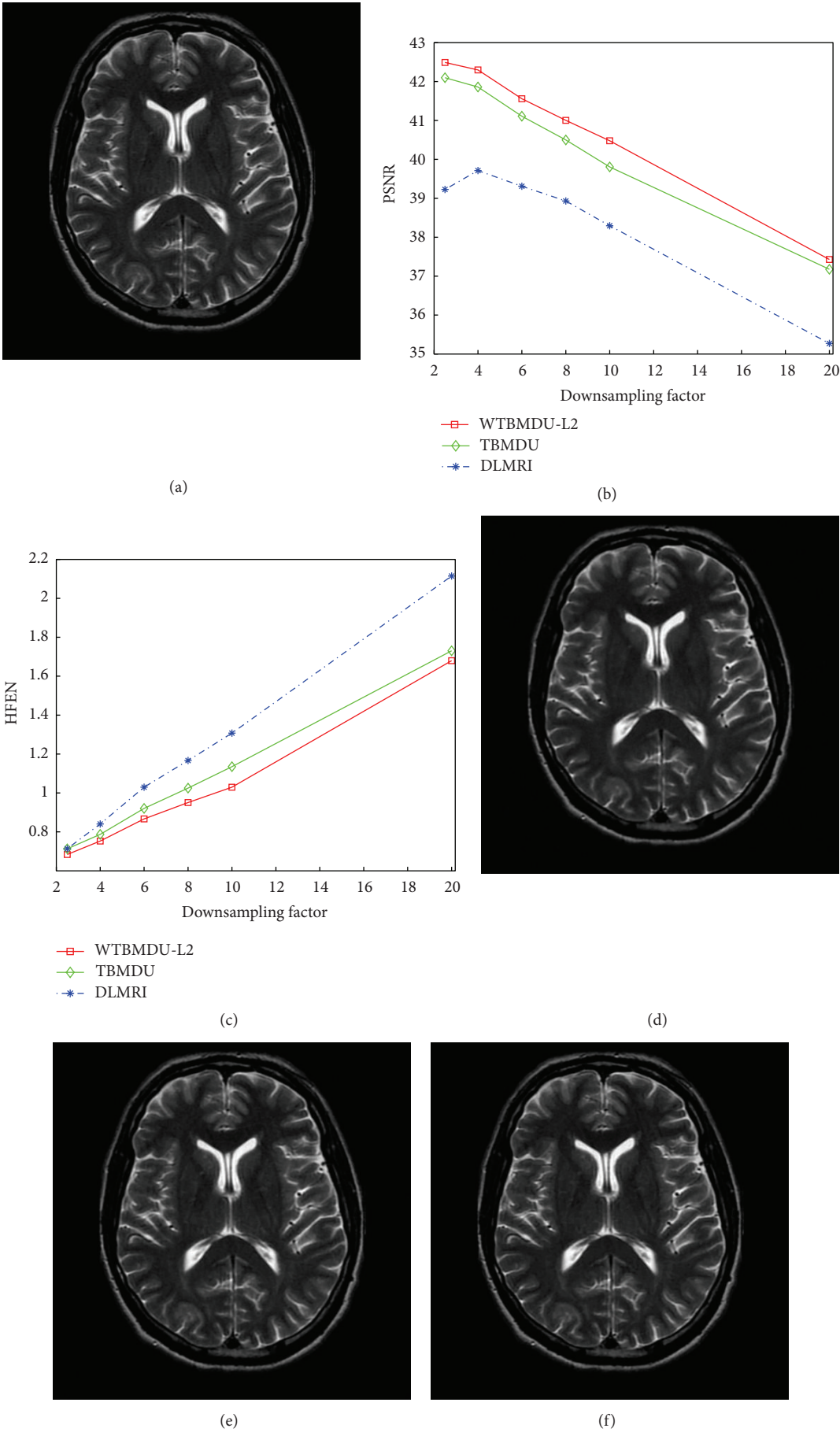


FIGURE 3: Continued.

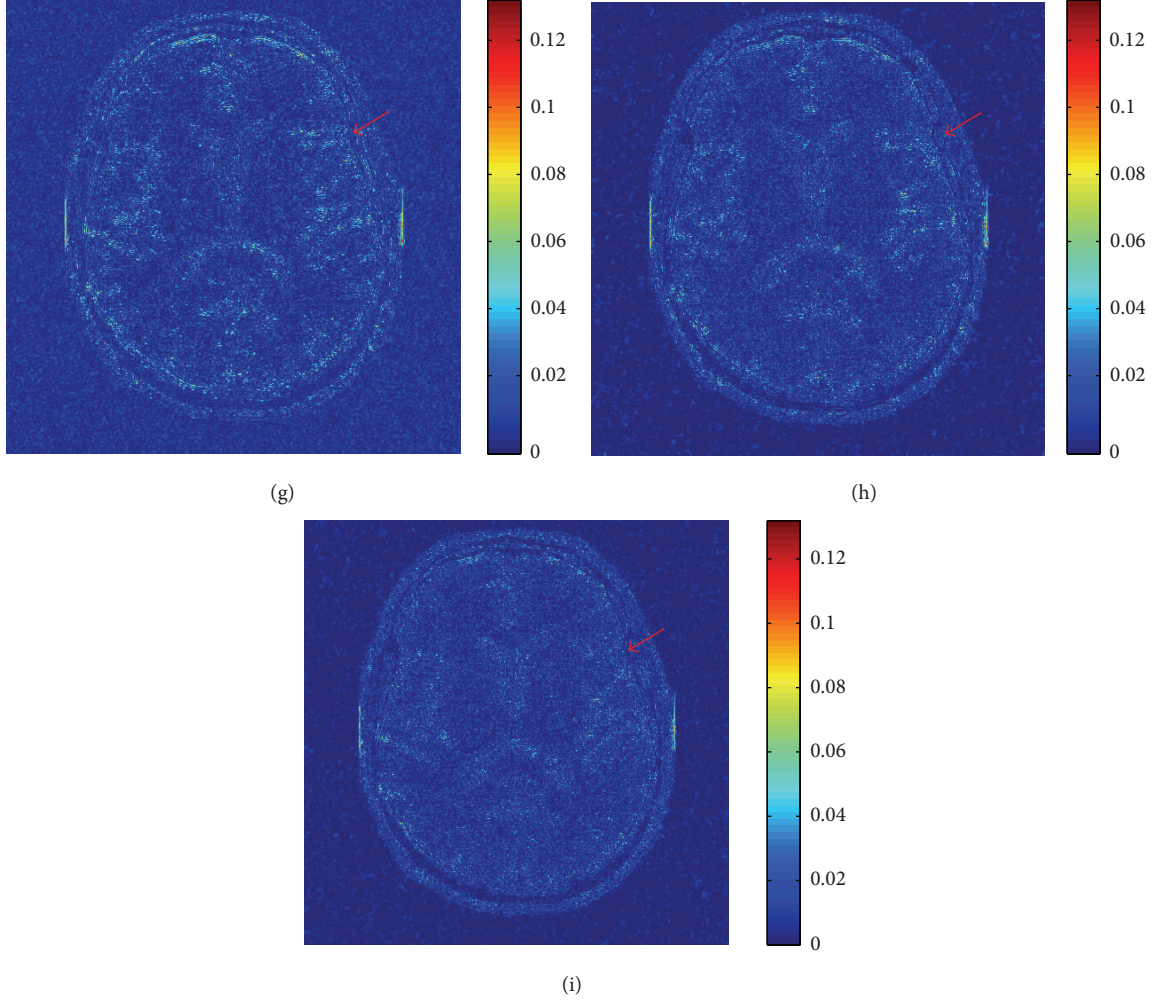


FIGURE 3: (a) The reference image. ((b) and (c)) The PSNR and HFEN value versus undersampling factor for three methods DLMRI, TBMDU, and WTBMDU at 2.5-, 4-, 6-, 8-, 10-, and 20-fold undersampling. ((d), (e), (f) and (g), (h), (i)) Reconstructions and corresponding error magnitudes for DLMRI, TBMDU, and WTBMDU at 10-fold undersampling, respectively.

