INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 17, Number 3, Pages 316–331

CS-MRI RECONSTRUCTION BASED ON THE CONSTRAINED TGV-SHEARLET SCHEME

TINGTING WU[1], ZHI-FENG PANG[2,*], YOUGUO WANG[1, 3], AND YU-FEI YANG[4]

Abstract. This paper proposes a new constrained total generalized variation (TGV)-shearlet model to the compressive sensing magnetic resonance imaging (MRI) reconstruction via the simple parameter estimation scheme. Due to the non-smooth term included in the proposed model, we employ the alternating direction method of multipliers (ADMM) scheme to split the original problem into some easily solvable subproblems in order to use the convenient soft thresholding operator and the fast Fourier transformation (FFT). Since the proposed numerical algorithm belongs to the framework of the classic ADMM, the convergence can be kept. Experimental results demonstrate that the proposed method outperforms the state-of-the-art unconstrained reconstruction methods in removing artifacts and achieves lower reconstruction errors on the tested dataset.

Key words. Magnetic resonance imaging, total generalized variation, shearlet transformation, alternating direction method of multipliers (ADMM), compressive sensing.

1. Introduction

Magnetic resonance imaging (MRI) is commonly used in radiology to visualize the internal structure and function of the body by noninvasive and nonionizing means. However, the widespread use of MRI is hindered by its intrinsic slow data acquisition process. So how to speed up the scanning time has been the key in the MRI research community. Recently, compressive sensing (CS) [3] has shown great potential in speeding up MRI by under-sampling k-space data. In the meantime, reducing the acquired data which compromises with its diagnostic value may result in degrading the image quality. Considering the above reasons, finding an inversion algorithm with good practical performance in terms of image quality and reconstruction speed is crucial in clinical applications.

Let u be an ideal image scaled in $[0, 1]$ and set $A = P\mathcal{F}$, where P is a selection matrix and $\mathcal{F}$ is the Fourier transformation matrix. Accordingly, the undersampling k-space data f involved in the sampling matrix A and the additive noise η can be boiled down to

$$f = Au + \eta. \tag{1}$$

From the view of the numerical computation, reconstructing u from f is an ill-posed problem since the operator A depends on imaging devices or data acquisition patterns, which usually leads to a large and ill-conditioned matrix. So some variational-PDE based models have been proposed to overcome these drawbacks.

In order to improve the scanning time of the variational-PDE based models, motivated by the compressed sensing (CS) theory, Lustig et al. [29] proposed an unconstrained model to reconstruct CS-MRI images as follows:

$$\min_u \|u\|_{TV} + \tau \left\|\Phi^\top u\right\|_1 + \frac{\eta}{2}\|Au - f\|_2^2, \tag{2}$$

Received by the editors July 8, 2018 and, in revised form, March 11, 2020.
2000 *Mathematics Subject Classification.* 90C25, 49M27, 68U10, 94A08.
* Corresponding author.

where $\|u\|_{TV} = \|\nabla u\|_1$ is the total variation [5, 7, 24, 35, 43, 48] (TV), Φ is the wavelet transformation, the superscript $\top$ denotes (conjugate) transpose of matrix. $\|\Phi^\top u\|_1$ is the ℓ_1-norm of the representation of u under the wavelet transformation Φ, $\tau > 0$ is a scalar which balances Φ sparsity and TV sparsity.

As we know, the TV-based regularization in the model (2) can handle few-views problems in the MRI reconstruction, which has the advantage to preserve edges when removing noises in homogeneous regions. However, it usually tends to cause staircase-like artifacts [22, 26, 28, 32, 35] due to their nature of favoring piecewise constant solutions. To alleviate the above drawbacks, the total generalized variation (TGV) in [2] has attracted much interest in image science. On the other hand, the continuous wavelet transformation of a distribution f decays rapidly near the points where f is smooth, while it decays slowly near the irregular points. This property allows the identification of the singular support of f. However, the continuous wavelet transformation is unable to describe the geometry of the singularity set of f and, in particular, to identify the wavefront set of a distribution [40]. Unlike the traditional wavelets used in the second regularized term of (2) lacking the ability to detect directionality, the shearlets provide a multidirectional as well as a multiscale decomposition for multi-dimension signals [17, 18]. There are two main advantages of using shearlets regularization in reconstruction: one is that shearlets allow for a lower redundant sparse tight frame representation than other related multiresolution representations, while still offering shift invariance and a directional analysis; another is that the shearlet representation can be used to decompose the space $L^2(\Omega)$ of images into a sequence of spaces, while we apply the soft thresholding operator conveniently to numerical algorithm. Obviously, shearlets are better candidates than wavelets, as shearlets have essentially optimal approximation errors for images that contain edges apart from discontinuities along curves. So following these observations, Guo et al. [19] coupled the TGV with the shearlet transformation to reconstruct high quality images from incomplete compressive sensing measurements as

$$\min_u \mathrm{TGV}_\alpha^2(u) + \beta \sum_{j=1}^{N} \|\mathcal{SH}_j(u)\|_1 + \frac{\nu}{2}\|Au - f\|_2^2, \tag{3}$$

where $\mathcal{SH}_j(u)$ is the jth subband of the shearlet transformation of u; $\beta > 0$ balances the shearlet transformation sparsity and the TGV sparsity; $\nu > 0$ is the regularization parameter.

In the model (3), the key is how to balance two parameters β and ν. In form, an improperly large weight for the data fidelity term results in serious residual artifacts, whereas an improperly small weight results in damaged edges and fine structures [8]. To overcome these drawbacks, it needs to turn to the following constrained optimization model as

$$\begin{aligned} \min_u \quad & \mathrm{TGV}_\alpha^2(u) + \beta \sum_{j=1}^{N} \|\mathcal{SH}_j(u)\|_1 \\ \text{s.t.} \quad & \|Au - f\|_2 \le \sigma, \end{aligned} \tag{4}$$

where σ implies some prior information of noise. Compared with unconstrained model (3), the model (4) can estimate the noise level σ more easily than finding a suitable parameter ν. These two models are equivalent in nature when choosing suitable penalty parameter ν. In fact, this equivalency transformation has been successfully applied to imaging and sparsity tasks [37, 41, 42, 45] for the

TV-based models. For example, Ng et al. [31] considered the constrained TV image restoration and reconstruction problem. Furthermore, researchers reformulated many restoration and reconstruction problems into linearly constrained convex programming models while utilizing inexact versions of the ADMM [46]. By imposing box constraints on TV-based models and solving the resulting constrained models, more accurate solutions can be guaranteed in [4].

From the viewpoint of numerical implementations, some difference operators included in the ℓ_1-norm lead the model (4) to be the non-smooth constrained problem. So it is difficult to find a direct method to solve it. One of popular approaches is to use the alternating direction method of multipliers (ADMM). It can be traced back to the alternating direction implicit techniques for solving elliptic and parabolic partial differential equations developed in [11, 14]. Due to its dual decomposability and strong convergence properties, the ADMM has recently been used in various areas of scientific computing including optimization, image processing and machine learning [6]. This method is also related to other methods such as the splitting-Bregman (SB) [15] method and the Douglas-Rachford method [9]. As a summarization, the contributions of this paper are as follows:

- Firstly, we propose a new constrained formulation of the TGV-shearlet based on the CS-MRI reconstruction via allowing easy parameter setting, good quantity and visual evaluation;
- Secondly, an effective ADMM scheme is proposed to directly optimize the constrained objective function with a fast and stable convergence result. Overall, the proposed method exhibits reasonable performance and outperforms the recent unconstrained counterparts when being applied to the MRI reconstruction.

The organization of the paper is as follows. We introduce the constrained TGV-Shearlet based MRI reconstruction model, propose the ADMM scheme for our constrained model, and give its convergence results in Section 2. Section 3 is devoted to implementation details of numerical experiments, followed by some conclusions in Section 4.

2. ADMM scheme for constrained TGV-shearlet based CS-MRI reconstruction

Here we need to give some overviews of the TGV regularization in order to conveniently explain the numerical method for the proposed model. As aforementioned, the TGV regularization can automatically balance first-order and high-order derivatives instead of using any fixed combination [44]. Hence, this process can yield visually pleasant results in images with piecewise polynomial intensities and sharp edges without staircase effects.

Definition 2.1. *For $k = 2$ and $\alpha > 0$, we see that*

(5)

$$\mathrm{TGV}_{\alpha}^{2}(u) = \sup\left\{ \int_{\Omega} u \mathrm{div}^{2} \boldsymbol{w} dx \,|\, w \in \mathcal{C}_{c}^{2}(\Omega, S^{d\times d}), \|\boldsymbol{w}\|_{\infty} \leq \alpha_0, \|\mathrm{div}\boldsymbol{w}\|_{\infty} \leq \alpha_1 \right\}.$$

In order to efficiently solve the second order $\mathrm{TGV}_{\alpha}^{2}$ based models in terms of ℓ_1 minimization, we need to derive another form of $\mathrm{TGV}_{\alpha}^{2}(u)$. For the notational convenience, we define spaces U, V, W as

$$U = C_c^2(\Omega, \mathbb{R}), \quad V = C_c^2(\Omega, \mathbb{R}^2) \text{ and } W = C_c^2(\Omega, S^{2\times 2})$$

and set $\mathbf{v} := (v_1, v_2)^{\top} \in V$ and $\mathbf{w} := (w_{11}, w_{12}; w_{21}, w_{22}) \in W$.

As one interposition, we replace the domain Ω by a discrete rectangular unit-length grid

$$\Omega = \{(i,j) : i,j \in \mathbb{N}, 1 \le i \le N_1, 1 \le j \le N_2\},$$

where $N_1, N_2 \ge 2$ are the image size. The scalar products can be defined in spaces V and W as

$$\begin{aligned} \mathbf{v},\mathbf{p} \in V : \langle \mathbf{v},\mathbf{p}\rangle_{\mathbf{v}} &= \langle v_1, p_1\rangle + \langle v_2, p_2\rangle, \\ \mathbf{w},\mathbf{q} \in W : \langle \mathbf{w},\mathbf{q}\rangle_{\mathbf{w}} &= \langle w_{11}, q_{11}\rangle + \langle w_{22}, q_{22}\rangle + 2\langle w_{12}, q_{12}\rangle. \end{aligned}$$

In order to discretize the TGV functional [10, 36], the forward and backward partial differentiation operators are introduced as:

$$(\nabla_x^+ u)_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j}, & \text{if } 1 \le i < N_1, \\ 0, & \text{if } i = N_1, \end{cases}$$
$$(\nabla_y^+ u)_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j}, & \text{if } 1 \le j < N_2, \\ 0, & \text{if } j = N_2, \end{cases}$$

as well as

$$(\nabla_x^- u)_{i,j} = \begin{cases} u_{1,j}, & \text{if } i = 1, \\ u_{i,j} - u_{i-1,j}, & \text{if } 1 < i < N_1, \\ -u_{N_1-1,j}, & \text{if } i = N_1, \end{cases}$$
$$(\nabla_y^- u)_{i,j} = \begin{cases} u_{i,1}, & \text{if } j = 1, \\ u_{i,j} - u_{i,j-1}, & \text{if } 1 < j < N_2, \\ -u_{i,N_2-1}, & \text{if } j = N_2. \end{cases}$$

Then, the gradient as well as symmetrised gradient can be expressed as

$$\begin{aligned} \nabla : U \to V, \quad & \nabla u = \begin{bmatrix} \nabla_x^+ u \\ \nabla_y^+ u \end{bmatrix}, \\ \varepsilon : V \to W, \quad & \varepsilon(\mathbf{v}) = \begin{bmatrix} \nabla_x^- v_1 & \frac{1}{2}(\nabla_y^- v_1 + \nabla_x^- v_2) \\ \frac{1}{2}(\nabla_y^- v_1 + \nabla_x^- v_2) & \nabla_y^- v_2 \end{bmatrix}, \\ \mathrm{div} : V \to U, \quad & \mathrm{div}\mathbf{v} = (\nabla_x^-)^\top v_1 + (\nabla_y^-)^\top v_2, \\ \widetilde{\mathrm{div}} : W \to V, \quad & \widetilde{\mathrm{div}}\mathbf{w} = \begin{bmatrix} (\nabla_x^+)^\top w_{11} + (\nabla_y^+)^\top w_{12} \\ (\nabla_x^+)^\top w_{12} + (\nabla_y^+)^\top w_{22} \end{bmatrix}. \end{aligned}$$

Let the variable $\mathbf{r} = \mathrm{div}\mathbf{w}$ in (5), then the discretization TGV_α^2 can be rewritten as:

$$\mathrm{TGV}_\alpha^2(u) = \max_{u,\mathbf{r},\mathbf{w}} \left\{ \langle u, \mathrm{div}\mathbf{r}\rangle \,|\, \widetilde{\mathrm{div}}\mathbf{w} = \mathbf{r}, \|\mathbf{w}\|_\infty \le \alpha_0, \|r\|_\infty \le \alpha_1 \right\}.$$

Through some computations, the $\mathrm{TGV}_\alpha^2(u)$ can be further reformulated as [19]:

$$\mathrm{TGV}_\alpha^2(u) = \min_{\mathbf{p}} \alpha_1 \|\nabla u - \mathbf{p}\|_1 + \alpha_0 \|\varepsilon(\mathbf{p})\|_1,$$

here $\|\cdot\|_1$ denotes the ℓ_1-norm [25]. Based on the above preparations, we now consider the discretization form of the model (4). This form is still a non-smooth optimization problem [47]. Furthermore, the gradient operator included in the ℓ_1-norm and the constrained terms lead to more numerical difficulties. So we need to introduce some auxiliary variables to split the original problem into some easily solvable subproblems. Formally, we introduce some new variables v, h_j, $\mathbf{w} = (w_1, w_2) \in V$

and $\mathbf{z} = \begin{bmatrix} z_1 & z_3 \\ z_3 & z_2 \end{bmatrix} \in W$ and define the convex set $K = \{v \in \mathbb{R}^N, \|v\|_2 \leq \sigma\}$, this discretization form can be written to another equivalent form as

$$
\begin{aligned}
\min_{u,\mathbf{p},\mathbf{w},\mathbf{z},h_j,v} \quad & \alpha_1\|\mathbf{w}\|_1 + \alpha_0\|\mathbf{z}\|_1 + \beta\sum_{j=1}^{N}\|h_j\|_1 \\
\text{s.t.} \quad & \mathbf{w} = \nabla u - \mathbf{p}, \\
& \mathbf{z} = \varepsilon(\mathbf{p}), \\
& h_j = \mathcal{SH}_j(u), \\
& v = f - Au, \\
& v \in K,
\end{aligned} \tag{6}
$$

where $\nabla u = (\nabla_x u, \nabla_y u)^\top$. Let χ_K denote the following indicator function of K:

$$
\chi_K(x) = \begin{cases} 0, & \text{if } x \in K, \\ \infty, & \text{if otherwise.} \end{cases}
$$

The augmented Lagrangian function of the problem (6) is given by

$$
\begin{aligned}
\mathcal{L}(\mathbf{w},\mathbf{z},h_j,v,u,\mathbf{p},\boldsymbol{\lambda}) \quad = \quad & \alpha_1\|\mathbf{w}\|_1 + \alpha_0\|\mathbf{z}\|_1 + \beta\sum_{j=1}^{N}\|h_j\|_1 + \chi_K(v) + \langle\boldsymbol{\lambda}_1, \nabla u - \mathbf{p} - \mathbf{w}\rangle \\
& + \langle\boldsymbol{\lambda}_2, \varepsilon(\mathbf{p}) - \mathbf{z}\rangle + \sum_{j=1}^{N}\langle\lambda_3, \mathcal{SH}_j(u) - h_j\rangle + \langle\lambda_4, Au + v - f\rangle \\
& + \frac{\mu_1}{2}\|\nabla u - \mathbf{p} - \mathbf{w}\|_2^2 + \frac{\mu_2}{2}\|\varepsilon(\mathbf{p}) - \mathbf{z}\|_2^2 + \frac{\mu_3}{2}\sum_{j=1}^{N}\|\mathcal{SH}_j(u) - h_j\|_2^2 \\
& + \frac{\mu_4}{2}\|Au + v - f\|_2^2,
\end{aligned} \tag{7}
$$

where $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \lambda_3, \lambda_4)$ are the Lagrangian multipliers [1, 16, 34], $\mu_i\,(i = 1,2,3,4)$ are the penalty parameters.