TABLE 1: The reconstruction PSNR values of the axial T2-weighted brain image at different sampling trajectories with the same undersampling percentage using different methods.

Sampling mask	2D random sampling	Spiral sampling	Cartesian sampling	Radial sampling
DLMRI	40.54	37.12	35.78	35.32
TBMDU	44.01	38.54	37.41	36.06
WTBMDU-L2, $p = 0.3$	45.42	38.31	37.32	36.55
WTBMDU-L2, $p = 0.5$	44.94	38.86	37.61	36.82
WTBMDU-L2, $p = 0.7$	44.78	39.22	37.70	36.95
WTBMDU-L1, $p = 0.3$	45.20	38.82	37.62	36.80
WTBMDU-L1, $p = 0.5$	44.88	39.20	37.73	37.16
WTBMDU-L1, $p = 0.7$	45.18	39.15	38.08	37.06

of trajectories although under the same undersampling rate. The largest and smallest improvements were achieved with

the 2D random and radial sampling, respectively, where roughly 5 dB and 2 dB were obtained. This indicates that the efficiency of dictionary learning methods may depend on the incoherence of the data acquisition. In the family of WTBMDU algorithms, the optimal p value is also different at various trajectories for both WTBMDU-L1 and WTBMDU-L2. To balance the numerical calculation and the selection of trajectories, $p = 0.5$ or $p = 0.7$ is a good option.

Figure 3 illustrates the performance of DLMRI, TBMDU, and WTBMDU at a range of acceleration factors including 2.5, 4, 6, 8, 10, and 20, where zero-mean complex white Gaussian noise with standard deviation $\sigma = 10.2$ was added to the 2D random sampled k -space [29]. Since the stopping rule for the outer loop of both TBMDU and WTBMDU is determined by $\|F_p u^k - f\|_2 < \sigma$, the number of outer iterations k_{\max} of TBMDU and WTBMDU-L2 with $p = 0.7$ take the values of 3, 3, 6, 7, 7, 11 and 4, 8, 8, 10, 11, 11 for the above six acceleration factors, respectively. For the quantitative comparison, the values of PSNR and HFEN as functions of acceleration factor are shown in Figures 3(b)

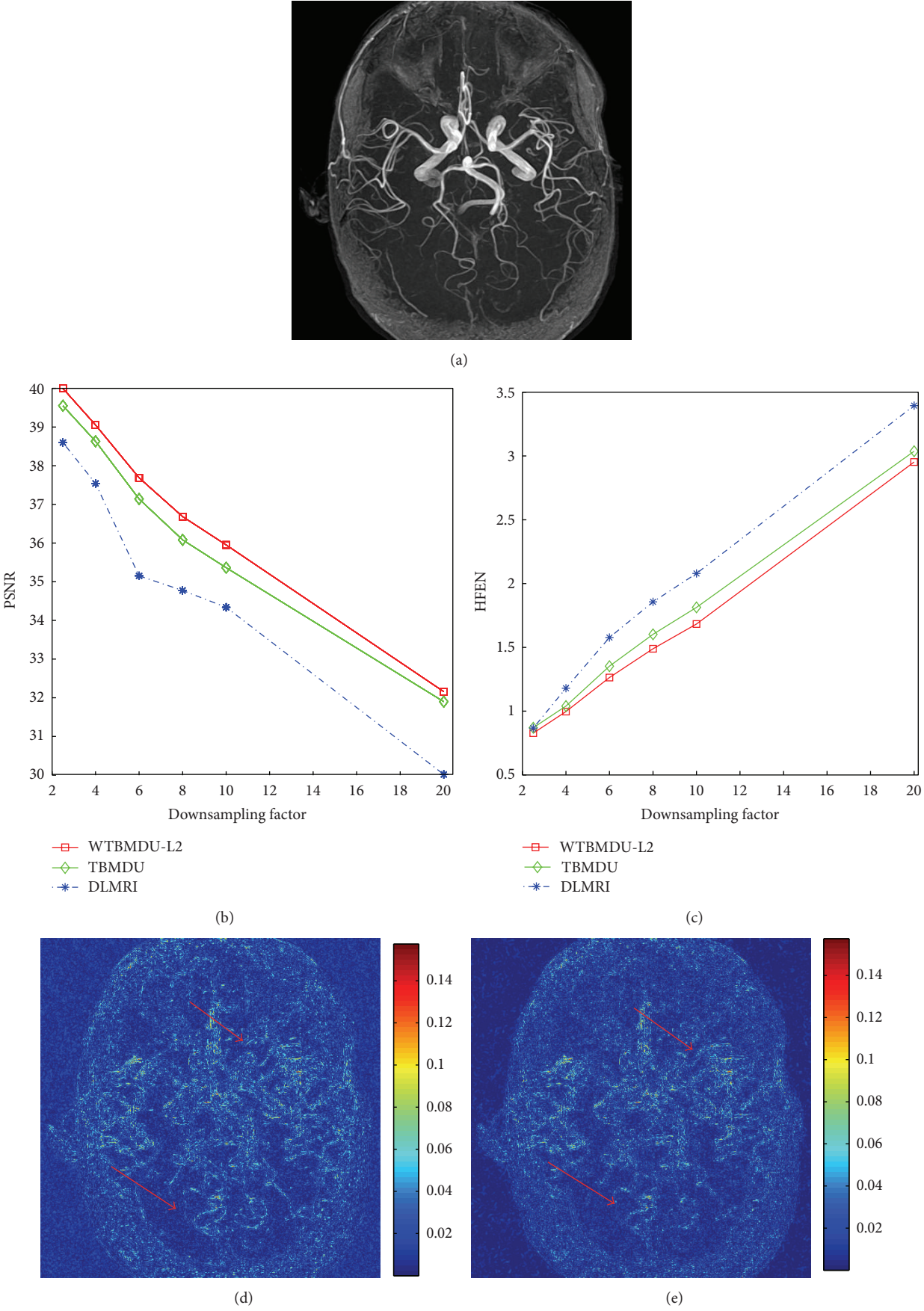


FIGURE 4: Continued.