The augmented Lagrangian method (ALM) for (6) is an iterative algorithm based on the iteration

$$
\begin{cases}
\left(\mathbf{w}^{k+1}, \mathbf{z}^{k+1}, h_j^{k+1}, v^{k+1}, u^{k+1}, \mathbf{p}^{k+1}\right) \in \underset{\mathbf{w},\mathbf{z},h_j,v,u,\mathbf{p}}{\operatorname{argmin}} \mathcal{L}\left(\mathbf{w},\mathbf{z},x_j,v,u,\mathbf{p};\boldsymbol{\lambda}_1^k,\boldsymbol{\lambda}_2^k,\lambda_3^k,\lambda_4^k\right), \\
\boldsymbol{\lambda}_1^{k+1} = \boldsymbol{\lambda}_1^k + \theta\mu_1\left(\nabla u^{k+1} - \mathbf{p}^{k+1} - \mathbf{w}^{k+1}\right), \\
\boldsymbol{\lambda}_2^{k+1} = \boldsymbol{\lambda}_2^k + \theta\mu_2\left(\varepsilon(\mathbf{p}^{k+1}) - \mathbf{z}^{k+1}\right), \\
\lambda_3^{k+1} = \lambda_3^k + \theta\mu_3\left(\mathcal{SH}_j\left(u^{k+1}\right) - h_j^{k+1}\right), \\
\lambda_4^{k+1} = \lambda_4^k + \theta\mu_4\left(Au^{k+1} + v^{k+1} - f\right).
\end{cases} \tag{8}
$$

To guarantee the convergence of the ALM, the minimization of each subproblem needs to be solved to certain high accuracy before the iterative updates of multipliers. In contrast, it is easier to minimize with respect to $\mathbf{w}, \mathbf{z}, h_j, v, u$ and $\mathbf{p}$ each, which can be grouped into two blocks $\mathbf{m} = \{\mathbf{w}, \mathbf{z}, h_j, v\}$ and $\mathbf{n} = \{u, \mathbf{p}\}$. For a fixed $\mathbf{n}$, the minimization with respect to $\mathbf{w}, \mathbf{z}, h_j$ and v can be carried out in parallel because all $\mathbf{w}, \mathbf{z}, h_j$ and v can be separated from one another in (7). The basic idea of the ADMM dates back to the work by Glowinski and Marocco [14], Gabay and Mercier [11], where they proposed the method by utilizing the separable structure of variables [13]. In the following, we consider how to solve these two blocks.

2.1. m-subproblem. Here we focus on minimizing $\mathcal{L}((\mathbf{w},\mathbf{z},h_j,v),(u,\mathbf{p}),\boldsymbol{\lambda})$ w.r.t. $(\mathbf{w},\mathbf{z},h_j,v)$. This writes

$$\mathbf{w}^{k+1} \in \underset{\mathbf{w}}{\operatorname{argmin}}\, \alpha_1\|\mathbf{w}\|_1 - \langle \boldsymbol{\lambda}_1^k, \mathbf{w}\rangle + \frac{\mu_1}{2}\left\|\nabla u^k - \mathbf{p}^k - \mathbf{w}\right\|_2^2, \tag{9}$$

$$\mathbf{z}^{k+1} \in \underset{\mathbf{z}}{\operatorname{argmin}}\, \alpha_0\|\mathbf{z}\|_1 - \langle \boldsymbol{\lambda}_2^k, \mathbf{z}\rangle + \frac{\mu_2}{2}\|\varepsilon(\mathbf{p}^k) - \mathbf{z}\|_2^2, \tag{10}$$

$$h_j^{k+1} \in \underset{h_j}{\operatorname{argmin}}\, \beta\|h_j\|_1 - \langle \lambda_3^k, h_j\rangle + \frac{\mu_3}{2}\|\mathcal{SH}_j(u^k) - h_j\|_2^2, \tag{11}$$

$$v^{k+1} \in \underset{v}{\operatorname{argmin}}\, \chi_K(v) + \langle \lambda_4^k, v\rangle + \frac{\mu_4}{2}\|Au^k + v - f\|_2^2. \tag{12}$$

The subproblems about $\mathbf{w},\mathbf{z},h_j$ in (9)-(11) are similar to the ℓ_1-regularized least squares problems and the solutions can be explicitly obtained by using the Shrinkage operator.

For fixed u and $\boldsymbol{\lambda}_1$, the minimizer $\mathbf{w}$ is given by

$$\mathbf{w}^{k+1} = Shrink_2(\nabla u^k - \mathbf{p}^k + \boldsymbol{\lambda}_1^k/\mu_1, \alpha_1/\mu_1), \tag{13}$$

where $Shrink_2(\cdot,\alpha_1/\mu_1)$ is the generalized shrinkage and defined as

$$Shrink_2(\boldsymbol{\xi},\alpha_1/\mu_1) \triangleq \max\{\|\boldsymbol{\xi}\|_2 - \alpha_1/\mu_1, 0\}\cdot\frac{\boldsymbol{\xi}}{\|\boldsymbol{\xi}\|_2}.$$

Considering the Frobenius norm $\|\cdot\|_F$ of a matrix, we find that the $\mathbf{z}$ subproblem (10) has the following closed form solution:

$$\mathbf{z}^{k+1} = Shrink_F(\varepsilon(\mathbf{p}^k) + \boldsymbol{\lambda}_2^k/\mu_2, \alpha_0/\mu_2), \tag{14}$$

where $\mathbf{z}^{k+1} \in W$ and

$$Shrink_F(\boldsymbol{\zeta},\alpha_0/\mu_2) \triangleq \max\{\|\boldsymbol{\zeta}\|_F - \alpha_0/\mu_2, 0\}\cdot\frac{\boldsymbol{\zeta}}{\|\boldsymbol{\zeta}\|_F}.$$

For the h_j subproblem, we can directly solve (11) using shrinkage operator

$$h_j^{k+1} = Shrink(\mathcal{SH}_j(u^k) + \lambda_3^k/\mu_3, \beta/\mu_3), \tag{15}$$

where the shrinkage operator $Shrink(\cdot,\beta/\mu_3)$ is defined by

$$Shrink(\varpi,\beta/\mu_3) \triangleq \max\{|\varpi| - \beta/\mu_3, 0\}\cdot\frac{\varpi}{|\varpi|}.$$

We note that the computational costs for these three kinds of Shrinkage operators are linear in N, which are effective for the ℓ_1-norm based problem without including any operators. Numerically, the jth subband of the shearlet transformation of u can be implemented efficiently in frequency domain by the component-wise multiplication [20]

$$\mathcal{SH}_j(u) = \mathcal{F}^{-1}(\hat{H}_j .* \hat{u}) = \mathcal{F}^{-1}\mathrm{diag}(\hat{H}_j)\mathcal{F}u = M_j u,$$

where $\hat{u}$ denotes the 2D Fourier transformation of u and $\hat{H}_j$ is the frequency domain shearlet based on the jth subband. $.*$ denotes the component-wise multiplication operator. Besides, let $\mathcal{F}$ and $\mathcal{F}^{-1}$ be the 2D Fourier transform operator and its inverse operator, respectively. In the numerical aspect, we present the $\mathcal{SH}_j(u)$ as vectorized version. M_j is a block circulant matrix which can be diagonalized by 2D Fourier transform under the periodic boundary condition, with "diag" being the diagonal operator [21].