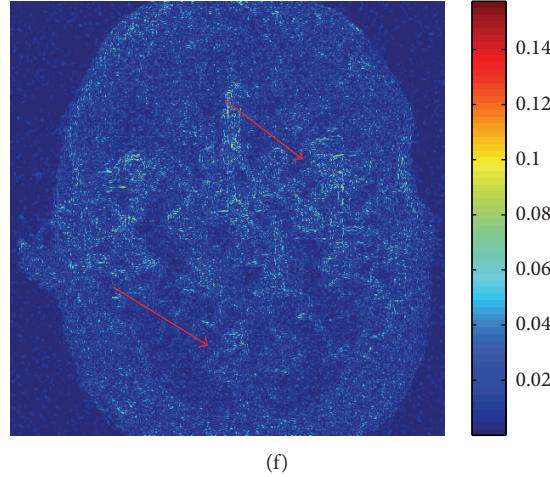


FIGURE 4: (a) The reference noncontrast MRA of the Circle of Willis. ((b) and (c)) The PSNR and HFEN values versus acceleration factor for the three methods DLMRI, TBMDU, and WTBM DU at 2.5-, 4-, 6-, 8-, 10-, and 20-fold acceleration. ((d), (e), and (f)) Reconstruction error magnitudes for DLMRI, TBMDU, and WTBM DU at 10-fold acceleration, respectively.

and 3(c). It can be seen that WTBM DU performs best at all tested acceleration factors. The reconstruction results and the corresponding error images under 10-fold acceleration using three methods are displayed in Figures 3(d), 3(e), and 3(f) and 3(g), 3(h), and 3(i), respectively.

As shown in the error images, the brighter the error image appears, the larger the deviation between the reconstruction and the reference image will be. WTBM DU (Figure 3(i)) presents less pixel errors and structure loss than that of DLMRI in Figure 3(g) and TBMDU in Figure 3(h), especially in regions indicated by red arrows. In general, it can be concluded that both WTBM DU-L2 and TBMDU offer strong preservation of details compared with that of DLMRI, while WTBM DU-L2 provides a crisper result of fine structures than that of TBMDU.

Figure 4 depicts the reconstruction results of a transverse slice of a noncontrast MR angiography (MRA) of the circle of Willis (COW) at the same experiment setting as in Figure 3. The MRA of the circle of Willis has much textural information such as the vessels on the middle region and fine-scale details on the bottom region. As can be seen from Figure 4(b), when the acceleration factor increased until 10-fold, the PSNR gap between TBMDU and WTBM DU increases synchronously. The PSNR improvement is also reflected in Figures 4(e) and 4(f) where much less errors appeared in the WTBM DU error image, indicating WTBM DU performs better in maintaining fine details. On the other hand, similar as observed in Figure 3, when the acceleration factor increased to as large as 20-fold, the advantage of the nonconvex optimization degraded and none of these methods can faithfully reconstruct the original image from such fewer k -space samples with additional noise.

To investigate the sensitivity of various methods to different levels of complex white Gaussian noise, DLMRI, TBMDU, and WTBM DU were applied to reconstruct a T2-weighted sagittal view of the lumbar spine under pseudo radial sampling at 6.09-fold acceleration. Figure 5(c) plots

the PSNR values of the recovered MR images by DLMRI (blue curves), TBMDU (green curves), and WTBM DU (red curves) at a sequence of different standard deviations ($\sigma = 2, 5, 8, 10, 14.2$). In the case of $\sigma = 2$, the PSNR of the image obtained by DLMRI is only 38.22 dB, TBMDU is 39.49 dB, and WTBM DU-L2 with $p = 0.7$ reaches 40.52 dB. Obviously, the difference gap between three methods is significant at low noise levels. The reconstructions and corresponding magnitudes of the error images with $\sigma = 8$ are shown in Figures 5(d), 5(e), 5(f), 5(g), 5(h), and 5(i). It can be observed that the skeletons in the top half part of the TBMDU reconstruction appear less obscured than those in the DLMRI results. Meanwhile, the reconstruction by WTBM DU is clearer and sharper than that by DLMRI and TBMDU and is relatively devoid of aliasing artifacts. This reveals that our method provides a more accurate reconstruction of image contrast and sharper anatomical depiction in noisy case.

4.2. Reconstruction of Experimental MRI Data (Complex-Valued). Figure 6 shows the comparison between different methods on a physical phantom, which is often used to assess the resolution of MR reconstruction. Figure 6(a) displays the fully sampled reconstruction of the physical phantom. Figures 6(c), 6(d), 6(e), and 6(f) exhibit the results of DLMRI, TBMDU, WTBM DU-L1, and WTBM DU-L2 with $p = 0.7$ at 80% undersampling ratio. The corresponding PSNR values are 18.66 dB, 26.82 dB, 28.89 dB, and 29.12 dB, respectively. We can find that the WTBM DU reconstructions exhibit higher resolution than those with DLMRI and TBMDU and are almost devoid of aliasing artifacts especially in zoomed-in regions.

To investigate the noise sensitivity of proposed method on the complex k -space data, complex Gaussian noise of $\sigma = 30$ was added to the k -space with 5-fold acceleration. In this case, the PSNR values of DLMRI, TBMDU, and WTBM DU-L2 with $p = 0.7$ methods are 17.93 dB, 22.94 dB, and 25.42 dB, respectively. The reconstructions of three methods

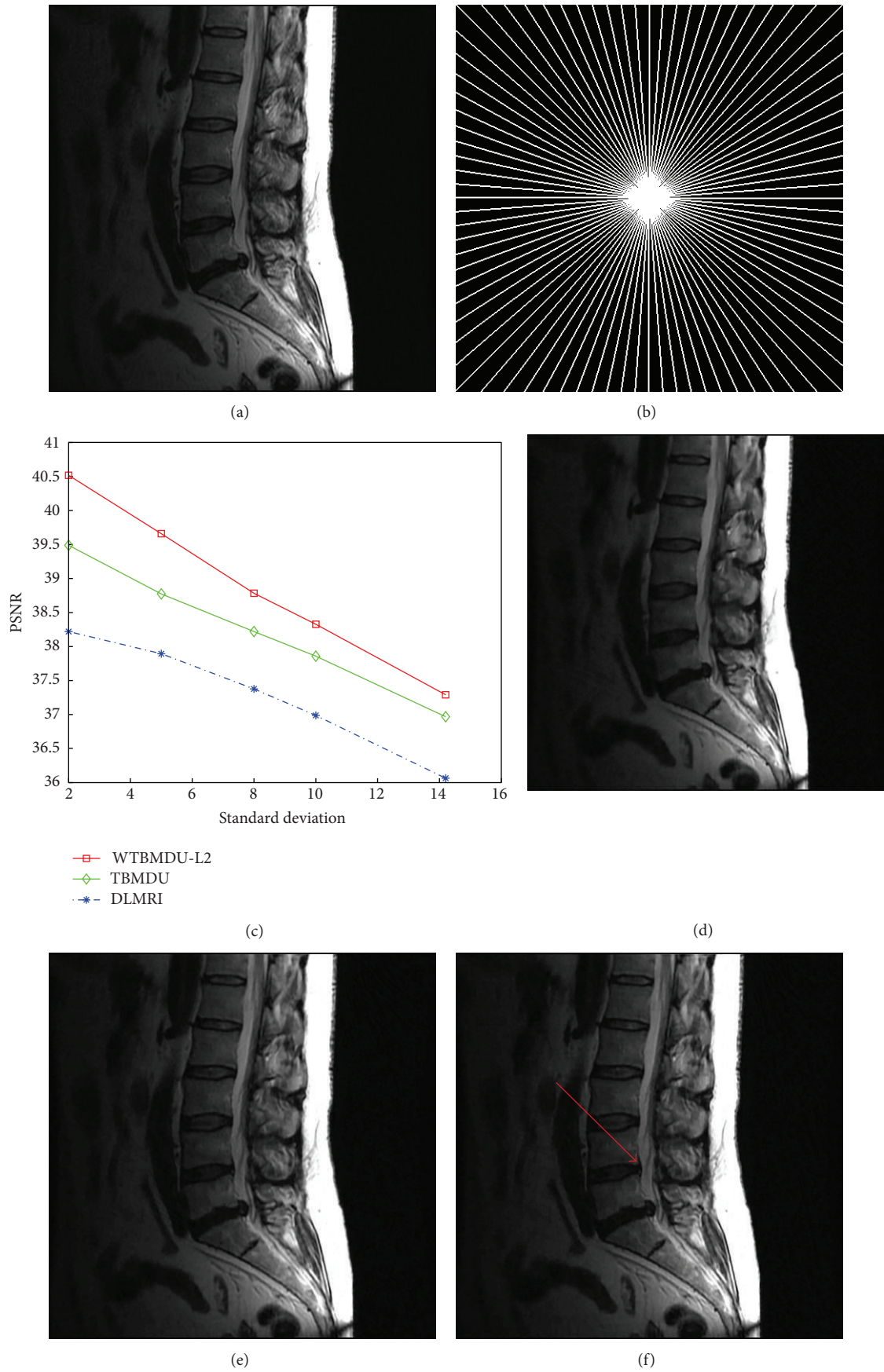


FIGURE 5: Continued.

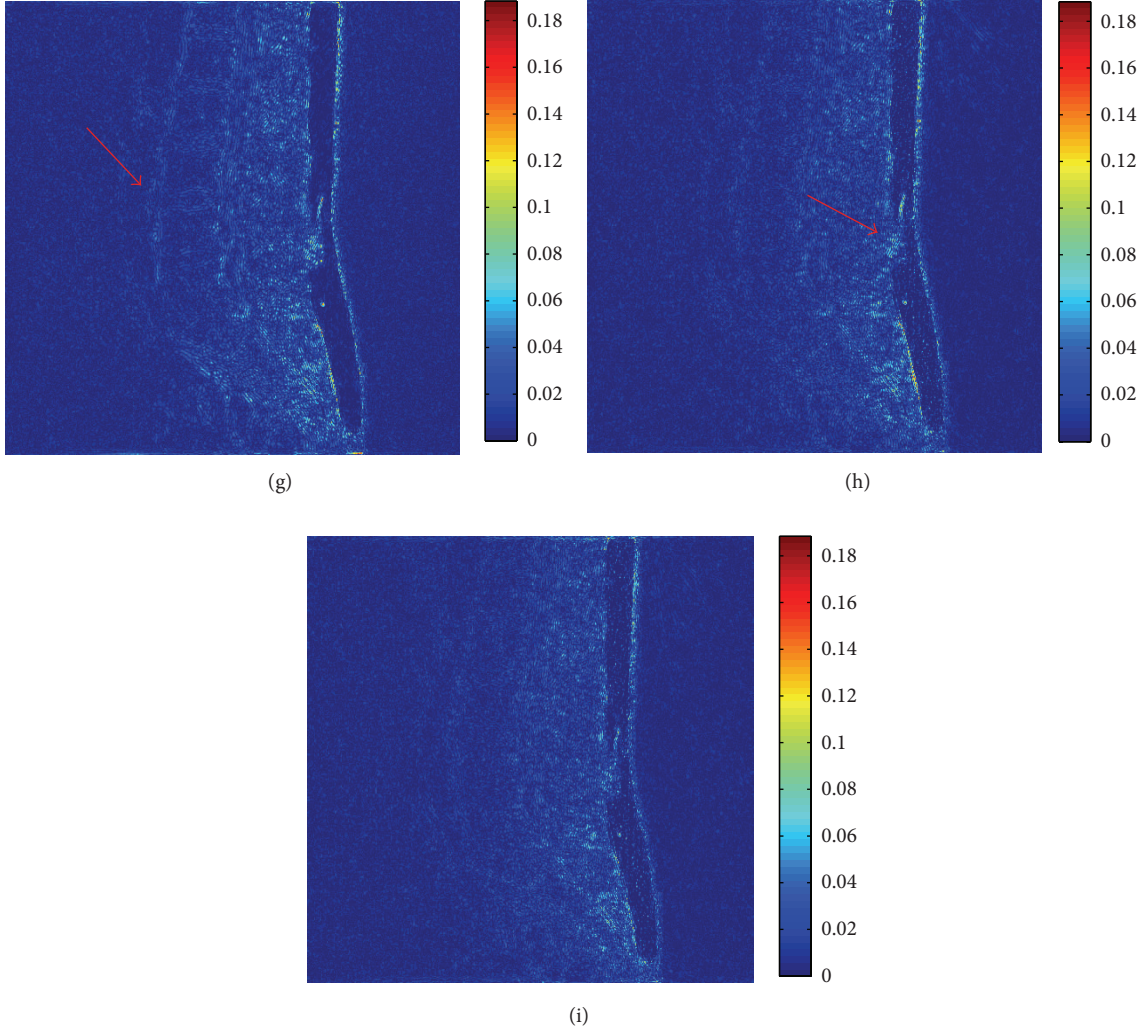


FIGURE 5: (a) The reference T2-weighted sagittal view of the lumbar spine. (b) Sampling mask in k -space with 6.09-fold undersampling. (c) PSNR versus noise level for DLMRI, TBMDU, and WTBMDU. ((d), (e), (f) and (g), (h), (i)) Reconstructions and corresponding error magnitudes for DLMRI, TBMDU, and WTBMDU with noise $\sigma = 8$.

are shown in Figure 7. The enlargements of two region-of-interests (ROIs) are presented in Figures 7(d) and 7(e). As can be observed, the DLMRI reconstruction exhibits more oscillating artifacts than that of the other two methods. Besides, the spot and circle in the bottom right of the WTBMDU reconstruction appear less obscured than those in the DLMRI and TBMDU results.

In Figures 8 and 9, two real brain data sets containing more fine-detailed structures [44, 46] were used for comparison. The images in Figures 8(a) and 9(a) are both the fully-sampled reconstruction of size 256×256 as references. A variable density Cartesian sampling trajectory with 80% undersampling was employed as shown in Figure 8(b). The reconstructions from two data sets using DLMRI, TBMDU and WTBMDU-L2 with $p = 0.7$ are displayed in Figures 8(b), 8(c), and 8(d) and Figures 9(b), 9(c), and 9(d), respectively. The corresponding PSNR values of three methods are 31.35 dB, 32.04 dB, and 32.69 dB for reconstructions in

Figure 8 and 30.58 dB, 31.59 dB, and 31.82 dB for those in Figure 9. Some regions of error in these subplots have been illustrated with red arrows, indicating that the reconstruction using WTBMDU shows better fidelity to the reference image than that of DLMRI and TBMDU. The enlargements of the reconstruction results with these methods are shown in Figures 8(f) and 9(e). It can be observed that visible aliasing artifacts along the phase encoding direction (horizontal in the image plane) in the DLMRI reconstruction is more pronounced than those of TBMDU reconstruction, while there are almost no artifacts in that of WTBMDU with $p = 0.7$. Close inspection of the zoomed-in images indicates that the weighted sparse regularization enhances the reconstruction quality.