Considering the inequality constraint $\|v\|_2 \leq \sigma$ in (12), the subproblem v is equivalently transformed to

$$\begin{aligned} \min_{v} \quad & \langle \lambda_4, v\rangle + \tfrac{\mu_4}{2}\|Au + v - f\|_2^2 \\ \text{s.t.} \quad & \|v\|_2 \leq \sigma. \end{aligned} \tag{16}$$

This is a constrained least square problem. Due to the structure of the above problem, this operation is a projection onto a ℓ_2-ball [42]. The subproblem v can be solved explicitly or very accurately via the projection operator

$$v^{k+1} = \min\left\{1, \frac{\sigma}{\|f - Au^k - \lambda_4^k/\mu_4\|_2}\right\} \cdot (f - Au^k - \lambda_4^k/\mu_4). \tag{17}$$

2.2. n-subproblem. Secondly, we try to minimize $\mathcal{L}((\mathbf{w}, \mathbf{z}, h_j, v), (u, \mathbf{p}), \lambda)$ w.r.t. $(u, \mathbf{p})$. The solutions $(u^{k+1}, \mathbf{p}^{k+1})$ satisfy the following minimization problem

$$\begin{aligned} (u^{k+1}, \mathbf{p}^{k+1}) \in \arg\min_{u,\mathbf{p}} & \langle \boldsymbol{\lambda}_1^k, \nabla u - \mathbf{p}\rangle + \langle \boldsymbol{\lambda}_2^k, \varepsilon(\mathbf{p})\rangle + \sum_{j=1}^{N} \langle \lambda_3^k, \mathcal{SH}_j(u)\rangle + \langle \lambda_4^k, Au\rangle \\ & + \frac{\mu_1}{2}\|\nabla u - \mathbf{p} - \mathbf{w}^{k+1}\|_2^2 + \frac{\mu_2}{2}\|\varepsilon(\mathbf{p}) - \mathbf{z}^{k+1}\|_2^2 \\ & + \frac{\mu_3}{2}\sum_{j=1}^{N} \|\mathcal{SH}_j(u) - h_j^{k+1}\|_2^2 + \frac{\mu_4}{2}\|Au + v^{k+1} - f\|_2^2. \end{aligned} \tag{18}$$

The minimizations with respect to $(u, \mathbf{p})$ should be simultaneously performed since updating u and $\mathbf{p}$ are coupled to each other. Then, $\mathbf{n}^{k+1} = (u^{k+1}, p_1^{k+1}, p_2^{k+1})^\top$ is the solution of a linear system:

$$\mathbf{B}\mathbf{n}^{k+1} = \mathbf{b}, \tag{19}$$

where

$$\mathbf{B} = \begin{bmatrix} B_1 & -\mu_1(\nabla_x^-)^\top & -\mu_1(\nabla_y^-)^\top \\ -\mu_1\nabla_x^+ & B_2 & \frac{1}{2}(\nabla_y^+)^\top\nabla_x^- \\ -\mu_1\nabla_y^+ & \frac{1}{2}(\nabla_x^+)^\top\nabla_y^- & B_3 \end{bmatrix}$$

and

$$\mathbf{b} = \begin{bmatrix} \sum_{j=1}^{N} M_j^* \left(\mu_3 h_j^{k+1} - \lambda_3^k\right) + \nabla^\top(\mu_1 w^{k+1} - \boldsymbol{\lambda}_1^k) - A^\top\lambda_4^k - \mu_4 A^\top(v^{k+1} - f) \\ (\lambda_1^k)_1 - \mu_1 w_1^{k+1} + (\nabla_x^+)^\top\left(\mu_2 z_1^{k+1} - (\lambda_2^k)_1\right) + (\nabla_y^+)^\top\left(\mu_2 z_3^{k+1} - (\lambda_2^k)_3\right) \\ (\lambda_1^k)_2 - \mu_1 w_2^{k+1} + (\nabla_y^+)^\top\left(\mu_2 z_2^{k+1} - (\lambda_2^k)_2\right) + (\nabla_x^+)^\top\left(\mu_2 z_3^{k+1} - (\lambda_2^k)_3\right) \end{bmatrix},$$

where $B_1 = \mu_3\sum_{j=1}^{N} M_j^* M_j + \mu_1\nabla^\top\nabla + \mu_4 A^\top A$, $B_2 = \mu_1 I + \mu_2(\nabla_x^+)^\top\nabla_x^- + \frac{1}{2}(\nabla_y^+)^\top\nabla_y^-$ and $B_3 = \mu_1 I + \mu_2(\nabla_y^+)^\top(\nabla_y^-) + \frac{1}{2}(\nabla_x^+)^\top\nabla_x^-$.

As we know, the coefficient matrix in (19) is block circular under the periodic boundary conditions for u, so it can be diagonalized by 2-D Fourier transforms $\mathcal{F}$. Considering $\mathcal{F}^\top\mathcal{F} = I$ and multiplying by $\mathcal{F}$ on both sides of (19), we obtain

$$\hat{\mathbf{B}}\mathcal{F}n^{k+1} = \mathcal{F}\mathbf{b}, \tag{20}$$

where $\hat{\mathbf{B}} = \mathcal{F}\mathbf{B}\mathcal{F}^\top$ is the blockwise diagonal. Then the equation (19) can be solved efficiently using Cramer's rule. Accordingly, the closed-form solutions are given by

$$n^{k+1} = \mathcal{F}^\top\hat{\mathbf{B}}^{-1}\mathcal{F}\mathbf{b}. \tag{21}$$

Remark 1. *In Algorithm 1, we choose the step length of* $\theta \in (0, (\sqrt{5}+1)/2)$. *The convergence of the ADMM with the step length* $\theta \in (0, (\sqrt{5}+1)/2)$ *was first established in* [12] *in the context of variational inequality, which covers the proof of the following theorem.*

Algorithm 1: ADMM for constrained TGV-shearlet based CS-MRI reconstruction

Input selection matrix P, data f. Choose model parameters α_1, α_0, $\mu_i (i = 1, \cdots, 4) > 0$, and $\theta \in (0, (\sqrt{5}+1)/2)$. Initialization: u^0, $(\lambda_i)^0 (i = 1, \cdots, 4)$, $p_i^0 (i = 1, 2)$. Set $k = 0$.

1. $\mathbf{w}^{k+1}$ is given by (13),
2. $\mathbf{z}^{k+1}$ is given by (14),
3. h_j^{k+1} is given by (15),
4. v is given by (17),
5. $u^{k+1}, p_1^{k+1}, p_2^{k+1}$ are given by (21),
6. update Lagrangian multipliers by
$$\boldsymbol{\lambda}_1^{k+1} = \boldsymbol{\lambda}_1^k + \theta\mu_1(\nabla u^{k+1} - \mathbf{p}^{k+1} - \mathbf{w}^{k+1}),$$
$$\boldsymbol{\lambda}_2^{k+1} = \boldsymbol{\lambda}_2^k + \theta\mu_2(\varepsilon(\mathbf{p}^{k+1}) - \mathbf{z}^{k+1}),$$
$$\lambda_3^{k+1} = \lambda_3^k + \theta\mu_3(\mathcal{SH}_j(u^{k+1}) - h_j^{k+1}), (j = 1, \cdots, N)$$
$$\lambda_4^{k+1} = \lambda_4^k + \theta\mu_4(Au^{k+1} + v^{k+1} - f).$$
7. Stop if it is convergent; otherwise, set $k := k + 1$ and go to step 1.