In summary, the results from our work clearly demonstrate the advantages of introducing iterative reweighting scheme and alternating direction method in nonconvex optimization over conventional methods, for constrained

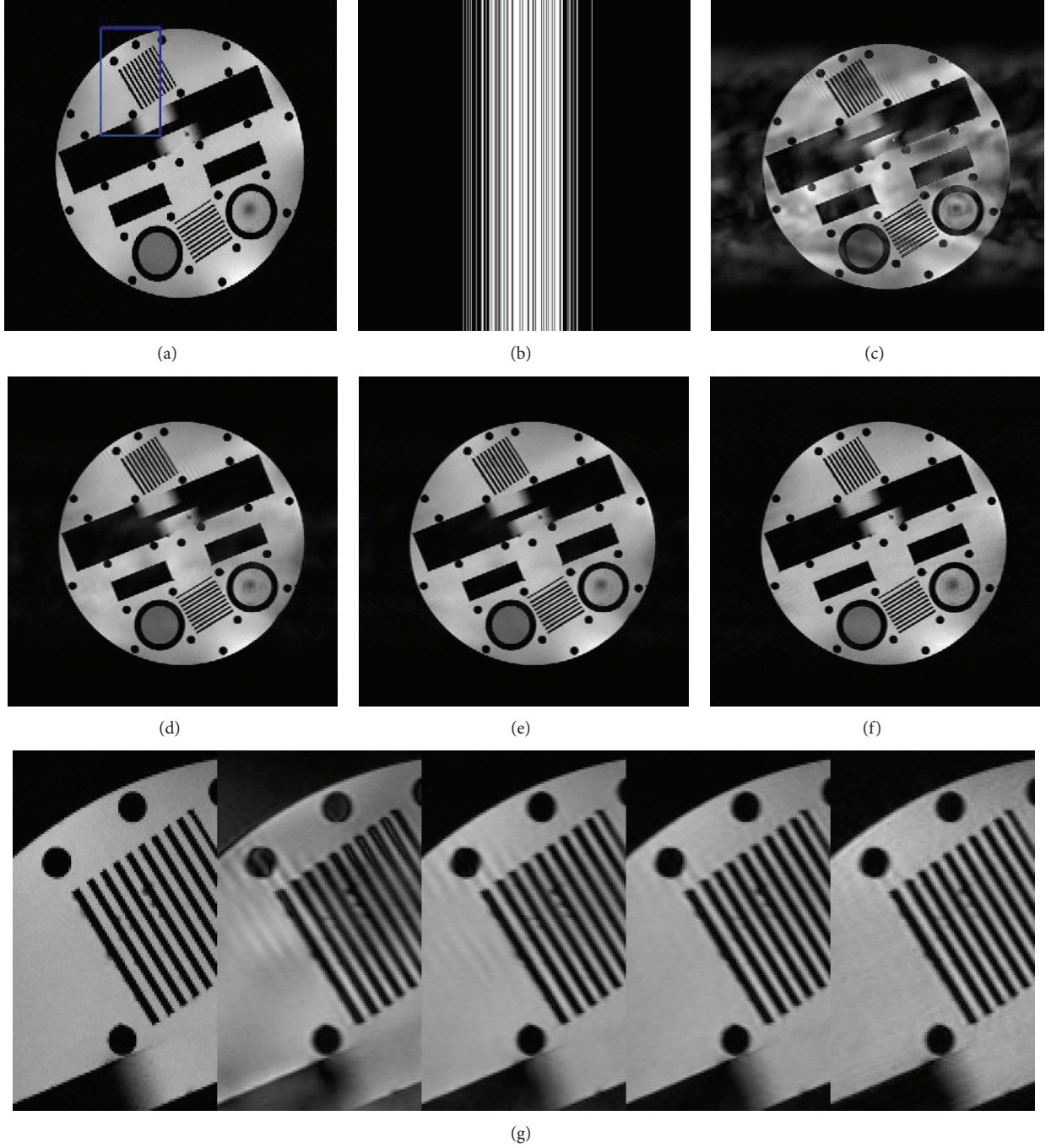


FIGURE 6: Reconstruction comparison on the real brain data. (a) Reference image corresponding to fully sampled reconstruction. (b) Fourier domain sampling pattern. ((c), (d), (e), and (f)) Reconstruction using DLMRI, TBMDU, WTBMDU-L1, and WTBMDU-L2 with $p = 0.7$ at 80% undersampling, respectively. (g) Enlargements of (a), (c), (d), (e), and (f).

image reconstruction from undersampled k -space data. Not only for the brain image containing large piecewise constant regions, as shown in Figures 1, 2, and 3, but also for the circle of Willis and lumbar spine composed of many texture features as shown in Figures 4 and 5, the proposed method visually provides more pleasant results than existing methods. By means of the Bregman iteration/AL methodology, it is notable

that the proposed method consistently supports superior results without any parameters tuned manually regardless of different sampling trajectories, varying sampling factors, and the existence of noise or not. We note that, in some circumstances, the PSNR value started to decrease a little bit after achieving the highest value. The reason of this phenomenon is that the solution of problem (1) does not

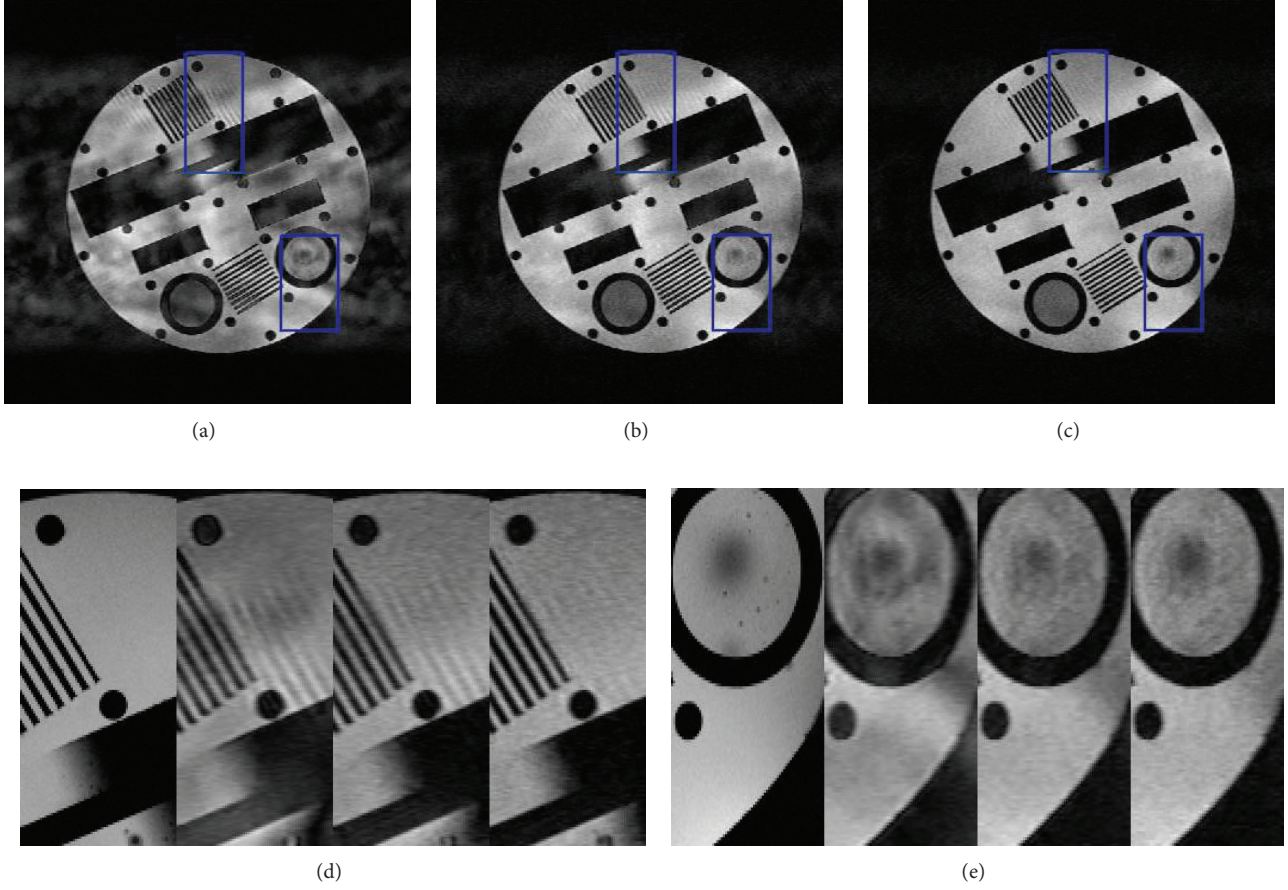


FIGURE 7: Comparison of reconstructing a physical phantom MR image with noise of $\sigma = 30$. ((a), (b), and (c)) Reconstructions using DLMRI, TBMDU, and WTBM DU at 5-fold undersampling, respectively. ((d) and (e)) Enlargements of the references (a), (b), and (c).

necessarily have higher PSNR than the intermediate iterates. In real-time imaging applications where the acquisition and reconstruction speed is crucial, a numerical algorithm which converges fast at the beginning iterations is highly appreciated and useful.

5. Conclusions and Discussions

This work presents a new regularizer combining the non-convex l_p pseudonorm and adaptive dictionary for MRI reconstruction from undersampled k -space data. Based on the potential success of TBMDU in the AL framework for CS-MRI, the proposed WTBM DU method extends the plain TBMDU to handle nonconvex l_p regularization by weighting the sparse coding of the coefficients. The strategy of employing iteratively reweighted l_1 or l_2 technique results in the WTBM DU-L1 and WTBM DU-L2 algorithms, respectively. High-accuracy reconstructions are obtained from significantly undersampled k -space data by minimizing the adaptively sparsity-promoting regularization criterion subject to the data-consistency constraint. In particular, compared to the DLMRI method that minimizes the l_0 regularization

by the greedy algorithm of matching pursuit, the iteratively reweighted strategy combined with AL framework yields much higher PSNR values. On the other hand, compared to the plain TBMDU for l_1 -norm minimization, the adaptively weighted method achieves 0.3–2.5 dB PSNR improvement and visually provides more pleasure results. Various experimental results demonstrate the superior performance of the method under a variety of sampling trajectories and k -space acceleration factors.

WTBM DU is very general since it represents an algorithmic framework that can be easily adapted to different reweighting strategies and nonconvex sparsity-inducing functions proposed in the previous literatures. Some extensions will be studied in future as follows: (i) for the reweighted l_1 strategy, using different nonconvex objective functions as conducted in [38, 41]; besides, the shrinkage can be generalized as in [47, 48]; (ii) for the reweighted l_2 strategy, trying different nonconvex objective functions as in [39]. Additionally, the proposed scheme can be easily extended to incorporate other prior (e.g., spike and slab prior used in [49], where the regularization consists of the l_2 -norm and l_0 -norm terms). It may be better for stabilizing the numerical process; (iii) the current iteratively reweighted strategies can

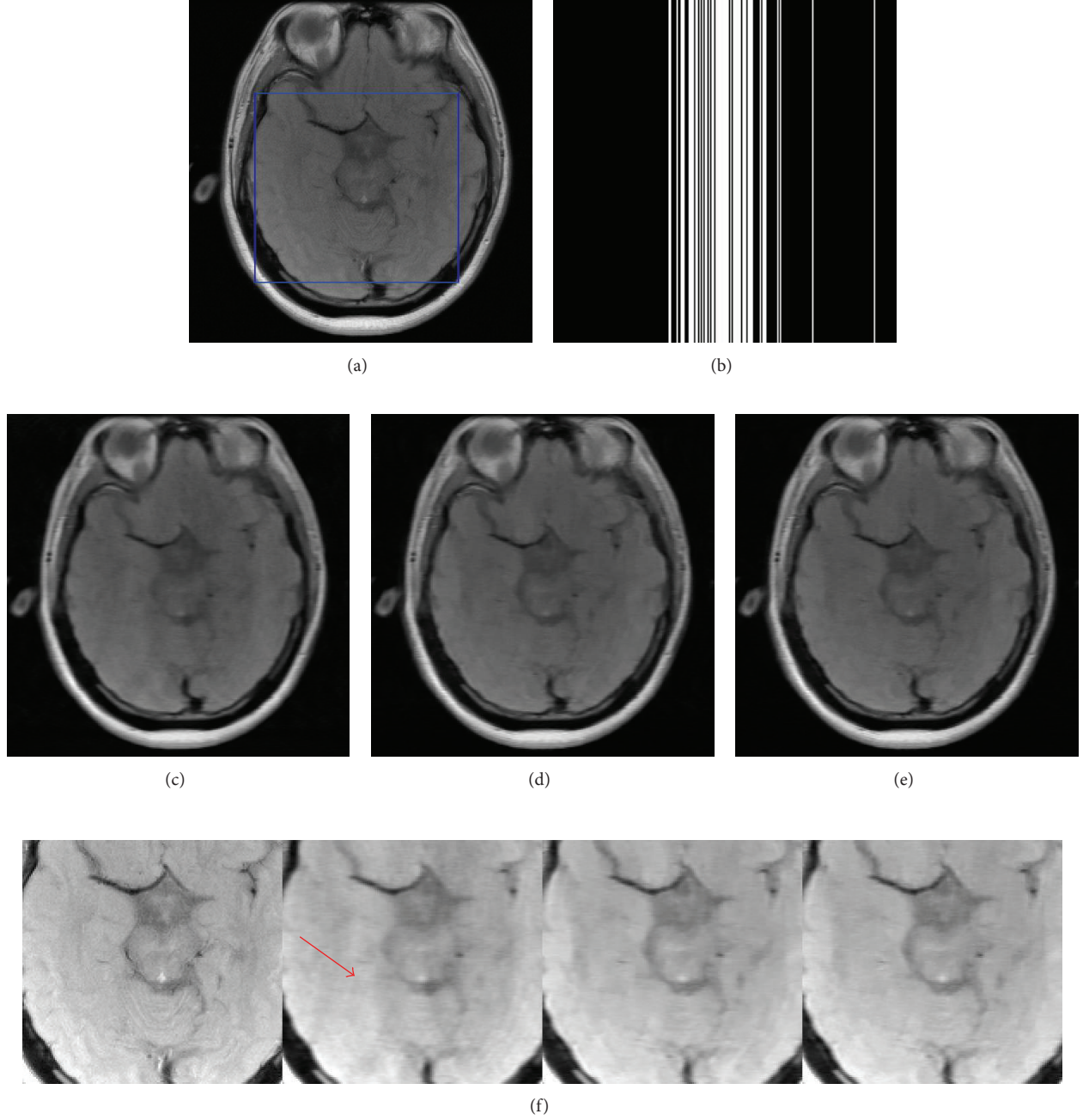


FIGURE 8: Reconstruction comparison of the real brain data. (a) Reference image corresponding to fully sampled reconstruction. (b) Sampling mask in k -space. ((c), (d), and (e)) Reconstruction using DLMRI, TBMDU, and WTBM DU-L2 with $p = 0.7$ at 80% undersampling, respectively. (f) Enlargements of (a), (c), (d), and (e).

be combined with the homotopic technique described in [14–16], where the value of p slowly decreases from 1 to 0.