Theorem 2.1. *For any μ_i $(i = 1, 2, 3, 4) > 0$ and $\theta \in (0, (\sqrt{5}+1)/2)$, the sequence $\left(\boldsymbol{w}^k, \boldsymbol{z}^k, h^k, v^k, u^k\right)$ generated by Algorithm 1 from any starting point $\left(\boldsymbol{w}^0, \boldsymbol{z}^0, h_j^0, v^0, u^0\right)$ converges to a solution of the problem (6).*

3. Experimental results

In this section, to show the improvement on our ADMM scheme for constrained TGV and shearlet transformation based on the CS-MRI reconstruction, we employ Algorithm 1 to some real clinical MRI reconstruction and compare it with the state-of-the-art TV wavelet transform ("UTVW" for short) [29], the unconstrained TGV shearlet model based results ("UTGVS" for short) proposed in [19]. Two in vivo MR images and one MR angiography image used in our experiment are shown in Fig. 1: brain MR image of size 256×256, foot MR image of size 512×512, and brain magnetic resonance angiography (BMRA) of size 512×512.

(a) Original brain MR image (b) Original foot MR image (c) Original BMRA image

FIGURE 1. Original images.

We perform the experiment under Matlab R2014a and Windows 10(x64) on a PC with an Intel Core (TM) i7-4712MQ CPU at 2.3GHz and 6.0GB of memory. Image reconstruction results will be analyzed and evaluated by quantitative and visual evaluation methods. The quantitative evaluation of the reconstructed image

is evaluated via the relative error (RE) and signal-to-noise ratio (SNR), which are defined as:

$$\mathrm{RE} = \frac{\|u - u_{true}\|_2^2}{\|u_{true}\|_2^2}$$

and

$$\mathrm{SNR} = 10 \log_{10} \frac{\|u_{true}\|_2^2}{\|u - u_{true}\|_2^2},$$

where u is the reconstructed image and u_{true} is the original image. Theoretically, the smaller RE and the larger SNR values normally indicate better performance in image reconstruction.

In our algorithm, $A = P\mathcal{F}$ consists of p rows of the $n \times n$ matrix corresponding to the full 2D discrete Fourier transform, where $p \ll n$. The p selected rows specify the selected frequencies at which the measurements in f are collected. The smaller the p, the lesser the amount of time required for an MR scanner to acquire f. The sampling ratio is defined to be p/n. The scanning duration is shorter if the sampling ratio is smaller. In MRI, we have certain freedom to select the rows, which correspond to certain frequencies. In our experiments, for simplicity, we use pseudo radial mask which can simulate randomness in acquisition for demonstration. Besides, we use the three-scale shearlet transformation which possesses 29 subbands: one for low frequency and 28 for high frequency, in the proposed algorithm. The stopping criterion used in the experiment consists of checking that the RE is less than 10^{-5}.

For the optimized selection of parameters, $\mu_i\,(i = 1, 2, 3, 4)$ are used to balance the data fidelity and two regularization terms, the values of them should be set in accordance with both the noise level in the observation and the sparsity level of the underlying image. Usually, the higher the noise level is, the smaller μ_i should be, where μ_1, μ_2 are often smaller than μ_3, μ_4, respectively. Actually, they are often empirically selected by visual inspection. Based on our experience, a simple way to choose them is to try different values and compare the recovered images, while $\mu_1, \mu_2 \in (10^{-3}, 10^{-5})$ and $\mu_3, \mu_4 \in (10, 10^4)$. The positive weights α_0, α_1 and β are used to balance the first, second derivatives and shearlet transformation. Proper values of them should be chosen based on sparsity feature of the underlying image. Generally, $\alpha_0, \alpha_1, \beta$ are chosen nearby 10^{-3} and 10^{-2} respectively in our experiments.

3.1. Noise level estimation. For a given discrete image, this parameter σ can be estimated as follows: Under the assumption that the noise is additive Gaussian white noise, we can estimate σ according to the image noise level η. In case of not having prior knowledge on noise level η, we need to estimate the noise level from image signals based on specific image characteristics. Generally, the estimation methods are classifiable into filter-based approaches, patch-based approaches, or some combination of them. In filter-based approaches [38, 39], the noisy image is firstly filtered using a high-pass filter to suppress the image structures. Then the noise variance is computed from the difference between the noisy image and the filtered image. The main difficulty of filter-based approaches is that the difference between the two images is assumed to be the noise, but this assumption is not always true, especially for images with complex structures or fine details. On the other hand, in patch-based approaches [27, 30, 33], images are decomposed into a number of patches. We can consider an image patch as a rectangular window in the image with a size of $N \times N$. The patch with the smallest standard deviation among decomposed patches has the least change of intensity. The intensity variation of

a homogenous patch is mainly caused by noise. In this paper, we try to capture the noise level by using patches-based principal component analysis (PCA) [23, 33] (e.g. the noise variance can be estimated as the smallest eigenvalue of the image block covariance matrix). In the following experiments, we estimated σ from the observed image, as discussed here.

3.2. Results of brain image reconstruction. In this example, we firstly test the brain MR image and select 18.72% (45 radial sampling lines) samples uniformly at random (as shown in Fig. 2 (b)). According to quantitative and visual quality evaluations, we manually select the optimal parameters $\sigma = 0.01$, $\beta = 10^{-2}$, $(\mu_1, \mu_2, \mu_3, \mu_4) = (10^{-3}, 10^{-5}, 10^3, 10^4)$, $(\alpha_0, \alpha_1) = (8 \times 10^{-4}, 10^{-3})$.

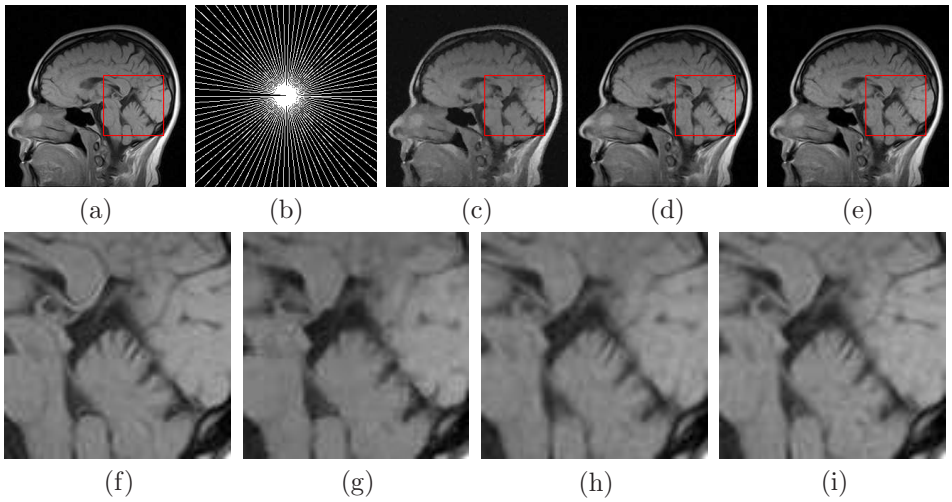
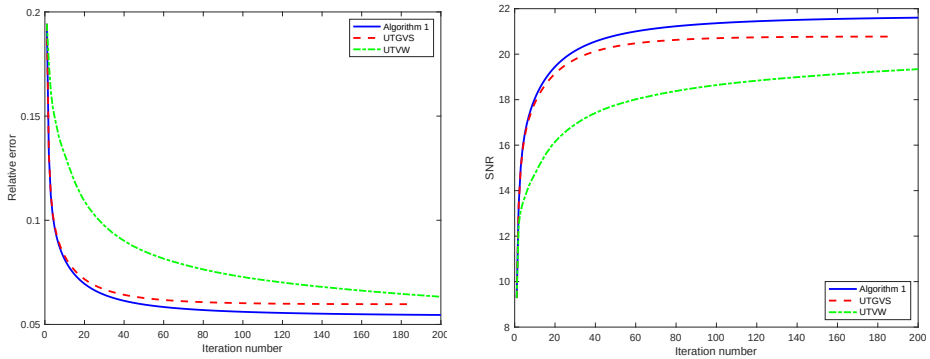