Another extension we will consider in future is to extend our proposed model to parallel MRI reconstruction. We note that a very recently published paper [21], under the work of Ramani and Fessler, addressed the regularized SENSE-reconstruction using AL methods. In order to effectively solve the unconstrained SENSE-reconstruction consisting of

TV/wavelets regularization term and data-fidelity term using the AL formalism, they split not only the regularization term, but also the Fourier encoding and spatial components in the data-fidelity term, by introducing auxiliary variables to decouple the data-domain components, and the regularization component, respectively. We believe that the regularizer of the sparse representations of overlapping image patches in our work can be naturally incorporated into their work in a unified AL framework.

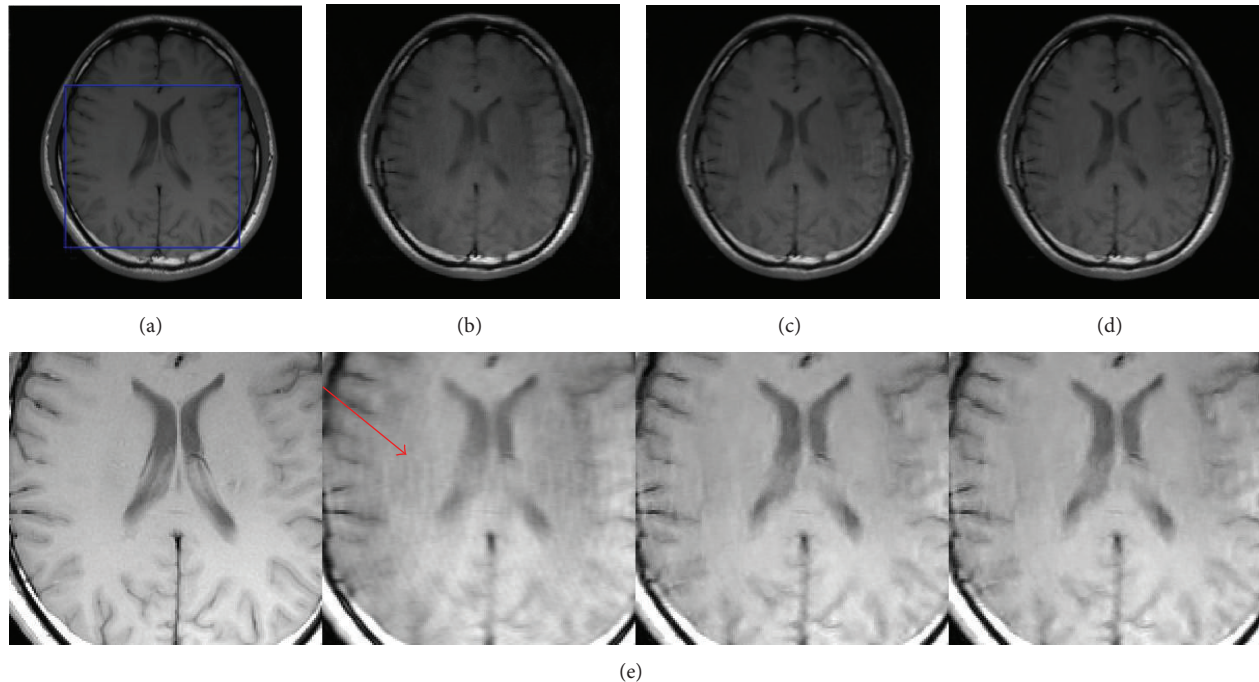


FIGURE 9: Reconstruction comparison of the real brain data. (a) Reference image corresponding to fully sampled reconstruction. ((b), (c), and (d)) Reconstruction using DLMRI, WTBMDU-L1, and WTBMDU-L2 with $p = 0.7$ at 80% undersampling, respectively. (e) Enlargements of (a), (b), (c), and (d).

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors would like to thank Dr. Leslie Ying at the State University of New York at Buffalo for helpful discussion and thank Ravishankar et al. for sharing their experiment materials and source codes. This work was partly supported by the National Natural Science Foundation of China (61362001, 51165033, 61340025, 61261010, 81120108012, and 11301508), the Natural Science Foundation of Jiangxi Province (20132BAB211030, 20121BBE50023, and 20122BAB211015), Technology Foundation of Department of Education in Jiangxi Province (nos. GJJ12006 and GJJ14196), Young Scientists Training Plan of Jiangxi province (no. 20142BCB23001), the Shenzhen Peacock Plan KQCX20120816155710259, and the Basic Research Program of Shenzhen JC201104220219A.

References

- [1] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: the application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.
- [2] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [3] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [4] S. L. Keeling, C. Clason, M. Hintermüller, F. Knoll, A. Laurain, and G. von Winckel, "An image space approach to Cartesian based parallel MR imaging with total variation regularization," *Medical Image Analysis*, vol. 16, no. 1, pp. 189–200, 2012.
- [5] D. Liang, B. Liu, J. Wang, and L. Ying, "Accelerating SENSE using compressed sensing," *Magnetic Resonance in Medicine*, vol. 62, no. 6, pp. 1574–1584, 2009.
- [6] F.-H. Lin, K. K. Kwong, J. W. Belliveau, and L. L. Wald, "Parallel imaging reconstruction using automatic regularization," *Magnetic Resonance in Medicine*, vol. 51, no. 3, pp. 559–567, 2004.
- [7] F.-H. Lin, F.-N. Wang, S. P. Ahlfors, M. S. Hämläinen, and J. W. Belliveau, "Parallel MRI reconstruction using variance partitioning regularization," *Magnetic Resonance in Medicine*, vol. 58, no. 4, pp. 735–744, 2007.
- [8] P. Qu, J. Luo, B. Zhang, J. Wang, and G. X. Shen, "An improved iterative SENSE reconstruction method," *Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering*, vol. 31, no. 1, pp. 44–50, 2007.
- [9] X. Qu, Y. Hou, F. Lam, D. Guo, J. Zhong, and Z. Chen, "Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator," *Medical Image Analysis*, vol. 18, no. 6, pp. 843–856, 2014.
- [10] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [11] L. Ying, D. Xu, and Z.-P. Liang, "On Tikhonov regularization for image reconstruction in parallel MRI," in *Proceedings of the IEEE Engineering in Medicine and Biology Conference (EMBC '04)*, pp. 1056–1059, September 2004.

- [12] K. T. Block, M. Uecker, and J. Frahm, "Undersampled radial MRI with multiple coils. Iterative image reconstruction using a total variation constraint," *Magnetic Resonance in Medicine*, vol. 57, no. 6, pp. 1086–1098, 2007.
- [13] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, "An efficient algorithm for compressed MR imaging using total variation and wavelets," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, June 2008.
- [14] J. Shi, X. Ren, G. Dai, J. Wang, and Z. Zhang, "A non-convex relaxation approach to sparse dictionary learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 1809–1816, June 2011.
- [15] J. Trzasko and A. Manduca, "Highly undersampled magnetic resonance image reconstruction via homotopic ℓ_0 -minimization," *IEEE Transactions on Medical Imaging*, vol. 28, no. 1, pp. 106–121, 2009.
- [16] A. Wong, A. Mishra, P. Fieguth, and D. A. Clausi, "Sparse reconstruction of breast MRI using homotopic ℓ_0 minimization in a regional sparsified domain," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 3, pp. 743–752, 2013.
- [17] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *Proceedings of the IEEE 33rd International Conference on Acoustics, Speech and Signal Processing (ICASSP '08)*, pp. 3869–3872, Las Vegas, Nev, USA, April 2008.
- [18] R. Chartrand, "Exact reconstruction of sparse signals via non-convex minimization," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 707–710, 2007.
- [19] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5–6, pp. 877–905, 2008.
- [20] M. Akcakaya, S. Nam, P. Hu et al., "Compressed sensing with wavelet domain dependencies for coronary MRI: a retrospective study," *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1090–1099, 2011.
- [21] S. Ramani and J. A. Fessler, "Parallel MR image reconstruction using augmented lagrangian methods," *IEEE Transactions on Medical Imaging*, vol. 30, no. 3, pp. 694–706, 2011.
- [22] F. Knoll, K. Bredies, T. Pock, and R. Stollberger, "Second order total generalized variation (TGV) for MRI," *Magnetic Resonance in Medicine*, vol. 65, no. 2, pp. 480–491, 2011.
- [23] D. Liang, H. Wang, Y. Chang, and L. Ying, "Sensitivity encoding reconstruction with nonlocal total variation regularization," *Magnetic Resonance in Medicine*, vol. 65, no. 5, pp. 1384–1392, 2011.
- [24] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [25] A. Bilgin, Y. Kim, F. Liu, and M. Nadar, "Dictionary design for compressed sensing MRI," in *Proceedings of the 18th Annual Meeting on International Society for Magnetic Resonance in Medicine (ISMRM '10)*, p. 4887, 2010.
- [26] Y. Chen, X. Ye, and F. Huang, "A novel method and fast algorithm for MR image reconstruction with significantly undersampled data," *Inverse Problems and Imaging*, vol. 4, no. 2, pp. 223–240, 2010.
- [27] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736–3745, 2006.
- [28] R. Otazo and D. Sodickson, "Adaptive compressed sensing MRI," in *Proceedings of the 18th Annual Meeting on International Society for Magnetic Resonance in Medicine (ISMRM '10)*, p. 4867, May 2010.
- [29] S. Ravishanker and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 30, no. 5, pp. 1028–1041, 2011.
- [30] Q. Liu, S. Wang, and J. Luo, "A novel predual dictionary learning algorithm," *Journal of Visual Communication and Image Representation*, vol. 23, no. 1, pp. 182–193, 2012.
- [31] Q. Liu, S. Wang, K. Yang, J. Luo, Y. Zhu, and D. Liang, "Highly Undersampled magnetic resonance image reconstruction using two-level Bregman method with dictionary updating," *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1290–1301, 2013.
- [32] Q. Liu, J. Luo, S. Wang, M. Xiao, and M. Ye, "An augmented Lagrangian multi-scale dictionary learning algorithm," *EURASIP Journal on Advances in Signal Processing*, vol. 1, pp. 1–16, 2011.
- [33] W. Yin, S. Osher, D. Goldfarb, and J. Darbon, "Bregman iterative algorithms for ℓ_1 -minimization with applications to compressed sensing," *SIAM Journal on Imaging Sciences*, vol. 1, no. 1, pp. 143–168, 2008.
- [34] 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*, vol. 57, no. 11, pp. 1413–1457, 2004.
- [35] A. Mehranian, H. Saligheh Rad, M. R. Ay, A. Rahmim, and H. Zaidi, "Smoothly clipped absolute deviation (SCAD) regularization for compressed sensing MRI using an augmented Lagrangian scheme," in *Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC '12)*, pp. 3646–3653, Anaheim, Calif, USA, November 2012.
- [36] I. F. Gorodnitsky and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm," *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.
- [37] Q. Lyu, Z. Lin, Y. She, and C. Zhang, "A comparison of typical ℓ_p minimization algorithms," *Neurocomputing*, vol. 119, pp. 413–424, 2013.
- [38] L. B. Montefusco, D. Lazzaro, and S. Papi, "A fast algorithm for nonconvex approaches to sparse recovery problems," *Signal Processing*, vol. 93, no. 9, pp. 2636–2647, 2013.
- [39] N. Mourad and J. P. Reilly, "Minimizing nonconvex functions for sparse vector reconstruction," *IEEE Transactions on Signal Processing*, vol. 58, no. 7, pp. 3485–3496, 2010.
- [40] B. D. Rao and K. Kreutz-Delgado, "An affine scaling methodology for best basis selection," *IEEE Transactions on Signal Processing*, vol. 47, no. 1, pp. 187–200, 1999.
- [41] H. Zou and R. Li, "One-step sparse estimates in nonconcave penalized likelihood models," *The Annals of Statistics*, vol. 36, no. 4, pp. 1509–1533, 2008.
- [42] P. Rodríguez and B. Wohlberg, "Efficient minimization method for a generalized total variation functional," *IEEE Transactions on Image Processing*, vol. 18, no. 2, pp. 322–332, 2009.
- [43] L. B. Montefusco, D. Lazzaro, and S. Papi, "Fast sparse image reconstruction using adaptive nonlinear filtering," *IEEE Transactions on Image Processing*, vol. 20, no. 2, pp. 534–544, 2011.
- [44] B. Liu, K. King, M. Steckner, J. Xie, J. Sheng, and L. Ying, "Regularized sensitivity encoding (SENSE) reconstruction using

- bregman iterations,” *Magnetic Resonance in Medicine*, vol. 61, no. 1, pp. 145–152, 2009.
- [45] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, “An iterative regularization method for total variation-based image restoration,” *A SIAM Interdisciplinary Journal*, vol. 4, no. 2, pp. 460–489, 2005.
- [46] J. X. Ji, B. S. Jong, and S. D. Rane, “PULSAR: a MATLAB toolbox for parallel magnetic resonance imaging using array coils and multiple channel receivers,” *Concepts in Magnetic Resonance Part B: Magnetic Resonance Engineering*, vol. 31, no. 1, pp. 24–36, 2007.
- [47] R. Chartrand, “Fast algorithms for nonconvex compressive sensing: MRI reconstruction from very few data,” in *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI '09)*, 2009.
- [48] S. Voronin and R. Chartrand, “A new generalized thresholding algorithm for inverse problems with sparsity constraints,” in *Proceedings of the 38th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '13)*, pp. 1636–1640, Vancouver, Canada, May 2013.
- [49] X. Lu, Y. Yuan, and P. Yan, “Sparse coding for image denoising using spike and slab prior,” *Neurocomputing*, vol. 106, pp. 12–20, 2013.