FIGURE 2. Reconstruction results for MR brain image from 18.72% spectral data. (a) Original brain; (b) Sampling mask; (c) UTVW result (RE=6.44%, SNR=19.93); (d) UTGVS result (RE=5.97%, SNR=20.77); (e) Algorithm 1 result (RE=5.45%, SNR=21.60); (f) Close-up view of (a); (g) Close-up view of (c); (h) Close-up view of (d); (i) Close-up view of (e).

Fig. 2 (c)-(e) show the visual comparisons of the reconstructed results by different methods in the brain image. Meantime, the local magnification views shown in Fig. 2 (g)-(i) illustrate that our proposed method produces a more natural-looking reconstruction with more regular structures compared with the UTVW and UTGVS. We plot the changes of RE and SNR with respect to iterations in Fig. 3. It shows that our constrained scheme in Algorithm 1 can reach a relatively lower relative error and higher SNR within fewer iterations, and the UTVW and UTGVS are slightly less efficient than our results. Especially, we tested different levels of sample ratios. The comparison results are presented, where the RE and the SNR of the recovered images to the true images are given in Table 1. These results and observations clearly demonstrate the efficiency and stability of our algorithm.

3.3. Results of foot image Reconstruction. In the second experiment, we test the foot MR image. The variable density sampling pattern, shown in Fig. 4 (b), is utilized to the k-space data of undersampling ratio 10.83% (51 radial sampling lines). Admissible reconstruction needs more projections than noiseless



(e1) (e2)

FIGURE 3. Convergence curves of MR brain image (18.72% spectral data): the y-axis is the relative error and SNR, the x-axis is the number of iteration.

TABLE 1. The values of MR brain image reconstructed from different sample ratios.

Sample ratio	Algorithm	*RE*	*SNR*
12.65%	UTVW	8.55%	17.31
	UTGVS	8.26%	17.89
	Algorithm 1	**7.77%**	**18.49**
14.65%	UTVW	8.14%	18.06
	UTGVS	7.62%	18.61
	Algorithm 1	**7.07%**	**19.33**
17.12%	UTVW	6.88%	19.38
	UTGVS	6.34%	20.25
	Algorithm 1	**5.91%**	**20.90**
18.76%	UTVW	6.11%	20.69
	UTGVS	5.75%	21.11
	Algorithm 1	**5.03%**	**21.83**
20.60%	UTVW	5.59%	21.25
	UTGVS	5.33%	21.79
	Algorithm 1	**4.94%**	**22.49**
21.76%	UTVW	5.42%	21.57
	UTGVS	5.13%	22.11
	Algorithm 1	**4.64%**	**23.02**
22.54%	UTVW	5.21%	22.09
	UTGVS	4.91%	22.51
	Algorithm 1	**4.43%**	**23.42**

cases because of inconsistencies in the data. To demonstrate the performance with additive noise using the proposed method, white Gaussian noise with the variance $\delta = 0.25$ is added into real and imaginary parts of original k-space data, respectively. We fix $\mu_1 = 10^{-3}, \mu_2 = 10^{-4}, \mu_3 = 10^3, \mu_4 = 10$, set $\alpha_0 = 7 \times 10^{-3}, \alpha_1 = 2 \times 10^{-3}$, and choose optimal $\beta = 1.5 \times 10^{-2}$. Generally, the higher the noise level is, the smaller μ_4 should be.

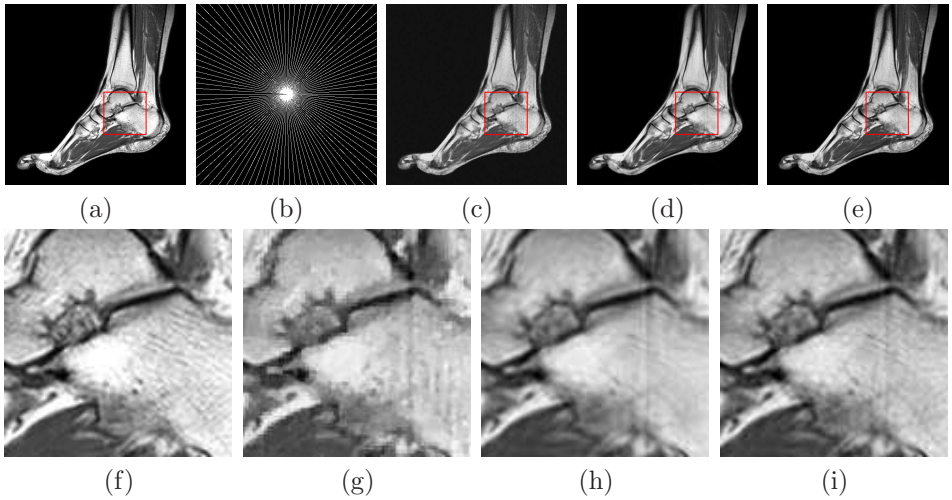


FIGURE 4. Reconstruction results for MR foot image from 10.83% spectral data with $\sigma = 0.25$. (a) Original foot; (b) Sampling mask; (c) UTVW result (RE=7.92%, SNR=20.06); (d) UTGVS result (RE=7.44%, SNR=20.98); (e) Algorithm 1 result (RE=6.13%, SNR=22.79); (f) Close-up view of (a); (g) Close-up view of (c); (h) Close-up view of (d); (i) Close-up view of (e).

The reconstructed images are shown in Fig. 4 (c)-(e), respectively. In Fig. 4 (g)-(i), we list the zoomed-in results of the red boxes in Fig. 4 (c)-(e). It is seen that Algorithm 1 can adequately reconstruct the image in the sense that most details and fine structures are accurately recovered from a small set of data samples compared with other strategies. In order to show the robustness of the proposed model, we test various noise levels in Table 2. For the noise-added image, better noise suppressing and sharper textures or edges are achieved using our proposed method than using others.

TABLE 2. The values of foot image reconstructed from different noise levels (Sample ratios=10.83%).

Noise level (σ)	Algorithm	RE	SNR
0.1	UTVW	7.09%	21.02
	UTGVS	6.64%	21.79
	Algorithm 1	**5.77%**	**23.20**
0.3	UTVW	8.39%	19.65
	UTGVS	7.45%	20.97
	Algorithm 1	**6.14%**	**22.77**
0.5	UTVW	9.51%	18.13
	UTGVS	8.29%	19.98
	Algorithm 1	**6.43%**	**22.38**

3.4. Results of BMRA Reconstruction. BMRA is a group of techniques based on the MRI to image blood vessels, which contains complex geometric structures,

limited spatial resolution and low image contrast. Furthermore, in comparison, we demonstrate the effectiveness of our proposed algorithm by testing one slice of BMRA image with multilevel structures. The sample ratio is set to be approximately 10.64% (50 radial sampling lines). Based on the selection by the authors and several tries, we choose moderate values, $\sigma = 0.01$, $\beta = 10^{-2}$, $(\mu_1, \mu_2, \mu_3, \mu_4) = (2 \times 10^{-3}, 10^{-5}, 10^3, 10^3)$, $(\alpha_0, \alpha_1) = (4 \times 10^{-4}, 10^{-3})$, which appear to give a best compromise between convergence speed and image quality of our constrained scheme.

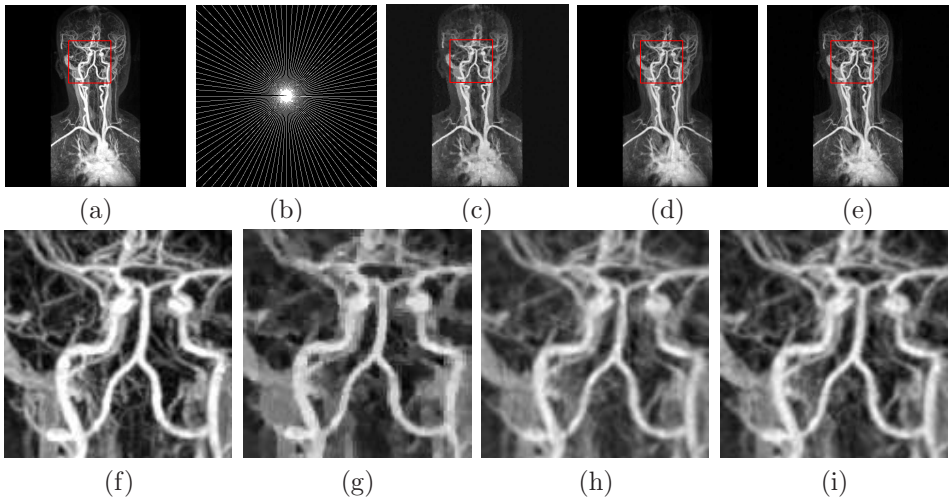
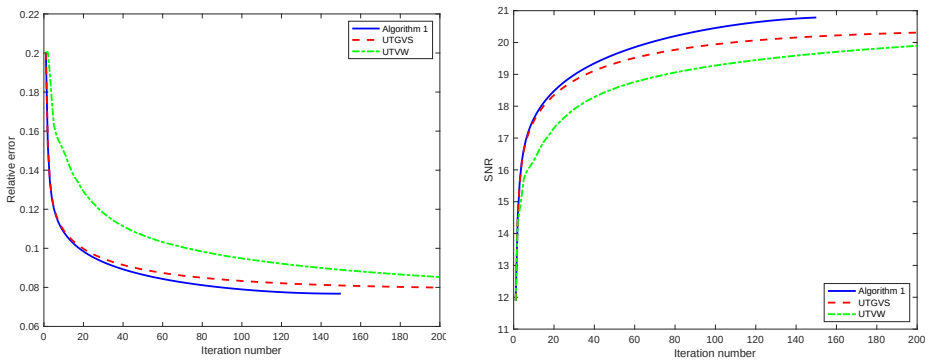


FIGURE 5. Reconstruction results for BMRA from 10.64% spectral data. (a) Original BMRA; (b) Sampling mask; (c) UTVW result (RE=8.37%, SNR=19.89); (d) UTGVS result (RE=7.99%, SNR=20.31); (e) Algorithm 1 result (RE=7.68%, SNR=20.79); (f) Close-up view of (a); (g) Close-up view of (c); (h) Close-up view of (d); (i) Close-up view of (e).

Here we compare the constrained and unconstrained models simultaneously for the BMRA image, as shown in Fig. 5 (c)-(e), and depict the corresponding zoomed-in regions in Fig. 5 (g)-(i). We observe that the unconstrained UTVW and UTGVS reconstruction results in patchy artifacts with some loss in fine details, while the results obtained by the constrained model are visibly better. For more details, the thin structures and junction parts are preserved due to the directional sensitivity of shearlets. In the meanwhile, the ringing artifacts along the geometric features are suppressed in our result from the contribution of the TGV. The comparison of RE and SNR are plotted in Fig. 6. The horizontal label is chosen as iteration number to show the convergence rate. These results indicate that Algorithm 1 performs better than the UTVW and UTGVS in the sense that it obtains better recovery results (smaller RE and higher SNR) within fewer iterations.

4. Conclusion

This paper proposed a constrained CS-MRI reconstruction model by combining the TGV and the shearlet transformation regularization. The constraint as the estimation of the prior noise level can efficiently penalize the data fitting term so we can obtain more robust reconstruction images than other state-of-the-art methods.



(e1) (e2)

FIGURE 6. Convergence curves of BMRA reconstruction (10.64% spectral data): the y-axis is the relative error and SNR, the x-axis is the number of iteration.

Since the proposed model is non-smooth, we use the ADMM and the projection scheme to transform it into several easy solvable subproblems and thereby fast convergence can be guaranteed. Our results in vivo results demonstrated that the proposed model can preserve more details and fine structures than the existing regularization methods while suppressing noise in applications of the MRI reconstruction.

Acknowledgments

The authors would like to thank two anonymous referees for their important and valuable comments, which led to great improvements of an early version of the article. This work is partially supported by the NSFC (Nos. 61971234, 11501301, 11401170, 61179027), the National Basic Research Program of China (973Program) (No. 2015CB856003), Hunan Provincial Key Laboratory of Mathematical Modeling and Analysis in Engineering (Changsha University of Science & Technology 2018MMAEYB03), the 1311 Talent Plan of NUPT, and the Natural Science Foundation of the Hunan Province of China (No. 2019JJ40323).

References

[1] D. P. BERTSEKAS, *Constrained optimization and Lagrange multiplier methods*, Academic press, 2014.

[2] K. BREDIES, K. KUNISCH, AND T. POCK, *Total generalized variation*, SIAM Journal on Imaging Sciences, 3 (2010), pp. 492–526.

[3] E. J. CANDÈS, J. ROMBERG, AND T. TAO, *Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information*, IEEE Transactions on Information Theory, 52 (2006), pp. 489–509.

[4] R. H. CHAN, M. TAO, AND X. YUAN, *Constrained total variation deblurring models and fast algorithms based on alternating direction method of multipliers*, SIAM Journal on imaging Sciences, 6 (2013), pp. 680–697.

[5] H. CHANG, Y. LOU, M. K. NG, AND T. ZENG, *Phase retrieval from incomplete magnitude information via total variation regularization*, SIAM Journal on Scientific Computing, 38 (2016), pp. A3672–A3695.

[6] C. CHEN, M. LI, X. LIU, AND Y. YE, *Extended admm and bcd for nonseparable convex minimization models with quadratic coupling terms: convergence analysis and insights*, Mathematical Programming, 173 (2019), pp. 37–77.

[7] L. CHEN AND T. ZENG, *A convex variational model for restoring blurred images with large rician noise*, Journal of Mathematical Imaging and Vision, 53 (2015), pp. 92–111.

[8] Y. Chen, X. Ye, and F. Huang, *A novel method and fast algorithm for mr image reconstruction with significantly under-sampled data*, Inverse Problems and Imaging, 4 (2010), pp. 223–240.
[9] J. Eckstein and D. P. Bertsekas, *On the douglas-rachford splitting method and the proximal point algorithm for maximal monotone operators*, Mathematical Programming, 55 (1992), pp. 293–318.
[10] W. Feng and H. Lei, *Single-image super-resolution with total generalised variation and shearlet regularisations*, IET, Image Processing, 8 (2014), pp. 833–845.
[11] D. Gabay and B. Mercier, *A dual algorithm for the solution of nonlinear variational problems via finite element approximation*, Computers & Mathematics with Applications, 2 (1976), pp. 17–40.
[12] R. Glowinski, *Numerical Methods for Nonlinear Variational Problems*, Springer, 1984.
[13] R. Glowinski, S. Luo, and X.-C. Tai, *Fast operator-splitting algorithms for variational imaging models: Some recent developments*, in Handbook of Numerical Analysis, vol. 20, Elsevier, 2019, pp. 191–232.
[14] R. Glowinski and A. Marroco, *Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de dirichlet non linéaires*, Revue française d'automatique, informatique, recherche opérationnelle. Analyse numérique, 9 (1975), pp. 41–76.
[15] T. Goldstein and S. Osher, *The split bregman method for l1-regularized problems*, SIAM journal on imaging sciences, 2 (2009), pp. 323–343.
[16] K. Guo, D. Han, and T. Wu, *Convergence of admm for optimization problems with nonseparable nonconvex objective and linear constraints*, Pacific Journal of Optimization, 14 (2018), pp. 489–506.
[17] K. Guo and D. Labate, *Optimally sparse multidimensional representation using shearlets*, SIAM journal on mathematical analysis, 39 (2007), pp. 298–318.
[18] K. Guo, D. Labate, and W.-Q. Lim, *Edge analysis and identification using the continuous shearlet transform*, Applied and Computational Harmonic Analysis, 27 (2009), pp. 24–46.
[19] W. Guo, J. Qin, and W. Yin, *A new detail-preserving regularization scheme*, SIAM Journal on Imaging Sciences, 7 (2014), pp. 1309–1334.
[20] S. Häuser and G. Steidl, *Convex multiclass segmentation with shearlet regularization*, International Journal of Computer Mathematics, 90 (2013), pp. 62–81.
[21] C. He, C.-H. Hu, and W. Zhang, *Adaptive shearlet-regularized image deblurring via alternating direction method*, in 2014 IEEE International Conference on Multimedia and Expo (ICME), IEEE, 2014, pp. 1–6.
[22] Z. Jia, M. K. Ng, and W. Wang, *Color image restoration by saturation-value total variation*, SIAM Journal on Imaging Sciences, 12 (2019), pp. 972–1000.
[23] B. Komander, D. A. Lorenz, and L. Vestweber, *Denoising of image gradients and total generalized variation denoising*, Journal of Mathematical Imaging and Vision, 61 (2019), pp. 21–39.
[24] Z. Li, F. Malgouyres, and T. Zeng, *Regularized non-local total variation and application in image restoration*, Journal of Mathematical Imaging and Vision, 59 (2017), pp. 296–317.
[25] J. Liu, F. Fang, and N. Du, *Color-to-gray conversion with perceptual preservation and dark channel prior*, INTERNATIONAL JOURNAL OF NUMERICAL ANALYSIS AND MODELING, 16 (2019), pp. 668–679.
[26] J. Liu, M. Yan, and T. Zeng, *Surface-aware blind image deblurring*, IEEE transactions on pattern analysis and machine intelligence, (2019).
[27] X. Liu, M. Tanaka, and M. Okutomi, *Single-image noise level estimation for blind denoising*, IEEE transactions on image processing, 22 (2013), pp. 5226–5237.
[28] S. Luo, Q. Lv, H. Chen, and J. Song, *Second-order total variation and primal-dual algorithm for ct image reconstruction.*, International Journal of Numerical Analysis & Modeling, 14 (2017).
[29] M. Lustig, D. Donoho, and J. M. Pauly, *Sparse mri: The application of compressed sensing for rapid mr imaging*, Magnetic resonance in medicine, 58 (2007), pp. 1182–1195.
[30] J. V. Manjón, P. Coupé, and A. Buades, *Mri noise estimation and denoising using non-local pca*, Medical image analysis, 22 (2015), pp. 35–47.
[31] M. K. Ng, P. Weiss, and X. Yuan, *Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods*, SIAM journal on Scientific Computing, 32 (2010), pp. 2710–2736.

[32] Z.-F. Pang, Y.-M. Zhou, T. Wu, and D.-J. Li, *Image denoising via a new anisotropic total-variation-based model*, Signal Processing: Image Communication, 74 (2019), pp. 140–152.
[33] S. Pyatykh, J. Hesser, and L. Zheng, *Image noise level estimation by principal component analysis*, IEEE Transactions on Image Processing, 22 (2013), pp. 687–699.
[34] R. T. Rockafellar, *Convex analysis*, Princeton university press, 2015.
[35] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D: Nonlinear Phenomena, 60 (1992), pp. 259–268.
[36] M.-G. Shama, T.-Z. Huang, J. Liu, and S. Wang, *A convex total generalized variation regularized model for multiplicative noise and blur removal*, Applied Mathematics and Computation, 276 (2016), pp. 109–121.
[37] L. Shen and S. Pan, *Inexact indefinite proximal admms for 2-block separable convex programs and applications to 4-block dnnsdps*, arXiv preprint arXiv:1505.04519, (2015).
[38] C. Sutour, C.-A. Deledalle, and J.-F. Aujol, *Estimation of the noise level function based on a nonparametric detection of homogeneous image regions*, SIAM Journal on Imaging Sciences, 8 (2015), pp. 2622–2661.
[39] S.-C. Tai and S.-M. Yang, *A fast method for image noise estimation using laplacian operator and adaptive edge detection*, in 3rd International Symposium on Communications, Control and Signal Processing, ISCCSP 2008., IEEE, 2008, pp. 1077–1081.
[40] B. Vandeghinste, B. Goossens, R. Van Holen, C. Vanhove, A. Pizurica, S. Vandenberghe, and S. Staelens, *Iterative ct reconstruction using shearlet-based regularization*, IEEE Transactions on Nuclear Science, 60 (2013), pp. 3305–3317.
[41] F. Wang, X.-L. Zhao, and M. K. Ng, *Multiplicative noise and blur removal by framelet decomposition and ℓ_1-based l-curve method*, IEEE Transactions on Image Processing, 25 (2016), pp. 4222–4232.
[42] P. Weiss, L. Blanc-Féraud, and G. Aubert, *Efficient schemes for total variation minimization under constraints in image processing*, SIAM journal on Scientific Computing, 31 (2009), pp. 2047–2080.
[43] Y. Wen, R. H. Chan, and T. Zeng, *Primal-dual algorithms for total variation based image restoration under poisson noise*, Science China Mathematics, 59 (2016), pp. 141–160.
[44] T. Wu and J. Shao, *Non-convex and convex coupling image segmentation via tgpv regularization and thresholding*, Advances in Applied Mathematics and Mechanics, (2020).
[45] T. Wu, D. Z. Wang, Z. Jin, and J. Zhang, *Solving constrained tv2l1-l2 mri signal reconstruction via an efficient alternating direction method of multipliers*, Numerical Mathematics: Theory, Methods and Applications, 10 (2017), pp. 895–912.
[46] T. Wu, W. Zhang, D. Z. Wang, and Y. Sun, *An efficient peaceman–rachford splitting method for constrained tgv-shearlet-based mri reconstruction*, Inverse Problems in Science and Engineering, 27 (2019), pp. 115–133.
[47] C. Zeng and C. Wu, *On the edge recovery property of noncovex nonsmooth regularization in image restoration*, SIAM Journal on Numerical Analysis, 56 (2018), pp. 1168–1182.
[48] W. Zhu, S. Shu, and L. Cheng, *An efficient proximity point algorithm for total-variation-based image restoration*, Advances in Applied Mathematics and Mechanics, 6 (2014), pp. 145–164.

[1] School of Science, Nanjing University of Posts and Telecommunications, Nanjing, 210023, China.
E-mail: wutt@njupt.edu.cn

[2] College of Mathematics and Statistics, Henan University, Kaifeng 475004, China.
E-mail: zhifengpang@163.com

[3] National Engineering Research Center of Communications and Networking, Nanjing University of Posts and Telecommunications, Nanjing 210023, China.
E-mail: wyg@njupt.edu.cn

[4] School of Computer Engineering and Applied Mathematics, Changsha University, Changsha 410003, China.
E-mail: yangyufei_hd@sina.com